

DOCTORAL THESIS

Smart Terrain Perception Using Hyperspectral Imaging

Dhanushka Chamara Liyanage

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
58/2021

Smart Terrain Perception Using Hyperspectral Imaging

DHANUSHKA CHAMARA LIYANAGE



Tallinn University of Technology
School of Engineering
Department of Electrical Power Engineering and Mechatronics
This dissertation was accepted for the defence of the degree on 14/11/2021

Supervisor: Prof. Mart Tamre
School of Engineering
Tallinn University of Technology
Tallinn, Estonia

Co-supervisor: Robert Hudjakov, Ph.D
School of Engineering
Tallinn University of Technology
Tallinn, Estonia

Opponents: Prof. Tomasz Piotr Kucner, Ph.D
Aalto University
Espoo, Finland

Prof. G. Anbarjafari, Ph.D
University of Tartu
Tartu, Estonia

Defence of the Thesis: 14/12/2021, Tallinn

Declaration:

Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology, has not been submitted for doctoral or equivalent academic degree.

Dhanushka Chamara Liyanage

.....
Signature

Copyright: Dhanushka Chamara Liyanage, 2021
ISSN 2585-6898 (publication)
ISBN 978-9949-83-765-6 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9949-83-766-3 (PDF)
Printed by Koopia Niini & Rauam

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
58/2021

Hüperspektraal-pilditehnika maastiku nutikaks tajumiseks

DHANUSHKA CHAMARA LIYANAGE



Contents

List of Publications	7
Author's Contribution to the Publications	8
Abbreviations	9
1 Introduction	10
1.1 Background and motivation	12
1.2 Problem statement and research objectives	13
1.3 Thesis contributions	15
1.4 Research methods.....	15
1.5 Thesis structure	16
2 Literature Review	17
2.1 Terrain perception	17
2.2 Semantic segmentation of terrains.....	18
2.2.1 Overview of semantic segmentation	18
2.2.2 Unstructured terrain semantic segmentation	19
2.2.3 Image annotation	20
2.3 Spectral imaging.....	20
2.4 Dimensionality reduction in hyperspectral images.....	21
2.5 Hyperspectral image visualisation	24
2.6 Hyperspectral image classification methods.....	24
2.6.1 Pixel-wise classification	24
2.6.2 Spectral-spatial classification	25
2.7 Chapter summary.....	26
3 Band Selection Method for Terrain Hyperspectral Imaging.....	27
3.1 Overview	27
3.2 Methodology.....	27
3.3 Classification method.....	32
3.4 Results	33
3.5 Comparison with benchmark datasets from remote sensing	35
3.6 Chapter summary.....	39
4 RGB Image Generation from Hyperspectral Datablocks.....	40
4.1 Methodology.....	40
4.1.1 Manifold alignment RGB image generation method	41
4.2 Results for RGB image generation	45
4.3 Chapter summary.....	52
5 Unstructured Terrain Semantic Segmentation	53
5.1 Methods	53
5.2 Performance matrices.....	53
5.3 Image datasets	54
5.4 Terrain segmentation using RGB images and manually labelled ground truth	55
5.5 Terrain segmentation with RGB images generated using HSI datablocks and manual labels	57
5.6 Terrain semantic segmentation based on the RGB images generated from the HSI visualization and pixel-wise classification	58
5.7 Spectral-spatial image classification	60

5.7.1 Classification method.....	61
5.7.2 Results of spectral-spatial classification.....	61
5.8 Chapter summary.....	63
6 Conclusion.....	64
6.1 Future works.....	65
List of Figures.....	66
List of Tables.....	67
References.....	68
Acknowledgements.....	76
Abstract.....	77
Lühikokkuvõte.....	79
Appendix.....	81
Curriculum Vitae.....	132
Elulookirjeldus.....	133

List of Publications

- I D. C. Liyanage, R. Hudjakov, M. Tamre, "Hyperspectral / Multispectral imaging methods for quality control" as a chapter of the book titled "Handbook of Research on New Investigations in Artificial Life, AI, and Machine Learning" by IGI global publisher, pp. 1–33, 2021, doi.org/10.4018/978-1-7998-8686-0.
- II D. C. Liyanage, R. Hudjakov, and M. Tamre, "Hyperspectral Image Band Selection Using Pooling," in 2020 International Conference Mechatronic Systems and Materials (MSM), 2020, pp. 1–6.
- III D. C. Liyanage, R. Hudjakov, and M. Tamre, "Hyperspectral Imaging Methods Improve RGB Image Semantic Segmentation of Unstructured Terrains," in 2020 International Conference Mechatronic Systems and Materials (MSM), 2020, pp. 1–5.
- IV A. Zahavi, A. Palshin, D. C. Liyanage, and M. Tamre, "Influence of illumination sources on hyperspectral imaging," Proc. 2019 20th Int. Conf. Res. Educ. Mechatronics, REM 2019, vol. 5, pp. 1–5, 2019.

Other publications:

- V Zahavi, A.; Najafi Haeri, S.; Chamara Liyanage, D.; Tamre, M. (2020). "A Dual-Arm Robot for Collaborative Vision-Based Object Classification". Proceedings of 17th Biennial Baltic Electronics Conference held in Tallinn, Estonia 2020. (BEC2020): 2020 17th Biennial Baltic Electronics Conference (BEC), Tallinn, Estonia, 6-8 Oct. 2020. IEEE. DOI: 10.1109/BEC49624.2020.9277067.
- VI Zahavi, A.; Al Afrange, F.; Najafi Haeri, S.; Ajeevan, U.; Chamara Liyanage, D. (2018). "ABB YuMi high-speed pick and place game in action". Proceedings of the 29th DAAAM International Symposium on Intelligent Manufacturing and Automation: 29th DAAAM International Symposium on Intelligent Manufacturing and Automation, Zadar, Croatia, EU, 24h-27th October 2018. Ed. Katalinic B. DAAAM International Vienna, 1216–1221. DOI: 10.2507/29th.daaam.proceedings.176.
- VII Dhanushka, Chamara Liyanage; Mart, Tamre; Indrek, Kivi (2017). Modelling and synthesis of belt-driven ship model testing carriage system. Proc. of 16th International Symp. Topical problems in the field of electrical and power engineering: Topical problems in the field of electrical and power engineering. Ed. Janis Zakis. Tallinn: TTÜ Dept. of Electrical Power Engineering and Mechatronics, 167–170.

Author's Contribution to the Publications

- I Dhanushka Liyanage is the primary author of the book chapter. The article discusses the hyperspectral and multispectral imaging methods in industrial applications for quality assessment. As the primary author of the chapter, Dhanushka carried out the literature review and furnished the article.
- II Dhanushka Liyanage is the primary author of the article who gathered test data, conducted experiments and compiled the article. The article proposes a novel method of band selection for hyperspectral images.
- III Dhanushka Liyanage is the principal author of the article. He created the concept of the work, gathered data and conducted experiments together with writing the article. The article discusses using hyperspectral image classification results as the ground truth for the RGB image semantic segmentation.
- IV Dhanushka Liyanage is one of the authors of this article who contributed to the work as a supervisor while conducting MHK0040 Machine Vision course as a teaching assistant.

Abbreviations

AI	- Artificial Intelligence
ALV	- Autonomous Land Vehicle
AGV	- Autonomous Ground Vehicle
CNN	- Convolutional Neural Network
DCNN	- Deep Convolutional Neural Network
DR	- Dimensionality Reduction
HSI	- Hyperspectral Imaging
IoU	- Intersection over Union
kNN	- k Nearest Neighbour
LiDAR	- Light Detection and Ranging
LPP	- Locality Preserving Projections
LWIR	- Long Wave InfraRed
MLP	- Multi-Layer Perceptron
MSI	- Multispectral Imaging
MWIR	- Medium Wave InfraRed
NIR	- Near InfraRed
PCA	- Principal Component Analysis
RCNN	- Recurrent Convolutional Neural Networks
ReLU	- Rectified Linear Unit
RGB	- Red Green Blue
SVM	- Support Vector Machines
SWIR	- Short Wave InfraRed
UGV	- Unmanned Ground Vehicle
VIS	- Visible Light
VNIR	- Visible Near InfraRed

1 Introduction

Without vision, it could have been extremely difficult for humans and animals to thrive. Similarly, computer vision plays a significant role in providing vision input for industrial robots. So does Autonomous Ground Vehicles (AGVs). All the cutting-edge driverless vehicles' autonomous systems heavily depend upon their vision sensors to perceive the environment for its autonomy apart from other sensor modalities. Some vehicle manufacturers solely rely on computer vision for self-driving [1, 2]. It was not technically feasible to achieve such a feat of self-driving without the latest developments in Artificial Intelligence (AI). The advancements in AI made it possible to solve highly complicated computer vision problems such as object detection and recognition. The AGVs are not only for passenger transportation, but also they are disrupting various application domains in transportation and logistics. Numerous research institutes, private companies, are working on developing AGVs in recent years. These autonomous mobile robots are meant to accomplish various tasks such as transportation of passengers, search and rescue operations, seeding, weeding, and harvesting in agriculture, parcels, and grocery delivery in the logistics sector, etc. Growing demand for different kinds of driverless vehicles fuelling the research and development of autonomy makes it one of the most active research areas today. Figure 1 shows different autonomous ground vehicles which are commercially available and in research these days.



(a)



(b)



(c)



(d)

Figure 1. (a) Military UGV developed by Rheinmetall AG[3], (b) agriculture robot weeding on a farm land[4], (c) Starship parcel delivery robot running on snow[5], (d) driverless shuttle bus in Japan for rural public transport [6].

Typically, AGVs can self-steer, self-regulate, avoid obstacles, and navigate without any human intervention. The first known research effort to develop an autonomous vehicle dates back to 1984 with the introduction of the Autonomous Land Vehicle (ALV) project by NavLab at Carnegie Mellon University [7]. It laid the groundwork to identify the main challenges of autonomous vehicle systems development, such as required sensing and computing capabilities.

Whether autonomous on-road vehicles or off-road vehicles, the fundamental technical problem is the environment perception. Perception is the ability of the autonomous system to collect information about its surroundings and extract relevant data [8]. The objective of a perception system is to enable the autonomous vehicle to navigate safely and efficiently. In contrast to the on-road autonomous vehicles, off-road AGVs face many more challenges in traversable terrain perception. The off-road environments are unstructured where it is not clear the traversable areas and non-traversable areas. Below, Figure 2 shows structured and unstructured terrains. The levels of hazards are often unpredictable, and constraints on mobility make it difficult for off-road navigation compared to on-road scenarios. Thus, off-road navigation demands a sophisticated perception system to realise complete autonomy, which is level 5 based on the vehicle autonomy levels classification proposed by the Society of Automotive Engineers (SAE). The levels of autonomy for road vehicles are not described here, which can be found on the SAE J3016 standard [9], and it is harmonised with ISO standards as ISO-PAS 22736 [10].



Figure 2. (a) A structured terrain where traversable and non-traversable areas can be easily distinguished. (Tallinn University of Technology parking lot) (b) Unstructured terrain where the traversable terrain is ambiguous (Location near Lasnamäe, Estonia).

The perception problem is currently being addressed by using various sensing modalities and processing methods. Nowadays, most AGVs depend upon Light Detection and Ranging sensors (LiDARs) and RGB colour cameras for localisation and environment perception [8, 11, 12]. Even the first autonomous vehicle, ALV, has used a TV camera and a laser scanner for terrain perception [7]. Different sensor modalities such as ultrasound sensors and radars are also used in specific applications to detect specific obstacles without being restricted to the RGB and LiDAR sensors mentioned above. Comprehensive details about sensors and measurements used on various unmanned systems are available in the reference [13]. Each of these sensors has its advantages and drawbacks. And some of them are prohibitively expensive for commercial applications.

Various terrain surfaces and conditions make it difficult for an off-road AGV to navigate with LiDAR sensors and RGB cameras alone. One of such limitations with LiDAR sensor-based perception is its inability to distinguish navigable obstacles such as grass and non-navigable obstacles like rocks and other hard objects [14]. The visual similarities between different terrain objects like construction rubble with painted objects may lead to false object detections, affecting path planning. In industrial vision applications, the Hue, Saturation and Values (HSV) colour model is used to overcome the issues such as illumination variations, shadows, reflections, etc. [15]. The HSV colour model has been used to detect traffic signs[16] and drivable roads[17] due to the advantages mentioned above.

Not only monocular RGB cameras but also various other imaging technologies are in use with mobile robots. Stereo cameras, infra-red cameras, omnidirectional cameras, and event cameras are examples of them [13]. Besides RGB and monochrome vision sensing technologies, spectral imaging is an advanced vision technology that combines both imaging and spectroscopy [18]. The spectral sensor collects image data on numerous spectral bands across the electromagnetic spectrum. Depending on the number of spectral bands present in the image, spectral imaging is classified into different classes. They are mainly Multispectral imaging (MSI), Hyperspectral imaging (HSI), and Ultraspectral Imaging (USI) methods. MSI and HSI are the widely used spectral imaging methods in the industry. Typically, MSI contains tens of spectral bands while HSI contains hundreds of spectral bands of narrow bandwidths [19]. Each pixel of such image is a high-dimensional vector that contains spectral reflectance in tens or hundreds of contiguous wavelength bands within a specific wavelength range. This technology can provide more information about the objects in the scene as it captures the reflectance on numerous spectral bands or image channels.

Combined spectra of the reflectance variation in each spectral band form a characteristic signature for each material which is the spectral signature. Such spectral characteristics can distinguish various terrain objects with better accuracy than RGB images. The spectral imaging technology initially appeared in the early 1970s with the Landsat satellite program, where it used a multispectral scanning system (MSS) comprised of visible and near-infrared spectral bands [20]. Spectral imaging is not limited to the visible light range in the electromagnetic spectrum compared with the RGB vision systems. Using appropriate wavelength ranges such as visible light, Near Infrared (NIR), Short Wave Infrared (SWIR), and Medium Wave Infrared (MWIR), it is possible to capture images that cannot be done using RGB cameras. Even though spectral imaging is widely used in precision agriculture, remote sensing, and various other industrial domains, it still finds little presence in AGV sensor suits used in mobile robotics.

1.1 Background and motivation

Multiple AGVs were introduced to the world from Estonia for package delivery [21, 22] military and civilian applications [23]. And Tallinn University of Technology is actively involved in research and development activities related to AGVs and UGVs. Ground vehicle autonomy is one of the hottest research topics in robotics. Therefore, improvements for the UGV perception based on hyperspectral imaging methods could be a valuable contribution.

Hyperspectral/multispectral imaging systems for autonomous driving has appeared in a few studies. Mainly for detecting a certain kind of obstacle that cannot be done otherwise. It is due to the abundance of information contained in the hyperspectral

image datacube. Typically, a hyperspectral datacube can be well over $\sim 100\text{MB}$ in file size. Pedestrian detection for self-driving vehicles is one such application where it uses a multispectral camera [24]. Based on terrain classification research by Winkers et al., HSI could be used for terrain perception by combining machine learning methods for image analysis [25]. Ice, mud, and loose gravel are often challenging terrain conditions for UGVs. In the reference [26], various imaging techniques were tried for mud detection. It has been concluded that the wettest regions of the terrain can be identified using SWIR wavelength range images, where wet mud from dry surroundings can be detected with RGB colour cameras.

In hyperspectral imaging, the image datacube, also called a hypercube, contain hundreds of spectral bands. These hypercubes often contain redundant data in most of the spectral bands. Thus, it requires selecting an optimal number of image bands from the datacube. Such band selection for terrain surfaces detection in various outdoor environments is still a challenging task.

The purpose of computer vision in AGVs is for scene understanding. Thus, a smart algorithm labels each pixel of an image, whether an RGB image or spectral image hypercube, with a terrain class. Such pixel-wise image labelling is called Semantic Segmentation [27]. Most of the existing scene understanding models in autonomous driving use RGB image semantic segmentation. Since current research has drawn significant attention for on-road autonomy, object detection and classification in structured environments have been thoroughly explored. Thus, there are a significant number of research publications on man-made object detection and semantic segmentation. They cover both unimodal approaches with single sensor input and multi-modal approaches by fusing multiple sensor inputs [28–30]. Hence, man-made object detection is not the primary objective for this thesis and instead, it focuses on detecting natural obstacles as they appear in their native environment. In contrast to 2D image classification, a combination of spectral-spatial feature-based classification could paint a more accurate picture of the terrain.

Self-driving vehicle manufacturers like Tesla, Inc already harnessing the power of modern machine learning methods such as deep neural networks for object detection, lane marking detection, etc. With state-of-the-art algorithms such as convolutional neural networks (CNN), it has been possible to extract the essential visual features crucial for autonomous navigation. Similarly, deep learning is becoming widely popular in HSI processing in various application domains [31].

1.2 Problem statement and research objectives

Hyperspectral imaging has excellent potential in terrain object detection, thus improving UGV perception, as explained in the previous section. Here are the research gaps found out in current research related to HSI for terrain perception.

- Most of the feature selection and feature extraction methods for HSI are developed based on various use cases. However, there has not been any such feature selection/ extraction methods developed for terrain classification. As autonomous vehicles have limitations in computational capabilities in their computing hardware resources, the methods should be efficient. Therefore, it requires developing a simple and effective feature selection algorithm for HSI terrain classification applications. The computation power for the perception system could be reduced by reducing the number of spectral bands in the images.

- The RGB imaging sensors work as the primary vision input in most of the ground vehicles available nowadays. These RGB images are used in semantic segmentation algorithms to classify various terrain objects, and these algorithms require large amounts of accurately labelled image data. Manual image labelling is a time-consuming and labour-intensive process, which hinders large scale dataset creation. There is a need for an efficient method to minimise the cost and effort of manual labelling.
- Although hyperspectral imaging is a novel technology in computer vision, still it finds little research effort put into spectral imaging to use this technology as a visual input for terrain perception in ground vehicles by the research community. There is a gap in terrain perception research to investigate the possibilities to use HSI segmentation to augment RGB image segmentation.
- Multispectral image semantic segmentation shows profound advantages over RGB image semantic segmentation in biomedical applications. The use of spectral images in the spectral-spatial classification method has not been validated for unstructured terrain segmentation. And there is no data to assess the effectiveness of spectral image based semantic segmentation accuracy over RGB for unstructured terrains.
- There are challenging environmental conditions such as black ice on driving terrain or wet soil encountered by both on-road and off-road vehicles. Driving on such terrain conditions poses a danger to both vehicles and operators. Whether there is a possibility to detect such challenging environments using spectral images should be explored.

Below are the main research objectives which will cover in this thesis.

1. Develop an effective band selection method to distinguish various terrain classes in unstructured environments. The method should be computationally simpler and efficient, thus enabling the outcome to implement on embedded computing hardware platforms with low computing capabilities often used in mobile robots.
2. Introduce an efficient method for image dataset generation using HSI.
 - a. Hyperspectral image pixel-wise classification in the spectral domain could be used as labels for image datasets.
 - b. RGB images could be created from HSI.
 - c. Thus, compare the effectiveness of such segmentation vs manually labelled image datasets.
3. Investigate state-of-the-art spectral-spatial HSI semantic segmentation networks based on deep convolution neural networks for terrain semantic segmentation. Draw necessary comparisons with RGB semantic segmentations.

1.3 Thesis contributions

This thesis made the following contributions to address above mentioned research questions.

1. One of the contributions of this thesis is an unsupervised band selection method to extract a few spectral bands from a hypercube. The proposed method can efficiently identify the most significant spectral bands for terrain image classification. The method was validated by classifying various terrain object classes found in off-road terrain scenarios (Publication I).
2. HSI pixel-wise classification using spectral data have been compared with RGB semantic segmentation, which shows that HSI pixel-wise classification is more accurate than that of the RGB semantic segmentation for terrain classes. Therefore, it can be summarised as the HSI pixel-wise classification can be used as ground truth labels for further semantic segmentation CNNs (Publication II).
3. The capabilities of HSI imaging were demonstrated as an alternative way to minimise manual image annotation for RGB image dataset preparation. The RGB images generated from 9-band HSI datacubes using manifold alignment methods proves that those images have a higher correlation to the actual RGB images captured from the same scene. The results show that the approach can be used effectively to train CNNs for RGB vision.
4. Demonstrated the effectiveness of HSI semantic segmentation in spectral-spatial domains for terrain classification scenarios. Introduced encoder-decoder based CNN architectures are highly effective in classifying such images trained with small image datasets.

The scope and boundaries of this thesis are as follows. Even though the work mainly focuses on ground vehicle application in off-road scenarios, it is not focused on a specific vehicle platform. The HSI processing and classification methods proposed in this thesis are not limited to ground vehicle navigation applications but equally can be applied to various other application domains.

This study was carried out at Tallinn University of Technology from 2017 to 2021. The outcomes of this doctoral research have been published in several IEEE conference proceedings, including MSM2020 and REM2019. And they are available on IEEE Xplore. The most relevant articles to the thesis are attached in the Appendix.

1.4 Research methods

The end goal of the thesis is to produce research outcomes that are feasible and viable to implement in real-world robotic applications. And the cost of the components used in the research is an important factor for future commercial viability. The hyperspectral cameras in the VNIR wavelength range are commercially cheaper than other cameras beyond the VNIR range. Although beyond the VNIR range, SWIR might provide more information of the scenes that cannot be captured in the VNIR range, it would limit the viability of real-life implementation. The hyperspectral imaging sensors in VNIR spectral range are based on silicon semiconductors, while the SWIR range is made of expensive semiconductors such as InGaAs, making those SWIR sensors pricey [32]. Therefore, this study focuses on hyperspectral imaging cameras in the VNIR wavelength range, which spans 400 – 1000 nm.

Following cameras were used in this research for HSI datacube acquisition:

- Ximea xiQ MQ022HG-IM-LS150-VISNIR camera from Ximea GmbH [33]
- Specim IQ mobile hyperspectral camera from Specim Spectral imaging Ltd [34]
- Resonon Pika II VNIR hyperspectral cameras from Resonon Inc [35]

Hyperspectral images captured using a portable camera tripod and mounting the hyperspectral camera at around 1,0 m height mimic the camera position of an actual AGV. The hyperspectral imaging sensor calibration was performed in-situ before taking images using the white reference calibration target provided by the camera manufacturer. The HSI hypercubes containing off-road terrain scenarios were captured at various locations in Estonia.

The RGB colour space will be used to compare the HSI perception methods.

All the algorithms, program codes, software test tools for this research were developed by the author. Furthermore, the experiments were carried out using the captured hyperspectral datacubes. MATLAB 2020b and its libraries were used as the primary programming tool, while Python and Visual C# programming languages have been used to develop software tools for various experiments.

The main processing computer is an HP Z4 workstation with Intel Xeon W-2123 3.6GHz Central Processing Unit (CPU), 32GB Random Access Memory RAM, and Nvidia P4000 GPU with 4GB RAM as hardware configuration.

1.5 Thesis structure

Chapter two of this thesis covers the existing literature review regarding state-of-the-art of terrain perception and spectral imaging. The same review further covers hyperspectral band selection approaches, classification methods, and semantic segmentation approaches on autonomous vehicle perception. The novel band selection method describes in the third chapter, where it presents experimental results and comparison results of image classification with existing methods. Hyperspectral images could be visualised in RGB colours to use on RGB semantic segmentation. The fifth chapter contains terrain segmentation using developed band selection methods and a comparison of RGB image semantic segmentation for off-road scenarios. The seventh chapter concludes the thesis with a summary of each contribution. The future research directions are also presented included in this chapter.

2 Literature Review

2.1 Terrain perception

Terrain perception is the core technical challenge for autonomy, and this topic has been in research for several decades. From early experimental autonomous vehicles, imaging and laser scanning were integral to terrain perception [7]. The RGB images and LiDAR point clouds are the most dominant inputs for terrain perception [36–39]. In addition to RGB imaging, several studies have explored the advantages of RGB-D imaging methods [30, 40, 41]. The majority of them focused on structured-environment perception, while a few studies went into unstructured environment perception.

On the contrary, an unstructured terrain object recognition method that has been presented by Erkan et al. [42] is purely based on vision sensors. As opposed to the impact of depth-sensing for off-road scene classification, Holder et al. suggest that RGB-D does not provide a significant advantage for detecting hazardous off-road terrains by taking several off-road datasets, including water puddles [43]. Alternatively, stereo vision for off-road object detection has been in multiple studies, including Defence Advanced Research Project Agency's Grand Challenge competition in 2005 [44], in agricultural environments [45] and a few other off-road autonomous driving projects [46]. Moreover, several studies have dived into beyond the visible light wavelength range imaging such as SWIR, MWIR and Long-Wave Infrared (LWIR) [26, 47, 48]. Objectives of such investigations were to detect various hazardous objects, hazardous terrains, safely navigate under adverse weather conditions and achieve passive terrain perception for sensitive applications.

The MSI/HSI based perception research attracted less attention than the RGB imaging-based perception research over the years. Here is a summary of recent developments in spectral imaging in terrain perception for autonomous navigation of ground vehicles.

One of the early research work into spectroscopic imaging for terrain perception had been realised using Acousto-optic tunable filters (AOTF) along with a Charge Coupled Device (CCD) camera in the late '90s [49]. Apart from the AOTF, polarising filters have been used in the same research to detect road hazards such as wet and icy roads.

For a natural terrain classification task, a multispectral imaging-based technique with machine learning has been used by Namin et al. [50]. The authors have used VNIR spectral range multispectral camera with seven spectral bands where they achieved a classification accuracy of 92 % and 89 % with SVM and Adaboost classification methods. The number of terrain classes used for the classification was ten which include natural and man-made terrain objects. Winkens et al. have proposed an HSI approach using a snapshot hyperspectral camera with 16 spectral image bands for drivable and un-drivable terrain detection. Even though the camera spectral range is limited to 470 – 620 nm, the authors presented successful results for a terrain classification of four classes using the Random Forest algorithm [51]. A four-channel multispectral camera has been used in vegetation and bare soli detection within the visible light wavelength range. The authors suggest that the Normalised Difference Vegetation Index (NDVI) index is successful vegetation and soil differentiation measure [26]. The research works mentioned above fulfil the needs for scene understanding further ahead of the vehicle, which would improve long-range path planning.

Long-range path planning, and nearby ground analysis for the traversability around the vehicle, are both important tasks in autonomous driving. The ground surface texture

around a vehicle captured using a mosaic hyperspectral camera with 25 image bands organised in a 5x5 mosaic pattern has been used to project the texture to the ground surface height map created using 3D LiDAR. Using the method suggested by Fuchs et al., the navigation system can map the terrain around the vehicle. The proposed approach suggests that semantic analysis of the environment can help detect and avoid risks posed by terrain right ahead of the vehicle [52].

2.2 Semantic segmentation of terrains

2.2.1 Overview of semantic segmentation

For AGV navigation, drivable and undrivable terrain should be identified by the perception systems, which typically comprised of RGB cameras and LiDAR sensors, as mentioned before. Scene understanding based on images is the objective of semantic segmentation. Assigning each image pixel with a certain category (class) label or semantic label is called semantic segmentation in computer vision. Semantic segmentation helps to locate the specific objects in their exact spatial locations in the image. In terrain semantic segmentation, the classes or categories are different terrain types such as gravel, dirt, mud, grass, etc. Such image segmentation is a common classification type problem in computer vision. State-of-the-art semantic segmentation methods are mostly CNN algorithms based on deep learning. The deep neural networks require a large number of images and labels as inputs.

With the introduction of Fully Convolutional Networks (FCN), the deep learning models revolutionised semantic segmentation with remarkable accuracy for RGB images [53]. During the past several years, deep learning methods evolved rapidly. As a result, there are numerous architectures introduced with various features. A deeper comparison between major network architectures, such as VGG, ReNet, ResNet, DenseNet, ResNeXt, and MobileNet, is discussed in a review article published by Hao et al. [54]. The same article discusses various novel methods based on the above CNN backbones, real-time methods, weakly supervised methods, popular public datasets, evaluation matrices of semantic segmentation, etc. A comprehensive description of the DCNN constituents such as convolution layers, pooling layers, fully connected layers, activation functions is presented by Kaymak et al. in their semantic segmentation review article, which is oriented towards autonomous driving applications [27].

The “2.6 Hyperspectral image classification methods” section will discuss the multi-layered image semantic segmentation methods.

Fully supervised learning methods proven to produce the highest accuracy in semantic segmentation, which requires accurately annotated image data labels. Even though there are software tools for image labelling, manually annotated images are the most accurate. Instead of strongly annotated label images, semantic segmentation accuracy has been investigated using a large number of weakly labelled and a small number of strongly labelled images showing 69 % IoU for the PASCAL VOC 2012 benchmark. Compared to a strongly annotated image dataset that yielded 70,3 % IoU, the weakly annotated fell short by a 1,3 % margin [55]. In this case, the weakly annotation was locating the objects on the image with bounding boxes that shows the coarse object locations.

2.2.2 Unstructured terrain semantic segmentation

Semantic segmentation of unstructured terrains is a complex task using RGB images due to numerous classes in the scene, which are of a mixture. When it comes to on-road terrain semantic segmentation, most of the classes, objects are man-made, such as roads, road signs, buildings, lamp posts, other road vehicles, etc. In terrain perception, most scene understanding models are developed for structured terrains or regular urban driving environments.

In contrast to the structured terrains, literature for unstructured terrain semantic segmentation is rather sparse. As a result, this review is based on three research articles published based on unstructured terrains [56–58]. There are different types of unstructured terrains. They are either urban driving environments that lack the structure or fully off-road terrains. Most developing countries have less structured driving roads, even in urban areas, which resembles unstructured terrains. Those unstructured terrains comprise gravel roads, gravel sidewalks, muddy spots, and water puddles, etc. Semantic segmentation of such terrains has been introduced with modified DeeplabV3+ DCNN by Baheti et al. and has been tested with India Driving Dataset (IDD), achieving 68.61 mean Intersection over Union (IoU) [56]. The same authors proposed another DCNN architecture called Eff-UNet for the same task, tested on the above mentioned IDD [57]. The Eff-UNet architecture is based on EfficientNet[59] encoder and UNet [60] decoder forming encoder-decoder architecture. With this architecture, the authors have achieved an accuracy of 62.76 mean IoU for IDD. The encoder-decoder architectures manifest that they are the most outstanding DCNN architectures so far to perform semantic segmentation. Following encoder-decoder architectures, Sgibnez et al. have proposed a lightweight DCNN architecture for an off-road terrain semantic segmentation [58]. The authors have experimented with multiple DCNN options as backbones for their encoder-decoder architecture based on ResNet, MobileNetV2, ShuffleNetV2 and EfficientNet-B0. All the reviewed articles above are based on RGB images as inputs.

There are several challenges for semantic segmentation performance when those classification algorithms run in real-time in real-world applications. The variation of environment lighting conditions, weather, camera parameters and shadows can greatly influence the classification outcome [61, 62]. A robustness assessment method for semantic segmentation was proposed by Wen Zhou et al., which uses a LiDAR. However, the proposed method has been validated on structured terrain conditions, where they validated road class under various environmental conditions [61].

U-Net Architecture

The U-Net architecture was introduced to biomedical image segmentation in 2015 [60]. A higher semantic segmentation accuracy with a smaller dataset is the most significant advantage of this architecture. Data augmentation has helped achieve the claimed accuracy level, which the authors have demonstrated the network accuracy of 92% and 77.6% using two biomedical image datasets containing 35 and 20 annotated training images, respectively.

The U-Net architecture has been built with a contracting encoder branch and an expanding decoder branch. Two convolution operations with Relu activation are used in every layer of the network. Convolution input layers are not padded where every convolution reduces the size of the output. Since it uses 2 x 2 max pooling and up-convolutions, the input images should have even X and Y dimensions. The skip connections help recover some of the fine details while decoding. The U-Net architecture is illustrated in Figure 3.

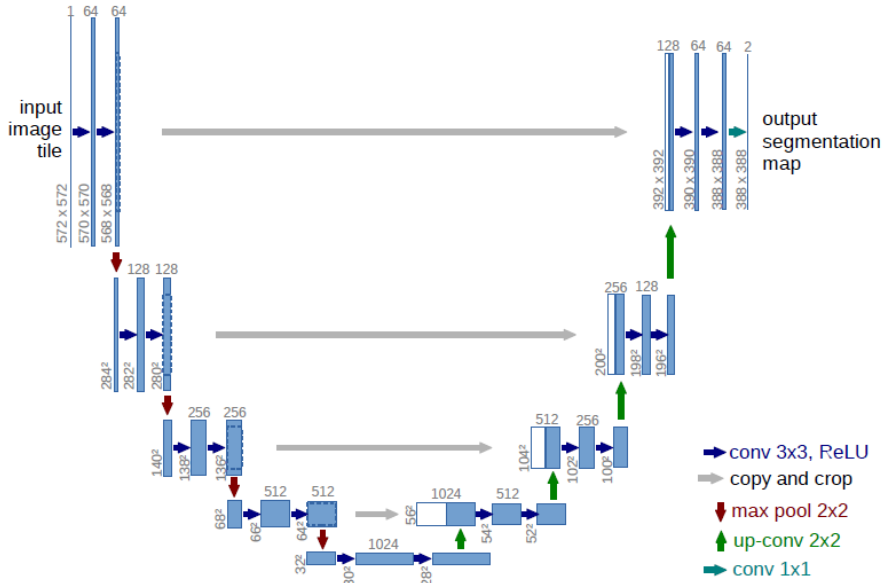


Figure 3. U-Net architecture for biomedical image segmentation [60].

2.2.3 Image annotation

Various tools have been developed to minimise the effort it takes for semantic image annotation, such as Supervisely [63], SuperAnnotate [64], MATLAB Image Labeller [65], LabelBox [66], LabelMe [67], and Microsoft Visual Object Tagging Tool (VoTT) [68]. These applications offer various shape matching tools such as circles, triangles, rectangles, and other polygons to define the object’s boundaries in the image. The polygon-based methods do not accurately define the object boundaries.

Image boundary-based semi-automated annotation tool proposed by Qin et al. claims that it reduces manual clicks to select image boundary points by 73%. The authors use edge detection and splitting algorithms to cluster the image regions, followed by a few manual selection clicks to annotate the image. This approach does not depend on any other input to assist annotation [69]. A semi-automated annotation tool called “EasyLabel” was proposed by Suchi et al. for an indoor object dataset [70]. Their method is based on RGB-D vision and is suitable for pixel and object-wise labelling of indoor objects. With the help of an object detector model and 3D layout estimator, Reza et al. have proposed a video data automatic annotation method. Their application is also focused on indoor scene labelling [71].

Based on the existing literature, the outdoor scene annotation methods have not gained enough attention.

2.3 Spectral imaging

Hyperspectral imaging technology is being successfully used in various industrial application areas such as precision agriculture [72, 73], food quality inspection [74–76], pharmaceutical production [77, 78], crime investigation [79, 80], waste sorting [81, 82] and so on. However, it is still finding little presence in autonomous driving.

In contrast to RGB imaging methods, MSI and HSI methods offer more information about the objects in the scene. The RGB imaging method provides only the spatial information of the scene in comparison to MSI or HSI. Moreover, RGB imaging only provides information within the visible light wavelength range, 400 – 760 nm. On the contrary, the MSI and HSI methods provide both spatial information and spectral information. These two methods are used within and beyond the visible light wavelength range. Below Table 1 summarises the capabilities of each image acquisition method based on the information available in the image relative to each other.

Table 1. Main differences between HSI, MSI, Spectroscopy and RGB imaging. Classification bullet rate (1-3) is a relative representation of the information available in each acquisition method [83].

Imaging method	Spectral information	Spatial information
Hyperspectral imaging	●●●	●●●
Multispectral imaging	●●	●●●
Spectroscopy	●●●	●
RGB imaging	●	●●●

Based on the number of spectral bands in the sensor, the constructions of MSI and HSI sensors and the spectral cameras is different. So do the image acquisition methods. Most MSI sensors are constructed with bandpass filter deposition in a mosaic pattern[84][85] for single-sensor cameras. There are multi-sensor cameras built using several single band sensors combined as an array [86]. The HSI cameras can be based on tunable filters, line scan sensors, snapshot imaging sensors. Line scan with push-broom method is widely used for HSI image acquisition, which differs from the area scan method used in RGB imaging and MSI. Line scanning is a slow process compared to the area scan that restricts the capture of images of dynamic objects. Different spectral image acquisition modes are discussed by Li et al. in their spectral imaging review article for biomedical imaging [87]. Even though HSI and MSI provide more information about the scene, both technologies have characteristic issues.

2.4 Dimensionality reduction in hyperspectral images

The HSI data cubes contain a large amount of spectral data in hundreds of contiguous spectral bands. Therefore, it provides high-resolution spectral characteristics of the object under investigation than other imaging methods. Even though there are advantages of high-resolution spectral characteristics, there are drawbacks as well. A large amount of data makes datacube processing computationally expensive. Moreover, the spectral data in contiguous bands are often redundant and correlated. Such a large amount of data causes the Hughes phenomenon. The Hughes phenomenon is when the classification accuracy increases by increasing the number of spectral and then drastically drops when it reaches a certain number of features in small sample size [88, 89].

Due to these reasons, there is a need for selecting optimum spectral bands which contains sufficient information to characterise the objects under investigation. This process is called dimensionality reduction, which transforms high dimensional data into low dimensional space. There are two different approaches to address this problem. One is feature extraction, while the other is feature selection or band selection [90]. The two methods are significantly different from each other. Feature extraction methods will not

contain original spectral information after the dimensionality reduction due to data transformation during the process [91].

On the other hand, the feature selection approach does not alter the original pixel data in the hyperspectral data cube, thus preserving the intrinsic information [91]. So that, one of the goals of this thesis is to find out the most suitable band selection/feature selection method for a terrain classification task. By selecting a limited number of spectral bands, it would be possible to develop an optimised multispectral sensor for UGV imaging applications, eventually enhancing the classification speed. The feature extraction methods are not suitable for this thesis work as they will not maintain the correlation between original spectral bands and the transformed data.

For autonomous ground vehicle applications, CNN training using numerous wavelength band groups according to terrain classes is a complex, labour, and time-intensive task. Mainly due to a large number of terrain classes available in unstructured terrain environments. The objective of band selection or feature selection is to identify the optimum spectral bands which provide the most distinctive and informative spectral characteristics to achieve the highest classification accuracy [92]. Numerous band selection methods have been introduced by the research community based on various applications and imaging conditions. These band selection methods are categorised into six different groups as ranking-based, searching-based, clustering-based, sparsity-based, embedding learning-based, and hybrid-scheme based methods [93].

The supervised band selection methods perform superior to the unsupervised methods [94]. However, in the case of terrain classification for unknown environments, supervised methods require retraining for new terrain object classes, which makes it less desirable over unsupervised band selection methods. Therefore, the ideal band selection method should be either an unsupervised or a semi-supervised method that uses less labelled data. Nevertheless, supervised methods are widely used in HSI band selection.

Most band selection methods are benchmarked with widely popular public datasets such as the Indian Pines dataset [95], Salinas valley dataset, and Pavia University dataset [96]. These datasets are mostly used in remote sensing research. Indian Pines dataset has been collected by AVIRIS sensor over North-western Indiana, which contains 224 spectral bands in visible/near-infrared wavelength range of 400 – 2500 nm. The dataset contains 16 classes [95]. Salinas valley scene also collected by AVIRIS sensor with 224 spectral bands spanning from visible – infrared wavelength range. The image contains 16 different classes [97]. Pavia dataset has collected over University of Pavia, Northern Italy by ROSIS – 3 sensor which contains nine different classes. The dataset contains 103 spectral bands in the 430 – 860 nm visible light wavelength range.

Unsupervised dimensionality reduction methods

Unsupervised band selection methods are developed based on different criteria. Mostly used criterion is the correlation between different spectral bands.

The Principal Component Analysis (PCA) methods transform data from a high dimensional space into a lower-dimensional space [98]. The first few dimensions of the resultant PCA conversion contains most of the information. However, the resultant image keeps the important characteristics of the dataset while dimensionally reducing PCA datacube. After PCA, the resultant image does not reveal its original intrinsic image band structure. This method is older and widely used in hyperspectral image dimensionality reduction as a benchmark.

Manifold learning is another non-linear dimensionality reduction approach. The spectral-spatial manifold reconstruction preserving embedding method has been

proposed by Huang et al., a novel manifold learning method. Their method proposes a novel distance criterion called spectral-spatial combined distance as the distance matrix. In this method, the HSI data is filtered using a weighted mean filter as the first step. Then, it uses spectral-spatial combined distance matrix to select spectral-spatial neighbours. A spectral-spatial adjacency graph is constructed using graph embedding in this method. This method would reveal the intrinsic manifold structure [99].

The similarity measure is one of the criteria for band selection, where it uses mutual information of spectral band and a reference band to select the most suitable spectral bands.

Supervised learning-based band selection methods

There are numerous band selection techniques based on the supervised learning method. Some of them use state-of-the-art CNN algorithms, which outperform most of the classical methods.

Band selection based on band-wise independent convolution and hard thresholding technique has been incorporated into CNN by Feng et al. The proposed method comprises band selection, feature extraction and classification [100].

A CNN based band selection method has been presented by Rui C et al. called CM-CNN, in which they propose a new structure in the neural network called contribution map. This method allows the extraction of discriminative spectral bands from the dataset. However, according to the published experimental results, the proposed CM-CNN method yields higher classification accuracy when selecting more than 40 spectral bands for the Indian Pines dataset and more than 30 spectral bands for the Pavia University dataset [14]. Unlike the other NN based methods, this method gives the most discriminative spectral bands while it cannot achieve higher classification accuracy for fewer number of (< 20) spectral bands.

Attention-based CNN method for HSI band selection could achieve comparable classification results with SVM. The attention-based CNN algorithm has been tested and validated on the Salinas valley dataset and Pavia university dataset. In that study, the researchers have compared the classification accuracy of three different CNN models, which did not show any statistical significance [101].

Band selection based on CNN with distance density proposed by Zhan, Y et al. show a better classification accuracy with a higher number of spectral bands. Especially the proposed BSCNN+ DA, which uses data augmentation with distance density calculation, outperforms all the other methods in their research. The most significant achievement in their study is that the method initially uses all the spectral bands in the datacube to train CNN. Then use distance density-based bands reduced datacube using Rectified Linear Unit (ReLU) activation function to iteratively select bands classifying the image and estimate the precision of each method to extract most significant spectral bands without retraining the CNN for every iteration. As the authors mentioned in their publication, their method is not effective when the number of selected spectral bands is less than 30. The other traditional methods can easily outperform the BSCNN+DA method if the selected spectral band count is less than 30. However, their method has other significant drawbacks as it needs to test all possible band combinations for every band partition in the datacube until it finds the optimum bands set. In other words, it requires a large number of iterations to identify the suitable bands, which makes their method computationally expensive [91].

2.5 Hyperspectral image visualisation

In order to train RGB semantic segmentation neural networks, it is required to prepare image datasets with RGB images and their ground truth. The ground truth is the actual pixel class which comes from image pixel-wise classification. The RGB image needs to be generated from the HSI datacube. The goals of such an RGB image generation method should be consistent rendering, edge preservation, computational ease, and natural colour palette, which are the most relevant in this research context [102].

Selecting three image bands from red, green, and blue wavelength ranges and mapping them to RGB channels is the simplest method to generate RGB images from HSI [103].

The hyperspectral image visualisation method proposed by Su et al. is based on visualisation orientated band selection followed by band similarity computation. In this case, they select three bands to produce trichromatic visualisation from three wavelength ranges for red, green, and blue, respectively. They use a band selection method that gives fewer spectral bands for each colour wavelength range, which is subsequently used to calculate the correlation coefficient to determine the similarity between the bands in the same wavelength range to choose one image band. Also, it looks for the dissimilarities between the selected bands in the other colour wavelength range, which can add most of the information into the final visual representation [104].

Several methods have been developed on filtering techniques [105, 106]. One of them is a most recent method based on PCA with edge-preserving filtering, which claimed to have better image contrast and original details of the image [105]. Similarly, using PCA and edge-preserving filtering, the bilateral filtering-based image fusion method has been proposed by Kotwal and others. Their method uses band weights at each pixel for image fusion [106].

Another image visualisation approach is the manifold alignment technique. In this case, the RGB images captured from the same or semantically similar site to the location where HSI images capture, the manifold alignment approach can transfer colour information from RGB images to HSI visualisations. Fusing the hyperspectral image with a high-resolution colour image using manifold alignment needs a few matching pixel pairs that present the same object in the scene [103].

2.6 Hyperspectral image classification methods

In hyperspectral imaging, achieving the highest classification accuracy is the main objective. There are various methods have been developed for hyperspectral image classification in various application domains. One of the challenges in HSI classification is the difficulty to extract of endmembers. The endmember is the pure signature of a particular material (object) class [107], often discussed in chemical/mineral analysis and other fields. The objects found in nature do not occur in their pure state, which adversely impacts classification accuracy.

2.6.1 Pixel-wise classification

Most of the HSI classification methods are developed for pixel-wise classification. From statistical methods to modern DCNN methods developed in the context of spectral signature-based discrimination.

Multilayer perceptron (MLP) networks are the simplest and the basis of artificial neural networks, which has been in use with diverse applications. These feed-forward

NNs were introduced for HSI classification decades ago. Even though they are simple, they can achieve high accuracy in hyperspectral image classification [108]. After the band selection, the hypercubes contain fewer spectral bands, which can efficiently classify the image despite some inherent pitfalls with the method.

Various CNN based classifiers have been developed in recent years. There are several feature extraction CNN architectures proposed in the reference [109]. The authors of the publication have proposed 1-D CNN, 2-D CNN and 3-D CNN for feature extraction and image classification.

The 1-D CNN is a single-pixel hyperspectral image classification network. The depth dimension of the datacube is the spectral distribution, which is 1-D image input for the NN [110]. These networks also contain convolution layers, activation function layers, and pooling layers. As mentioned before in the supervised band selection methods section, the distance density based band selection method uses 1-D CNN for classification [91]. Furthermore, the 2-D CNNs look into spatial features of each layer while 3-D CNNs learn both spectral-spatial features [109].

2.6.2 Spectral-spatial classification

Hyperspectral image classification based on both spectral-spatial features is the most appropriate classification approach for terrain perception. As it is mentioned before, the endmember identification for terrain object classes is not a viable solution. Since image pixels in the hypercube may have mixed with noise and neighbouring spectral data cause spectral mixing, the classification result based on spectral signature may not always be accurate. A combination of spectral-spatial features can overcome such complications. In contrast to the classical methods based on pixel-wise spectral data classification, the spectral-spatial combination looks at both spatial image features on two spatial dimensions and spectral image features on the third dimension. Therefore spectral-spatial combined classifiers perform better than pixel-wise spectral image classifiers [111]. There are numerous spectral-spatial HSI classification methods have been published so far. The earliest spectral-spatial joint classification of HSI was researched several decades ago, where Landsat imagery data were classified using neighbouring pixel patterns and the Markov approach [112].

However, HSI classifications related to terrain perception for autonomous navigation applications is nearly non-existent compared to remote sensing applications. The spectral-spatial classification methods reviewed in this state-of-the-art review are based on the methods developed for remote sensing applications and other multi-band image classification applications such as medical image classification. The key criterion for the optimal image segmentation is the computational time, and classification accuracy of the method benchmarked on each domain's popular datasets.

The state-of-the-art classification architectures are mostly DNN architectures. In the remote sensing application domain, the classification methods are mostly benchmarked on popular remote sensing datasets such as Indian Pines, Pavia University, KSC, and Salinas datasets. The most efficient models in remote sensing are 3D CNNs which achieved the highest classification accuracies, well over 95% for popular datasets and with the lowest processing times [113–115]. These architectures use image patches or smaller windows of the image as inputs to the classification network. Thus, the resultant is a pixel class that gives the centre pixel class of the input patch as the outcome.

2.7 Chapter summary

There are various band selection methods available for HSI. However, they are optimised for various applications. Those methods were developed without aiming to use embedded computing hardware, which can be used on an AGV. Therefore, it is needed to develop an efficient band selection method aimed at discriminating terrain surface classes. Based on the research, the data sparsity-based methods are simpler yet effective for band selection. Thus, novel band selection should be based on the data sparsity-based method.

The RGB image ground truth labelling is still being done by manually using various labelling software tools. However, it is a costly process to generate a large image dataset. The possibilities to use hyperspectral imaging to assist in identifying the ground truth has not been explored. In the situation of HSI use in terrain segmentation, most of the perception might still be dependent on RGB imaging due to faster classification speed. Under such conditions, there will be a need for RGB semantic segmentation CNN training for unstructured terrain classification cases. Due to the difference in spatial resolutions of HSI and RGB images, there might be a need for image labels for both cases separately. As some CNN models will not work with varying spatial resolutions, creating two image datasets would be difficult due to the intensive manual work required for such operation. There has not been any research on that area to utilise HSI based CNN model for RGB classification. Such development needs to make RGB images from HSI datacubes. The RGB image generation from HSI in unstructured terrain context also has not been explored before.

The U-Net DCNN architecture has proven successful in semantic segmentation with a limited number of training samples. Furthermore, the U-Net architecture accommodates multi-channel images, which resembles HSI datacubes. Since the hyperspectral dataset in this research is not a large image dataset, the U-Net architecture would be the most appropriate solution for HSI classification. There are spectral-spatial segmentation models that have been developed for multi-layered images. Still, those techniques have not been widely applied in HSI image classification for autonomous vehicle terrain perception scenarios.

3 Band Selection Method for Terrain Hyperspectral Imaging

3.1 Overview

One of the objectives of this thesis is to investigate band selection methods that can be used to reduce the number of spectral bands of HSI datacubes while preserving intrinsic characteristics. The suitability of the method should be evaluated based on its performance in unstructured terrain classification applications. According to the literature review in the preceding chapter, numerous band selection techniques are available for various applications. The method proposed in this chapter is similar to the pooling operation commonly used in CNN algorithms.

The experimental unstructured terrain datasets were acquired using Specim IQ mobile hyperspectral camera. These datasets represent off-road terrain in Estonia as this research focused on off-road terrain classification.



Figure 4. The RGB image of the terrain HSI dataset.

3.2 Methodology

The concept of the band selection method is to use the pooling operation in the spectral axis. Thus, detect the most significant peaks from each pooling kernel and extract distinct spectral bands.

The dataset was calibrated for relative reflectance using white and dark references taken while capturing the image. The data cube contains 12-bit resolution reflectance data for each band. Therefore, the calibrated image was kept with the same resolution. In (1) I_i denotes the calibrated reflection intensity at i^{th} band while I_{oi} is the reflectance of the original image at i^{th} wavelength band. Similarly, W_i and D_i denote the white reference reflectance and dark reference reflectance at the respective band.

$$I_i = 4095 \times \left(\frac{I_{oi} - D_i}{W_i - D_i} \right) \quad (1)$$

The objective of band selection is to extract a certain number of spectral bands which characterise the objects in the scene to define spectral bands for a multispectral sensor. As the first step, it is required to define the desired number of spectral bands, i.e. 9, 16 or 25 bands.

The number of pooling iterations is calculated based on the desired bands count, together with the size of the pooling window for each iteration. A hyperspectral image is a 3D datacube with height (h) and width (w) as spatial dimensions while the number of bands on the third dimension or spectral dimension. Since band selection is performed on the spectral axis, it is required to unfold the image. If the image contains only a few classes that are necessary to distinguish, then the entire datacube could be unfolded and used to find out the characteristic spectral bands. On the other hand, if the datacube contains various objects which do not need to be distinguished, then the classes of interest should be selected from the images.

After unfolding, it is a stack of pixels with spectral distributions. The HSI images used in this chapter have 204 spectral bands, and each pixel's spectral distribution shows the reflectance intensities for all 204 bands.

The original data cube columns are converted into rows of the modified image as in Figure 5. The purpose of unfolding is to perform pooling operations only on spectral data instead of spatial data.

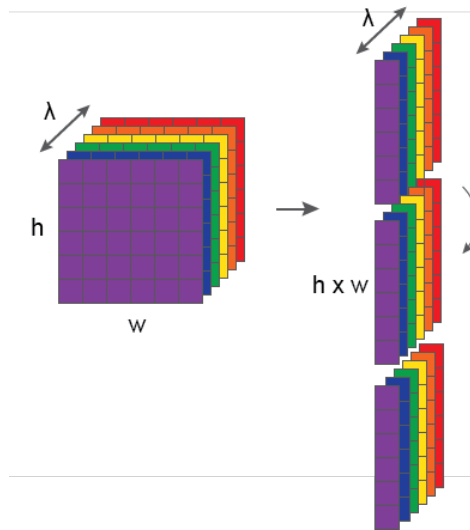


Figure 5. Spectral data cube unfolding.

Max Pooling

This method finds the largest feature in the pooling kernel. The kernels are chosen as non-overlapping. Max reflectance of the filter is calculated using (2).

$$Y_{max} = \max_{i=0,\dots,h,j=0,\dots,w} X_{i,j} \quad (2)$$

Min - Max Pooling (Proposed method)

In contrast to the max-pooling method, the min-max pooling method involves searching for the local minima and maxima (local extremum) in the pooling kernel. Peaks and valleys or the significant changes in the reflectance characteristic curves are the most useful spectral features for band selection. Therefore minimum points in the curve or valleys need to be taken into consideration too. This method searches for both minima and maxima in the current n^{th} kernel. The minima are calculated using (3).

$$Y_{min} = \min_{i=0,..,h,j=0,..,w} X_{i,j} \quad (3)$$

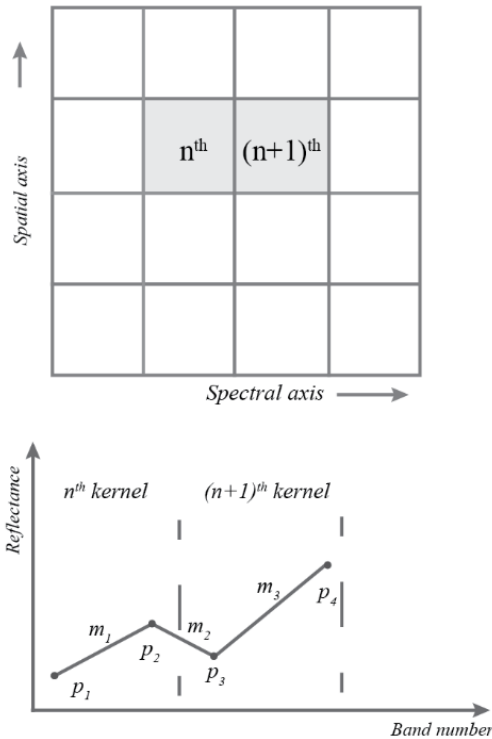


Figure 6. Min-max pooling method.

Similarly, maxima and minima for the $(n + 1)^{th}$ kernel will be calculated using the same equations (2) and (3). Find the maxima and minima coordinates in the spectral axis, which are the band numbers. The adjacent gradients are calculated between three consecutive extreme locations. Depending on the calculated gradient differences between extreme points, the band of either maxima or minima will be selected from the partition.

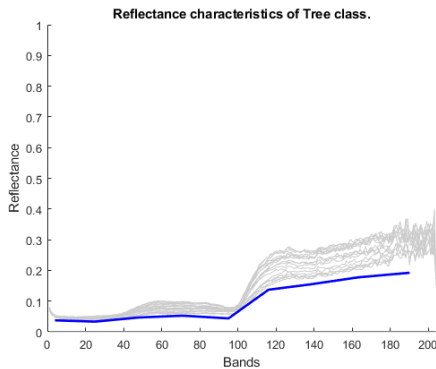
In this algorithm, let the consecutive points be p_1 , p_2 and p_3 and the gradients of p_1p_2 as m_1 , p_2p_3 as m_2 . Figure 6 shows how these points are located on the spectral axis. If the gradient difference is more than 0.15, which was obtained empirically, the point p_2 will be more significant in the kernel: otherwise, p_1 is more significant. The vector B is taken from (4). Below coefficient, ϵ is obtained empirically and depending on the classification accuracy for chosen classes. For the terrain classification, the ϵ was

chosen as 0,15. After obtaining band vectors for all n number of samples which is the number of pixels in the image, the maximum occurrence bands will be taken as the final band set.

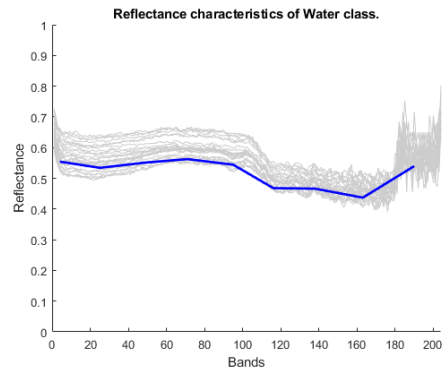
$$B = \begin{cases} p2, & \text{if } m2 - m1 > \varepsilon \\ p1, & \text{otherwise} \end{cases} \quad (4)$$

This gradient search technique ignores the band with low reflectance intensity if two similar points lie on consecutive bands. Even though it can catch subtle points in the spectral signature, it has a downside over max pooling. This method takes considerable computing time as it requires calculating gradients in each iteration.

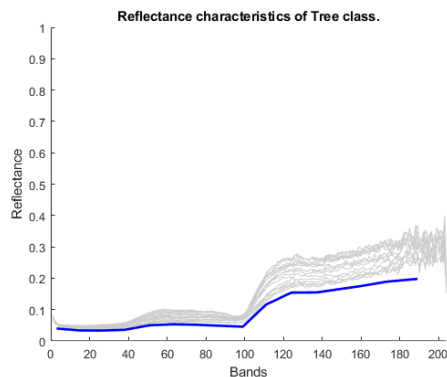
Single hyperspectral datacube often does not contain all the terrain classes. Therefore multiple datacubes are used to extract training data. Instead of using entire datacubes, a few image patches representing each class are selected and implemented the above-explained method. The spectra of two selected terrain classes, trees and water, are illustrated below in Figure 7 with different band counts. The thick blue line corresponds to the average spectra in 9, 16 and 25 bands characteristic curves, while the thick black line shows average spectra for the 204 bands characteristic curve of the terrain class.



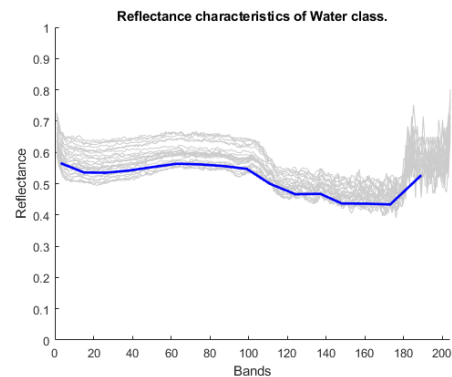
(a) Tree class with 9 bands



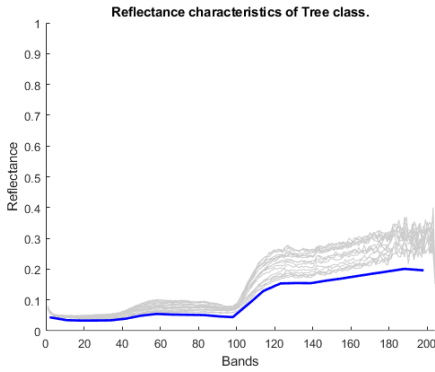
(b) Water class with 9 bands.



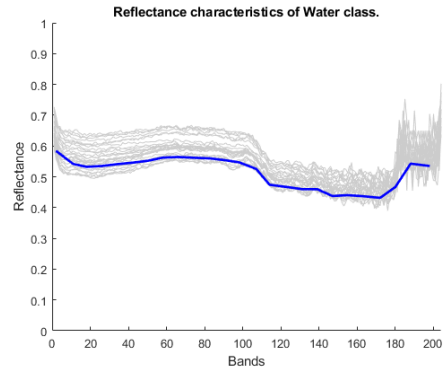
(c) Tree class with 16 bands.



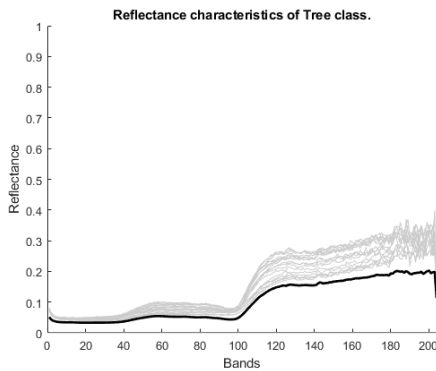
(d) Water class with 16 bands



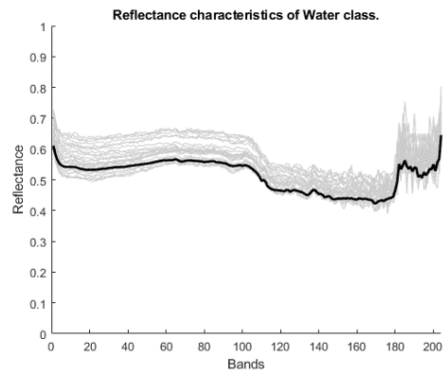
(e) Tree class with 25 bands.



(f) Water class with 25 bands.



(g) Tree class with 204 bands (all the bands).



(h) Water class with 204 bands (all the bands).

Figure 7. Spectral characteristics of Tree and Water terrain classes.

Apart from the proposed method, another method introduced by different group researchers is used in the experiments for comparison. It is the Distance density + CNN method [91]. The selected spectral wavelength bands for each method are in Table 2.

The selection of 9, 16 and 25 spectral bands is based on the possibility to develop a custom multispectral sensor. When constructing such a multispectral sensor, the bandpass filters are deposited in a mosaic pattern, a square matrix. The HSI acquisition uses the line scan method, while the MSI acquisition uses a much faster area scan technique. MSI classification is faster than HSI due to the fewer spectral bands in the image that can be used in AGV applications for faster processing, which leads to achieving real-time terrain scene classification.

Table 2. Selected spectral bands from various band selection methods.

Number of bands	Method	Spectral bands
9	Min-Max Pooling	2, 26, 46, 70, 96, 117, 139, 162, 190
16	Min-Max Pooling	2, 15, 27, 38, 50, 63, 76, 87, 99, 111, 124, 137, 148, 160, 173, 189
25	Min-Max Pooling	2, 11, 19, 27, 34, 42, 50, 58, 67, 75, 83, 91, 98, 107, 114, 123, 131, 139, 147, 155, 163, 172, 180, 188, 198
9	Distance density + CNN	22 61 127 134 162 168 174 184 192 195
16	Distance density + CNN	26 54 114 121 125 149 154 158 166 180 186 191 192 196 197 199
25	Distance density + CNN	15 17 57 71 112 118 123 125 136 143 154 163 164 169 171 172 175 180 183 184 185 187 188 191 192 200

3.3 Classification method

A supervised machine learning method is used for HSI classification. The HSI classifier in this task is a spectral-spatial CNN model, and it will be referred to as SS_CNN throughout this thesis. The classifier is as follows. The image input layer accepts 5 x 5 pixels spatial resolution and d number of channels. It is a 5 x 5 pixels window sliding over the image and classifying the centre pixel of the sliding window. After the input image layer, there are two convolution and Relu layers. Both convolution layers contain 3 x 3 kernels and 16 filters. The second Relu layer is followed by a fully connected layer and a softmax layer, classifying the pixels according to the number of classes. The architecture is described in Table 3.

Table 3. CNN classifier architecture for HSI pixel-wise classification.

Layer	CNN architecture							
	In	c1	r1	c2	r2	fc	softmax	out
Kernel size, Filters	5x5	3x3, 16 filters		3x3, 16 filters		d		

In this pixel-wise classification method, the training data are individual image pixels, and all the training pixels are collected from the same image. As shown in Figure 8, the small patches of pixels are extracted for training and validation.



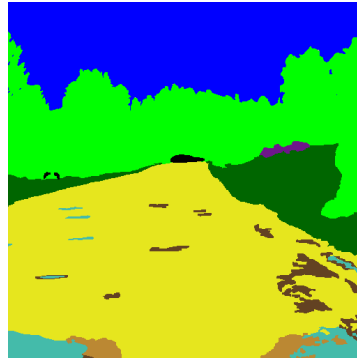
Figure 8. Places of the image where training sample pixels captured.

3.4 Results

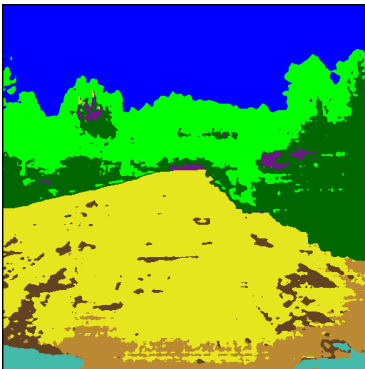
Below, Figure 9 shows the classification results for the terrain data set in Figure 4 using SS_CNN classification. The presented test image contains nine terrain classes. Three different band counts are used to create minimum band images and perform image classification. According to Figure 9 classification image results, the three different band variants show similar visual outcomes. Same image patch coordinates are used to extract training data in all the test cases.



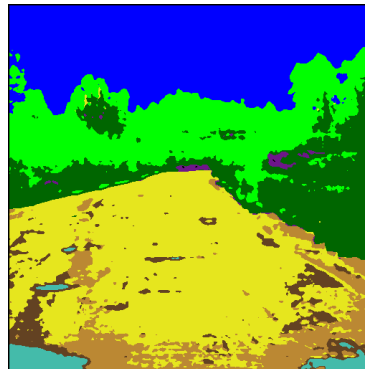
RGB colour composite image of HSI terrain dataset



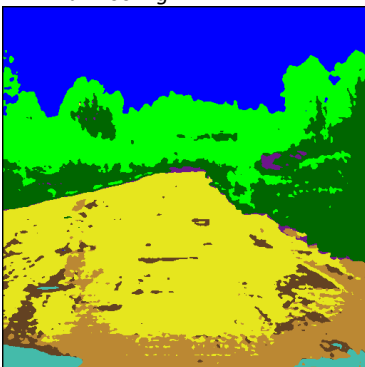
Ground truth of the dataset



Classification result for 9 bands using Min-Max Pooling



Classification result for 16 bands using Min-Max Pooling



Classification result for 25 bands using Min-Max Pooling

Figure 9. Ground truth and classification results for each band set.

Table 4. Classification pixel accuracy for different numbers of spectral bands.

Terrain class	9 bands		16 bands		25 bands		Legend
	Precision	Recall	Precision	Recall	Precision	Recall	
Undefined	0	0	0	0,02	0	0	
Grass	0,46	0,81	0,47	0,83	0,46	0,77	
Trees	0,95	0,63	0,95	0,65	0,94	0,67	
Rocks	0,28	0,65	0,31	0,49	0,25	0,50	
Water	0,98	0,66	0,94	0,65	0,99	0,66	
Sky	0,94	0,97	0,95	0,97	0,96	0,98	
Gravel	0,17	0,88	0,13	0,84	0,11	0,91	
Dirt	0,97	0,79	0,97	0,76	0,98	0,71	
Mud	0,28	0,55	0,25	0,52	0,28	0,54	
Macro avg	0,56	0,66	0,55	0,64	0,55	0,64	

The classification accuracy matrices were calculated using the below equations (5) and (6). The class-wise pixel accuracies and overall pixel accuracy are presented in Table 4. The overall accuracy of the dataset is affected by the class imbalance.

$$Precision = \frac{\text{True positives}}{(\text{True positives} + \text{False positives})} \quad (5)$$

$$Recall = \frac{\text{True positives}}{(\text{True positives} + \text{False negatives})} \quad (6)$$

One important fact is, ground truth labels for the selected dataset are not accurate for “dirt” and “mud” classes. Since these two classes are visually similar and the only difference between “dirt” and “mud” is that dirt is a dry terrain and mud is a wet terrain, it is indistinguishable for the human eyes by looking at RGB images. Considering both terrain classes are drivable terrains, the classification accuracy difference is acceptable.

The chosen dataset is classified using the SS_CNN method by selecting 16 spectral bands from two different band selection methods. The classification accuracy for the selected dataset using the min-max pooling and the distance density method is presented in Table 5.

Table 5. Classification pixel accuracy using different band selection methods.

Terrain class	Distance density method [91]		16 bands Min-max pooling	
	Precision	Recall	Precision	Recall
Undefined	0	0	0	0,02
Grass	0,28	0,61	0,47	0,83
Trees	0,82	0,44	0,95	0,65
Rocks	0,64	0,26	0,31	0,49
Water	0,98	0,63	0,94	0,65
Sky	0,94	0,98	0,95	0,97
Gravel	0,14	0,91	0,13	0,84
Dirt	0,95	0,79	0,97	0,76
Mud	0,35	0,42	0,25	0,52
Macro avg	0,56	0,56	0,55	0,64

The classification results for different band selection methods are shown in below Figure 10.

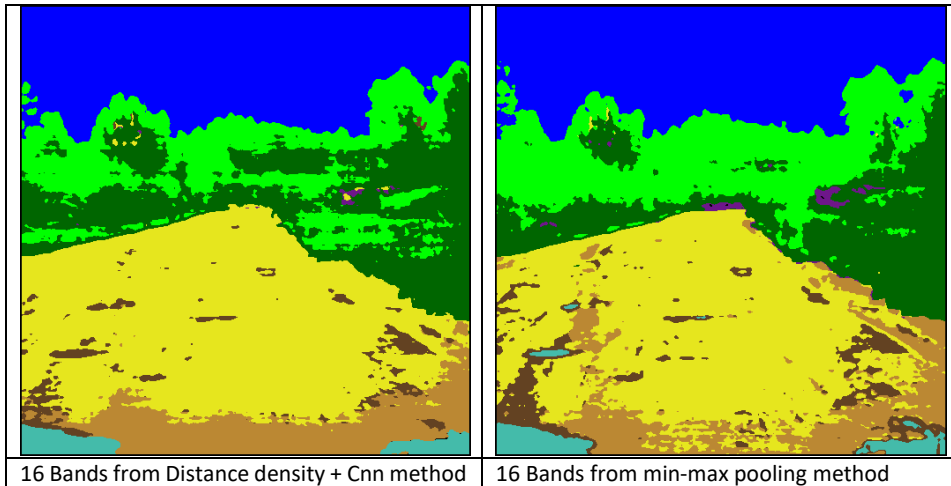


Figure 10. Classification results for different band selection methods.

The proposed band selection method gives a slightly better recall value than the other method. Most importantly, it gives higher classification average results for all the classes, even with fewer spectral bands. The objective of the method is to distinguish spectral signatures of closely correlated classes along with band selection, and this objective has been achieved with the above results.

3.5 Comparison with benchmark datasets from remote sensing

Popular remote sensing datasets, such as the Indian Pines dataset, are often used to compare the performance of different dimensionality reduction methods. The proposed pooling method in this study was used for the Indian Pines dataset dimensionality reduction [116], thus evaluating classification capabilities. The corrected Indian Pines dataset contains 145 x 145 pixels with 200 spectral bands, which exclude spectral bands containing water absorption spectral signature. The false colour representation of the dataset is shown in Figure 11.

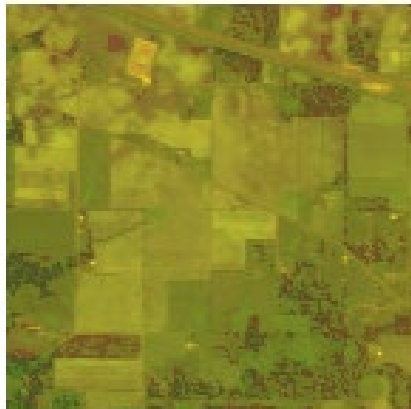


Figure 11. The false colour representation of the Indian pines dataset.

Due to the low resolution of the dataset, the amount of training data for all 17 classes were not sufficient in the original image. In order to minimise the impact of limited training data on the classification, the top nine classes are selected based on the number of spatial pixels available in each class. Table 6 gives the details of the classes in the Indian pines dataset. A batch of 144 pixels from each class are selected for band selection and classification. An equal pixel count from each class creates a balanced dataset. Training and validation samples are taken by splitting the above sample dataset. The ground truth for both the original image and with selected classes are shown in Figure 12. The non-selected classes are grouped with the “Background” class.

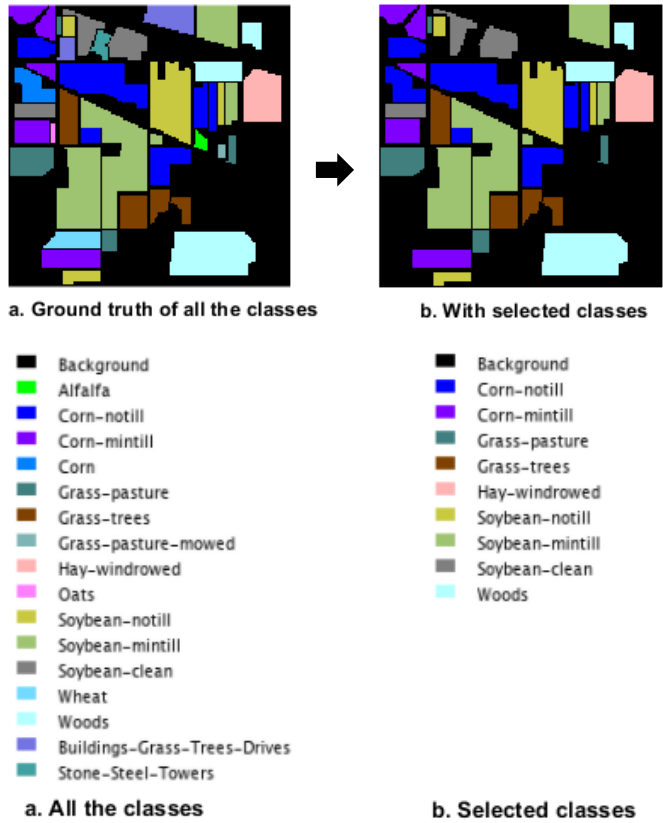


Figure 12. Ground truth of AVIRIS Indian Pines dataset from Purdue University [116].

Selected spectral bands for the Indian Pines dataset using Min-max pooling are as below in Table 7.

Table 6. Indian Pines dataset classes and the number of samples for classification.

#	Class	Total no of samples	No of training samples
1	Alfalfa	46	-
2	Corn-notill	1428	144
3	Corn-mintill	830	144
4	Corn	237	-
5	Grass-pasture	483	144
6	Grass-trees	730	144
7	Grass-pasture-mowed	28	-
8	Hay-windrowed	478	144
9	Oats	20	-
10	Soybean-notill	972	144
11	Soybean-mintill	2455	144
12	Soybean-clean	593	144
13	Wheat	205	-
14	Woods	1265	144
15	Buildings-Grass-Trees-Drives	386	-
16	Stone-Steel-Towers	93	-

Table 7. Selected spectral bands from the Indian Pines dataset.

	No of bands - d	Selected bands
1	9	9, 23, 48, 67, 90, 111, 133, 155, 177
2	16	7, 13, 29, 39, 51, 62, 74, 87, 98, 111, 123, 133, 152, 158, 170, 184
3	25	2, 11, 17, 26, 35, 43, 50, 58, 66, 75, 81, 91, 97, 105, 118, 122, 129, 137, 149, 153, 163, 170, 178, 181, 188

The training dataset shuffles at every epoch during CNN training. Classification results are shown in Table 8.

Table 8. The HSI classification with a different number of image bands.

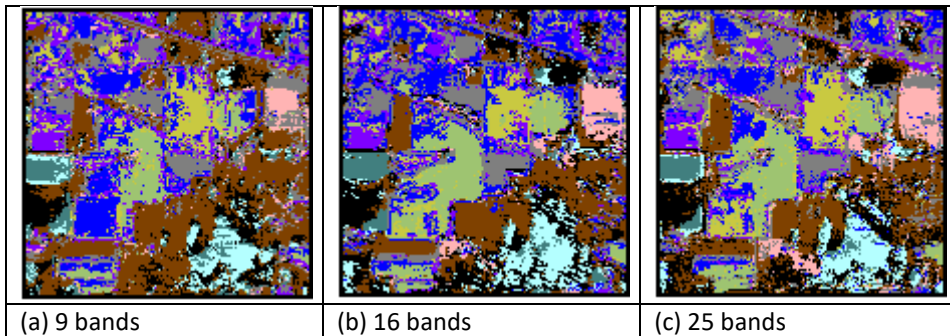


Table 9. Classification pixel-wise accuracy for the Indian Pines dataset based on a different number of bands.

Class	9 bands %	16 bands %	25 bands %
Corn-notill	23.24	34.35	27.57
Corn-mintill	28.70	29.21	26.51
Grass-pasture	38.67	47.22	33.97
Grass-trees	64.38	70.41	75.61
Hay-windrowed	38.28	44.35	75.10
Soybean-notill	36.60	46.84	60.80
Soybean-mintill	34.56	36.26	31.58
Soybean-clean	35.24	31.87	32.88
Woods	20.94	19.52	26.72

Above Table 9 presents the pixel-wise classification accuracy for each class in the dataset. The neural network methods require an enormous amount of labelled data which impacts the classification outcome. It is evident that classification accuracies are low due to the lack of samples taken for training and validation. The lower pixel resolution of the image is the reason for the smaller training dataset. It is noticeable that the classification accuracy increases along with the number of spectral bands for some terrain classes. However, it is not the same pattern for all the classes.

Furthermore, several classes are highly correlated with each other. Figure 13 shows the spectral signatures of all the terrain classes. Plot legend in Figure 13 follows the label colours in Figure 12 ground truth. In the case of less correlating spectral signatures, the proposed method is effective but not effective for highly correlating spectral signatures. The Min-max pooling method took 2.7 milliseconds to obtain the spectral bands.

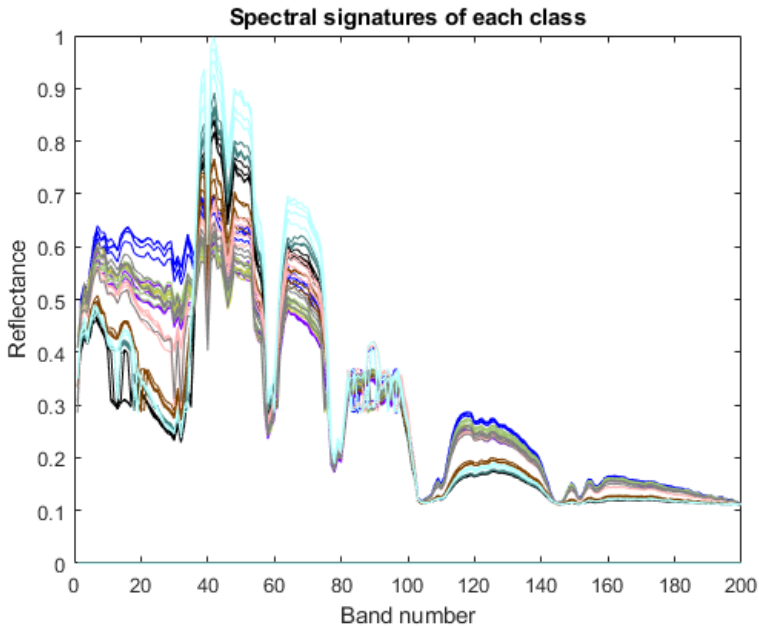


Figure 13. Spectral signatures of Indian Pines dataset classes.

3.6 Chapter summary

In this chapter, a simple yet effective band selection method is proposed for hyperspectral images. In contrast to various existing band selection methods, the proposed method is mathematically simple to implement. This method will ease the computation burden on the processing computer and thus be helpful for the deployment of an embedded computer with lower computing power. According to experiment results, the band selection method could identify the optimum spectral bands in 2.7 milliseconds. The effectiveness of the proposed method was presented for both terrain hyperspectral images and a remote sensing dataset.

4 RGB Image Generation from Hyperspectral Databudes

One of the objectives of this work is to investigate the possibilities of using the hyperspectral imaging method to prepare training inputs to train RGB semantic segmentation networks, which can eliminate labour intensive image annotation processes. Even though the HSI camera used in this research can produce RGB images using its built-in RGB camera, they are not in the same spatial resolution. Moreover, there can be vertical and horizontal alignment mismatches between the hyperspectral image and the RGB image. Therefore, it is necessary to generate RGB images from the HSI databcube. Below, Figure 14 illustrates the concept of using HSI classification as an input for the RGB perception systems.

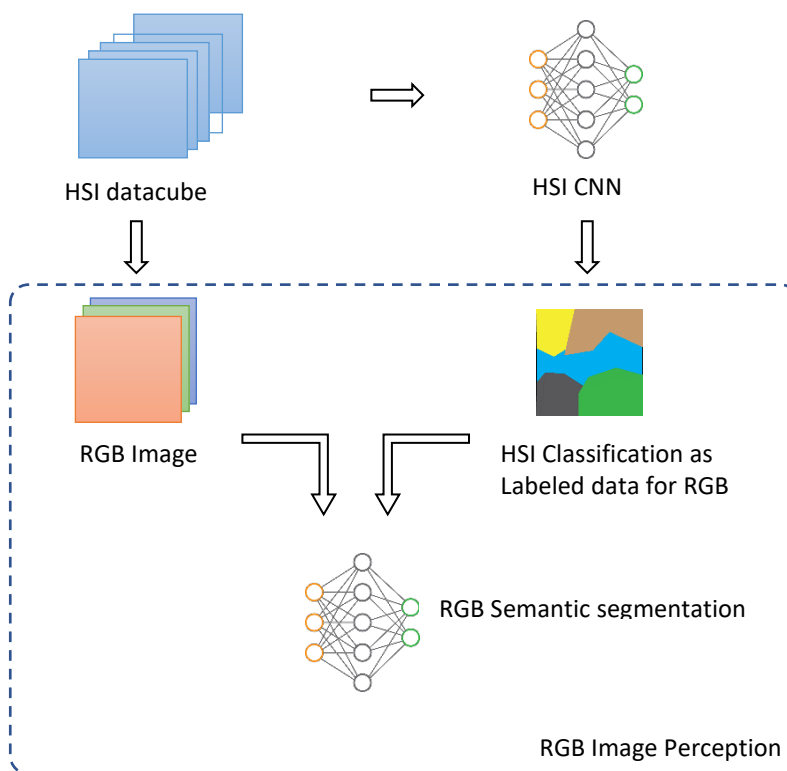


Figure 14. HSI classification as an input for the RGB semantic segmentation.

4.1 Methodology

Several trichromatic image visualization methods based on HSI databcubes were explored. One method was the manifold alignment technique, which uses HSI databcube and RGB images are captured from the same environment to transfer colour characteristics from RGB to HSI. The RGB image generation method uses band selection to form a smaller databcube. Subsequently, the manifold alignment method is used to create an RGB image. In this study, the min-max pooling method was used for band selection. Several other methods are also used in this research to determine which band selection performs

sufficiently close to the RGB representation of the scene. Several other RGB image generation methods have been implemented for the comparison of the chosen method.

Below, Figure 15 a and b show the test images used to implement the manifold alignment technique.



Figure 15 a. RGB image taken from the same site.



Figure 15 b. HSI datacube visualized using three image bands.

As in the above images, Figure 15 a. and Figure 15 b, the RGB image and HSI datacube coming from the Specim camera are not accurately aligned vertically and horizontally. The resolution of the RGB images is 645 x 645 pixels, while HSI datacube spatial resolution is 512 x 512 pixels. Therefore, the HSI classification result and RGB image from the Specim camera cannot implement the above proposed unsupervised image labelling due to the difference in resolution.

4.1.1 Manifold alignment RGB image generation method

The manifold alignment method for RGB image visualization from the hyperspectral image is based on reference [103]. This method use pixel pairs from HSI datacube and RGB image. The RGB image should have been taken from the same or semantically similar environments as the HSI datacube.

The concept of manifold alignment is projecting both image data to a shared embedding space with lower dimensions than the original data. This concept is illustrated in Figure 16. There are several methods to obtain the manifold structure of the underlying data. In this task, the natural colour image was obtained by using Locality Preserving Projections (LPP) method [117]. Here, it has been implemented feature-level manifold alignment with semi-supervised pixel-groups selection.

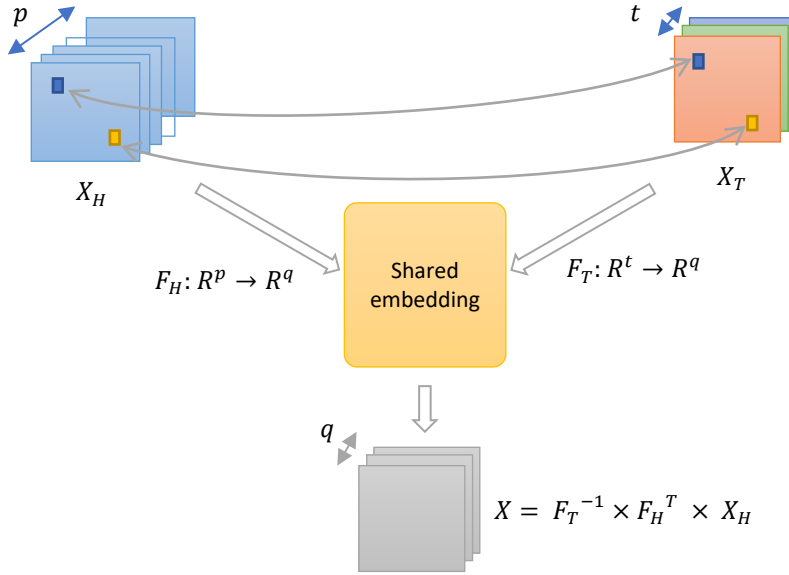


Figure 16. Manifold alignment for HSI and RGB images.

The explanation of the algorithm is as follows. The HSI dataset is represented as a matrix X_H , where $X_H \in R^{p \times n}$, p is the number of selected hyperspectral image bands, and n is the number of samples. The RGB image is represented as a matrix X_T where $X_T \in R^{t \times n}$, t is the number of colour channels which is three, and n is the number of samples. Since the pixels are selected as pairs from each image, the numbers of HSI samples n_S and n_R are equal to n . The objective of the LPP is to find two projection functions $F_H : R^p \rightarrow R^q$ and $F_T : R^t \rightarrow R^q$ where q is low dimensional space.

The locality preserving projections method keeps local neighbourhood relationships in each image. The adjacency graphs for each image G_H and G_T are constructed and obtained the weighted adjacency matrices for each image as $W_H(i, j)$, and $W_T(i, j)$ respectively. The weighted adjacency matrices' distance measures follow k-nearest neighbour. The hyperspectral image pixel distances are calculated as spectral angle (SA), while RGB image pixel distances were obtained using Euclidean distance.

The spectral angle between two hyperspectral pixels is calculated using (7). The i^{th} and j^{th} pixels of hyperspectral image datacube are denoted as $x_H(i)$ and $x_H(j)$.

$$SA_{(x_H(i), x_H(j))} = \cos^{-1} \left(\frac{\sum_{i,j=0}^n x_H(i) \cdot x_H(j)}{\sqrt{\sum_{i=0}^n x_H(i)^2} \cdot \sqrt{\sum_{j=0}^n x_H(j)^2}} \right) \quad (7)$$

The weight coefficient matrix for the HSI image is

$$w_H(i, j) = e^{\left(\frac{-SA_{(x_H(i), x_H(j))}}{\sigma^2} \right)} \quad (8)$$

The weighted adjacency matrix is a $n \times n$ matrix. The spectral angle cosine between two pixels in hyperspectral is the highest ($\cong 1$) when the two pixels are similar to each other.

$$W_H(i, j) = \begin{cases} 1, & \max w_H(i, j) \\ 0, & \text{Otherwise} \end{cases} \quad (9)$$

For the RGB image, i^{th} and j^{th} pixel colour vectors are denoted as $x_T(i)$ and $x_T(j)$ distance measures calculated using the below equation (10).

$$d_{i,j} = \sqrt{\sum_{i,j=0}^n (x_T(i) - x_T(j))^2} \quad (10)$$

The weight coefficients are calculated as

$$w_T(i, j) = e^{\left(\frac{-d_{i,j}}{\sigma^2}\right)} \quad (11)$$

Then the weighted adjacency matrix is obtained as follows.

$$W_T(i, j) = \begin{cases} 1, & \min w_T(i, j) \\ 0, & \text{Otherwise} \end{cases} \quad (12)$$

Similar to the HSI image weighted adjacency matrix, the RGB image weighted adjacency matrix is also an $n \times n$ matrix. In the case of RGB images, the smaller the Euclidean distance, the similar the pixels to each other. Therefore, the minimum distance is used to obtain the weighted adjacency matrix.

The correspondence matrix W_{HT} represents matching pairs between two images. If the selected pixels from each image space form a matching pair, the correspondence is 1, otherwise 0. The correspondence matrix is obtained from the below expression (13).

$$W_{HT}(i, j) = \begin{cases} 1, & \text{if } X_H(i) \text{ and } X_T(i) \text{ forms a matching pair} \\ 0, & \text{Otherwise} \end{cases} \quad (13)$$

Since an equal number of pixels were selected from both images, the correspondence matrix W_{HT} is also an $n \times n$ matrix.

The objective function to calculate the projection functions can be derived as below.

$$\begin{aligned} E(F_H, F_T) &= \sum_{i,j=0}^n \|F_H^T x_H^i - F_H^T x_H^j\|^2 \times \alpha_1 W_H(i, j) \\ &+ \sum_{m,n=0}^n \|F_T^T x_T^m - F_T^T x_T^n\|^2 \times \alpha_1 W_T(m, n) \\ &+ \sum_{i,m=0}^n \|F_H^T x_H^i - F_T^T x_T^m\|^2 \times \alpha_2 W_{HT}(i, m) \end{aligned} \quad (14)$$

The joint weighted matrix W is obtained below to minimize the objective function.

$$W = \begin{bmatrix} \alpha_1 W_H & \alpha_2 W_{HT} \\ \alpha_2 W_{HT}^T & \alpha_1 W_T \end{bmatrix} \quad (15)$$

Then, the objective function becomes (16).

$$E(F) = \sum_{i,j=0}^n \|F^T x^i - F^T x^j\|^2 \times W(i,j) \quad (16)$$

The diagonal matrix D is defined as (17).

$$D_{i,i} = \sum_{i=0}^n W(i,j) \quad (17)$$

Then, L is the Laplacian matrix which is $L = D - W$. The minimization of the above (16) has been proved in [103] to be equivalent to (18). Thus, it can be solved as a generalized eigenvalue problem.

$$\arg \min \text{trace}(F^T X L X^T F) \quad (18)$$

$$X L X^T F = \lambda X D X^T F \quad (19)$$

Joint projection matrix F is (20)

$$F = \begin{bmatrix} F_H \\ F_T \end{bmatrix} \quad (20)$$

And X is defined as (21)

$$X = \begin{bmatrix} X_H & 0 \\ 0 & X_T \end{bmatrix} \quad (21)$$

The smallest Eigenvalues λ matrix provide the optimum manifold projection functions. For the shared low dimensional embedding space, the number of dimensions q is selected as 3, which is for three RGB channels.

The pixels were selected from the same regions of both images as below Figure 17. Regions were defined from several parts of the image and randomly extracted N number of samples from each region.

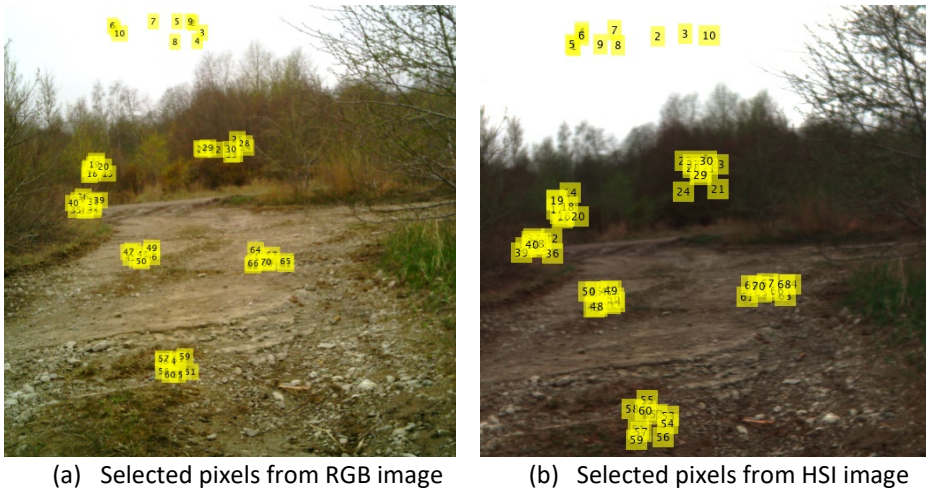


Figure 17. Pixel-pair regions for both RGB and HSI.

4.2 Results for RGB image generation

For the image generation, only visible light wavelength range HSI bands can be used. The below RGB image was generated using a few spectral bands selected using band selection methods described in the previous chapter. Moreover, it further narrowed down to 9 spectral bands in the visible light wavelength range.



Figure 18. RGB representation using manifold alignment method.

Several methods other than manifold alignment have been experimented with to compare the RGB image visualization accuracy for the HSI datacube. The compared HSI visualization methods are bilateral filtering, using band selection with correlation measure, and three spectral bands representing the highest intensities in RGB space.

The colour representation has been evaluated using two comparison matrices. They are root mean square error (RMSE) and correlation coefficients between each colour plane.



RGB image captured from the same location.



Bilateral filtering method.



Manifold alignment method using nine spectral bands.



RGB image generated from three spectral bands where each colour channel has peak intensity.

Figure 19. Generated RGB visualization images from various methods.

One colour representation accuracy evaluation technique is each colour channel's root mean square error [118]. The RMSE between the generated RGB image and the original RGB image captured from the same location has been calculated using equation (22). The $img1$ and $img2$ are RGB images with $[m, n]$ spatial resolution. The lower the RMSE value, the similar the image colour channel of both images.

$$RMSE = \sqrt{\frac{1}{MN} \sum_{m,n=0}^{M,N} [X_{img1} - X_{img2}]^2} \quad (22)$$

The RMSE estimates for each colour channel are tabulated in Table 10.

Table 10. RMSE for each image generation method.

RMSE	Three image bands from the HSI cube	Manifold alignment	Bilateral filter
R	0.0098	0.0508	0.0156
G	0.0605	0.0117	0.0684
B	0.0137	0.0117	0.0254
Average	0.0280	0.0247	0.0365

Even though the red colour channel of the RGB image generated using the manifold alignment method shows the highest RMSE than the others, the average RMSE for all three colour channels is smaller. Thus, the manifold alignment method gives better visual representation over the other two methods.

The second evaluation criterion is Pearson's correlation coefficient of the two images. The Pearson's correlation coefficient is calculated using the below equation (23).

$$C = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2 \sum(Y - \bar{Y})^2}} \quad (23)$$

The two images are denoted as X and Y where \bar{X} and \bar{Y} represent the mean values of the images. Pearson's Correlation coefficient for each colour channel between visualized image and RGB image captured from the same site is presented in Table 11 below.

Table 11. Pearson's correlation coefficients for each colour channel.

Correlation Coefficient	Three image bands from the HSI cube	Manifold alignment	Bilateral filter
R	0.7226	0.7326	0.7167
G	0.7494	0.7848	0.7430
B	0.7944	0.8419	0.8028

The Pearson's correlation coefficient is an indication of similarities between two different datasets. The higher the correlation coefficient, the similar the images. The value of the correlation coefficient can be within the range of -1 to 1. The above analysis for the selected image confirms that the manifold alignment method gives the highest correlation coefficient for all three image channels in the RGB image inputs.

The optimum number of HSI bands for the manifold alignment method

The HSI datacube contains 204 spectral bands, and the RGB image contains three colour channels. The number of image bands used for the manifold alignment algorithm influences the clarity of the RGB visualization from the HSI image. Several spectral band combinations were used to investigate the appropriate band count to form an RGB image reflecting natural colours. The min-max pooling method has been used for spectral band selection, which was a part of this research. All the image bands in the datacube, then 9, 16, and 25 spectral bands constituting the entire wavelength range, were used for the experiments. Furthermore, those image bands within the visible light wavelength range are used to determine the suitability of VNIR (400 – 1000 nm) or VIS (400 – 750 nm) wavelength range for the RGB generation. Below, Figure 20 shows several combinations of HSI to RGB image formation with various spectral band options.



9 spectral bands



16 spectral bands



25 spectral bands



204 spectral bands (entire datacube)

Figure 20. RGB visualization of HSI datacube in VNIR range bands.

The above images were generated by taking the spectral bands within the 400 – 1000 nm wavelength range. With all 204 image bands, the manifold learning method cannot successfully reconstruct an RGB image because Eigenvector decomposition yields complex and negative Eigenvalues for the optimization function.

Table 12. Pearson’s correlation coefficient of RGB images based on the number of spectral bands.

Correlation Coefficient	9 bands	16 bands	25 bands
R	0.7326	0.6710	0.6673
G	0.7848	0.7343	0.7288
B	0.8419	0.8228	0.8252
<i>Average</i>	0.7864	0.7427	0.7404

Above Pearson’s correlation coefficients imply that the images with the lowest spectral band counts provide the most similar RGB image representation.

The influence of NIR image bands on RGB visualization has been made by extracting only visible light spectral bands and forming RGB images using those bands. The visible light wavelength ranges from 380 – 750 nm, and hence, spectral bands within 400 – 750 nm have been used to form below RGB images in Figure 21.

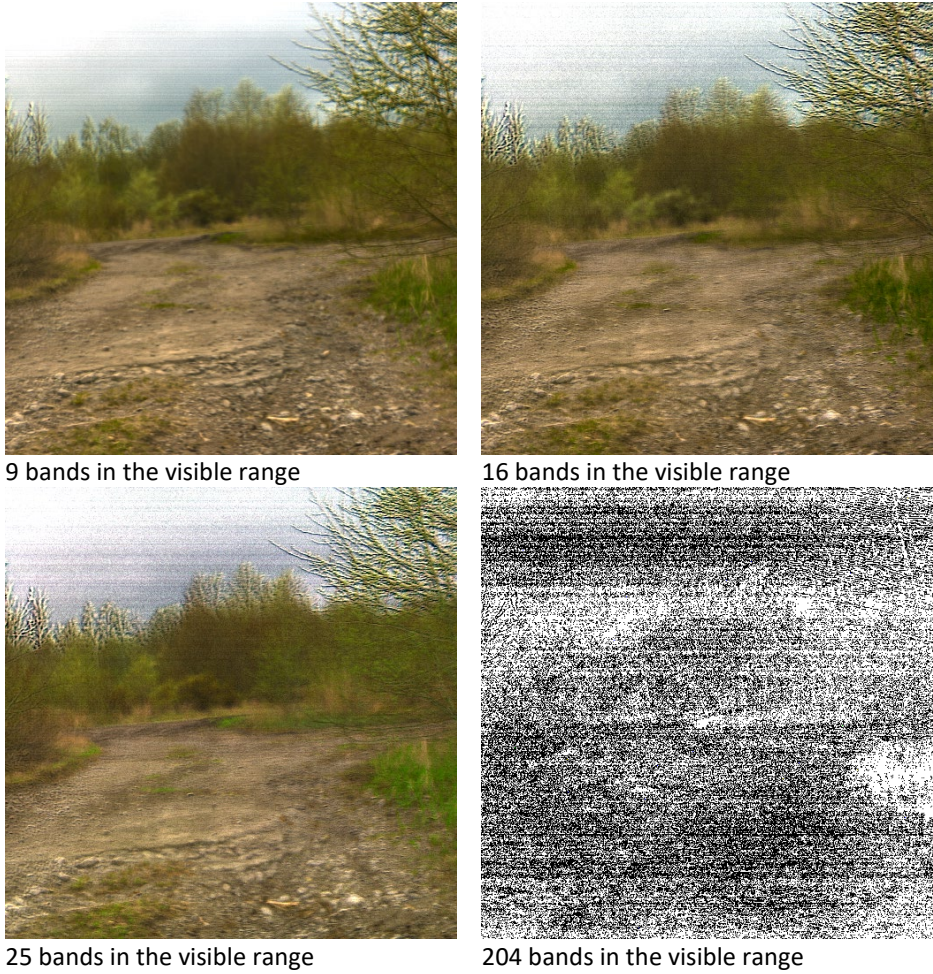


Figure 21. RGB image generation from the visible light range bands.

Table 13. Pearson's correlation coefficient for the RGB images generated by using visible light spectral bands.

Correlation Coefficient	9 bands	16 bands	25 bands
R	0.6994	0.6576	0.6528
G	0.7746	0.7387	0.7373
B	0.8314	0.8163	0.8146
Average	0.7685	0.7375	0.7349

Comparing Table 12 and Table 13 shows that the entire VNIR spectral range image bands carry more information than visible light wavelength bands alone for manifold alignment.

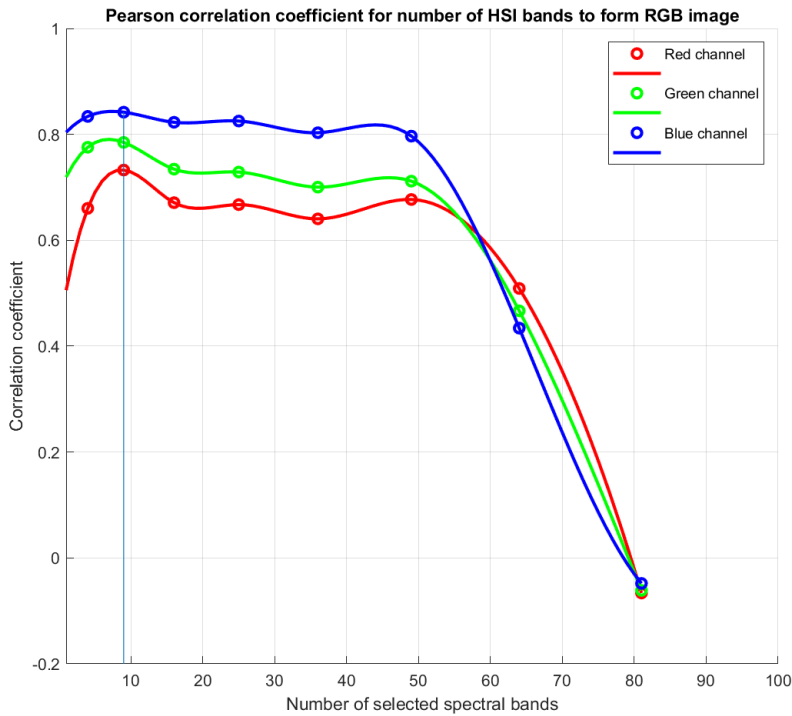


Figure 22. Pearson's correlation coefficient for each colour channel between original RGB and generated RGB from HSI.

Based on above Figure 22, it is evident that the optimum number of spectral bands from the HSI datacube should be nine bands for the manifold alignment method. Below Table 14 consists of RGB images generated using the chosen nine-band HSI datacubes of off-road terrain scenes. The correlation coefficients for the generated RGB images are presented in Table 15.

Table 14. Unstructured terrain HSI datacubes visualized in RGB using manifold alignment.













Image	RGB image from the location	HSI visualized in RGB using three image bands	HSI visualized in RGB using manifold alignment
1			
2			
3			
4			

Table 15. Correlation coefficients for above Table 10 image set.

Image	Correlation coefficient			
	Red channel	Green channel	Blue channel	Average
1	0,1209	0,2706	0,5707	0,3207
2	0,4691	0,5825	0,7041	0,5852
3	0,4930	0,5894	0,7396	0,6073
4	0,6458	0,7280	0,7702	0,7147

4.3 Chapter summary

Several different methods were explored to generate RGB images from HSI datacubes. The objective of this RGB image generation was to use them with RGB semantic segmentation CNN networks instead of high-resolution RGB images captured from the same locations.

The techniques for RGB image generation are bilateral filtering, selecting three different channels from each R, G and B region from the visible light wavelength range and manifold alignment. Among those three methods, the manifold alignment method using locality preserving projections yields the optimum similarity for HSI visualization as an RGB image. The experiment results show that the optimum number of spectral bands for RGB image generation was nine spectral image bands. The Person's correlation coefficient for the 9bands image set was 0.7685.

The datacube contains 204 spectral bands, which contain all the information about the scene. However, 204 image bands show that it cannot project all the data points to a common embedding. The use of the entire hypercube resulted in poor image quality compared to a lower number of spectral bands. Furthermore, it is evident that including the NIR range spectral bands of HSI datacube for the manifold projection gives a better correlation than the images generated from image bands taken from visible light wavelength range alone. A lesser number of image bands requires less time to process the image and hence improves the overall efficiency of the image generation process.

The average correlation coefficient could be further improved by increasing the number of pixel pairs from both images.

5 Unstructured Terrain Semantic Segmentation

5.1 Methods

The problem of scene understanding has been addressed using various methods from the beginning of computer vision. Starting from colour threshold methods in the early '90s to modern DCNNs, semantic segmentation methods improved gradually. Most of these DCNNs are developed for a specific application. For instance, the U-Net architecture was proposed for biomedical image segmentation with a few training images in the dataset [60]. The RefineNet was proposed by Lin et al. from the University of Adelaide, which improved a common drawback resulting from previous semantic segmentation DCNNs with output image blur [119]. Even though they perform well with one type of segmentation problem, the same network may perform poorly with another classification task area, such as terrain segmentation.

This comparison study selects several state-of-the-art semantic segmentation networks. They are SegNet, U-Net and DeepLabv3+ with Resnet18. The DeepLab V3+ DCNN architecture uses an encode-decoder structure to extract object boundaries while recovering most spatial features. With atrous separable convolution operation, this architecture can capture information from a larger field of view [120]. Since DeepLab V3+ is one of the most effective DCNN architectures in semantic segmentation, it has been chosen to perform terrain image semantic segmentation for RGB cases.

The overall deep learning-based image segmentation will comprise four different combinations.

- RGB images 645 x 645 px resolution with manually annotated labels
- RGB images generated from hyperspectral image datacubes and manually annotated labels
- RGB images generated from hyperspectral datacubes with hyperspectral image classification result as labels. HSI classification has been done with spectral data alone
- Spectral-Spatial classification with manually annotated labels

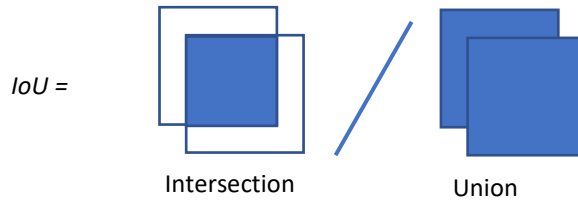
5.2 Performance matrices

Several accuracy matrices were used to quantify the accuracy of semantic segmentation. The accuracy matrices are pixel-wise accuracy, intersection over union (IoU) and F1 score.

TP-true positive, TN – true negative, FP – false positive and FN-false negative

$$\text{Overall accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (24)$$

$$\text{F1 score} = \frac{2TP}{2TP + FN + FP} \quad (25)$$



$$IoU = \frac{\text{Overlap}}{\text{Union}}$$

$$IoU = \frac{TP}{TP + FN + FP} \quad (26)$$

5.3 Image datasets

For semantic labelling, twelve terrain classes were selected. In comparison to existing similar public off-road RGB datasets, a similar number of classes were used in this research. The man-made objects are grouped into the “objects” class. Moreover, the terrains with very few occurrences are categorized as “undefined” along with the other unknown classes. The shadows of various constructions and trees are also annotated as “undefined”.

The complete list of terrain classes used in this research with label colours is shown in Table 16.

Table 16. List of terrain classes.

	Class	Red	Green	Blue	Hex code	Label Colour
1	Undefined	0	0	0	#000000	
2	Grass	0	102	0	#006600	
3	Concrete	170	170	170	#aaaaaa	
4	Asphalt	64	64	64	#404040	
5	Trees	0	255	0	#00ff00	
6	Rocks	110	22	138	#6e168a	
7	Water	68	187	170	#44bbaa	
8	Sky	0	0	255	#0000ff	
9	Gravel	187	136	51	#bb8833	
10	Object	192	64	64	#c04040	
11	Dirt	108	64	20	#e6e61e	
12	Mud	99	66	34	#634222	

Approximately 500 images were captured during the study. Since extensive manual work was required for labelling, 152 images were initially labelled for HSI and RGB datasets, respectively. The development of the dataset continues in order to improve semantic segmentation accuracy.

The entire dataset contains carefully annotated hyperspectral images of terrain scenes. For DCNN training, validation and testing, the dataset was split into three parts. The number of training images was 122, which is 80 % of the image dataset, while 15 images were used for testing and 15 used for validation which constitutes 10 % for each set. Since the number of images in the dataset is small for DCNN training, the dataset was augmented to increase the number of images for training. Figure 23 represents the constitution of each terrain class in the dataset. Since DCNN models require a large dataset for training, the dataset is augmented by transformations.

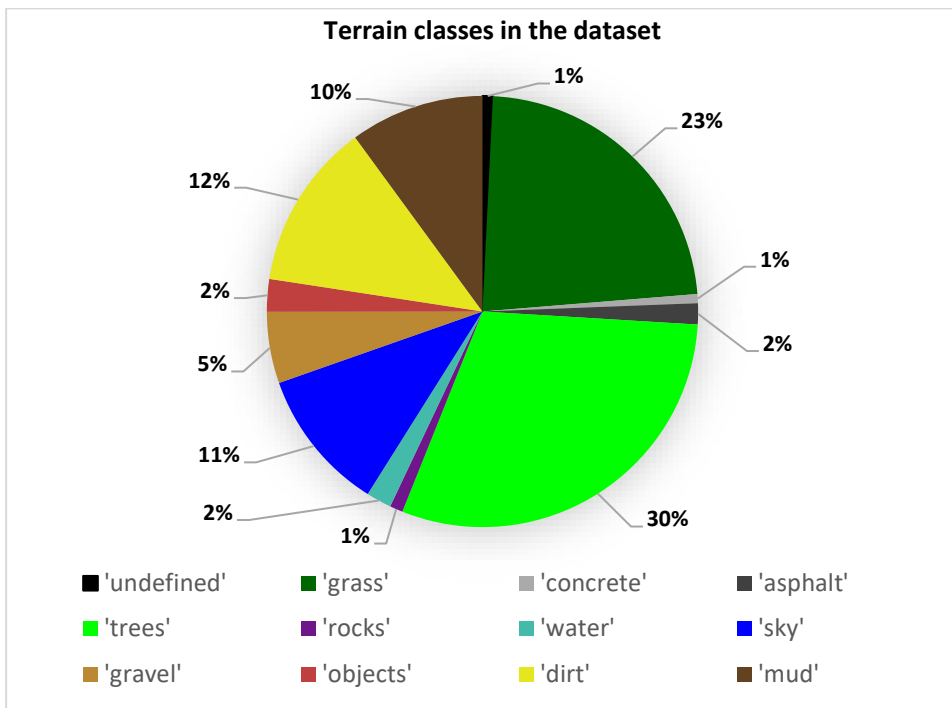


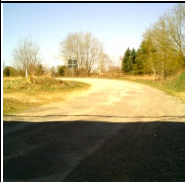





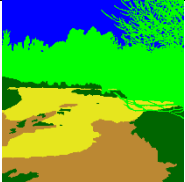
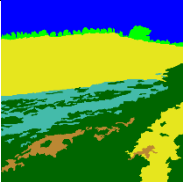
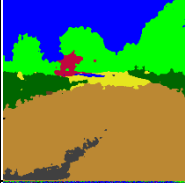
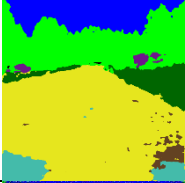
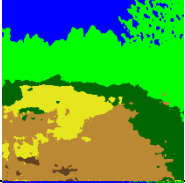

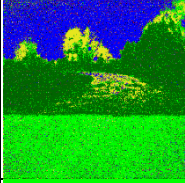
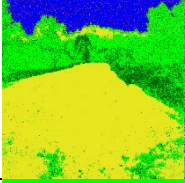
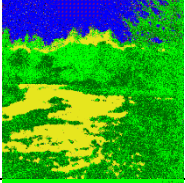
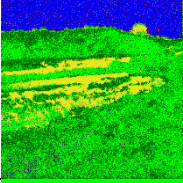
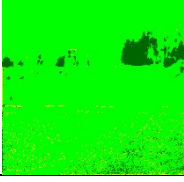
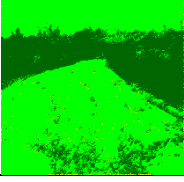
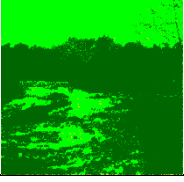
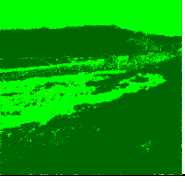
Figure 23. The composition of the terrain dataset.

5.4 Terrain segmentation using RGB images and manually labelled ground truth

State-of-the-art RGB image semantic segmentation networks were employed to compare the terrain segmentation results of the HSI classification with the RGB dataset derived from the HSI. The semantic segmentation DCNNs were trained using hand-annotated labels and RGB images, followed by an evaluation of the classification accuracy. The images were taken from Specim IQ camera with 645 x 645 pixels resolution in RGB colour space. These images were simultaneously captured while acquiring the HSI datacubes from the terrain scene.

Three different most popular semantic segmentation architectures were chosen for the experiments. They are Deeplab V3+ with ResNet18, SegNet and U-Net with three colour channels. The DCNN architectures were used with a depth of three (3 colour channels) in the input layer. The input images were resized to 640 x 640 pixels for the U-Net network while other networks were fed with 645 x 645 pixels resolution three-channel images.

Table 17. Sample images from the RGB image dataset and classification results.

	Image 1	Image 2	Image 3	Image 4
Input RGB				
Ground truth				
Deep lab V3+ Resnet18				
SegNet				
U-Net				

The validation accuracy for the dataset yielded 80,7 % using the ResNet18 model, while other models gave much worse results. Table 18 summarizes the classification accuracy for each model, while Table 17 shows a part of the dataset classification results.

Table 18. Segmentation accuracy for higher resolution RGB images.

Model	Mean IoU	Mean F1 score	Mean accuracy	Global accuracy
Deep lab V3+ Resnet18	0,5112	0,4832	0,6985	0,8068
Segnet	0,1340	0,2196	0,2010	0,4412
U-Net	0,0999	0,1756	0,0891	0,2501

With the above results, it is evident that the Deeplab v3+ ResNet18 model has higher accuracy compared to the other CNN architectures.

5.5 Terrain segmentation with RGB images generated using HSI datacubes and manual labels

In this section, terrain semantic segmentation based on the RGB image dataset generated using HSI datacubes will be discussed. As explained in the previous chapter, RGB image semantic segmentation DCNNs are faster than multi-band spectral image classification. Therefore, RGB based segmentation along with HSI segmentation can achieve better efficiency in developing such classification pipelines.

Training RGB image dataset was generated using the manifold alignment method, which uses spectral images and high-resolution RGB images captured from the same location. The number of bands used in this research is nine spectral bands. As mentioned in the preceding chapter, the nine bands provide the highest correlation to the RGB images. Manually annotated labels were used for training.

Table 19. Segmentation accuracy matrices.

Model	Mean IoU	Mean F1 score	Mean Accuracy	Global accuracy
Deeplab V3+ resnet18	0,3313	0,4013	0,4488	0,7057
Segnet	0,2812	0,2156	0,1869	0,4448
U-Net	0,2784	0,2179	0,1875	0,4784

Table 20 shows the input image set and the classification results. Overall validation accuracy for the RGB image dataset generated using HSI achieved 70.6 % with the Deeplab v3+ with ResNet18 model. Table 19 shows the accuracy of the segmentation.

The segmentation accuracy is considerably low for the SegNet and U-Net models. However, the overall pixel accuracy of this classification is slightly higher than the RGB image classification in the previous section for the SegNet model. Moreover, U-Net gives better overall pixel accuracy. Due to class imbalance in the terrain dataset is so significant that the overall pixel accuracy matrix does not paint a correct picture of the classification performance. The above sample images prove that the mean IoU gives a better overview of the segmentation accuracy.

Table 20. Sample images from the dataset and results.

	Image 1	Image 2	Image 3	Image 4
Input RGB				
Ground truth				
Deeplab V3+ ResNet 18				
SegNet				
U-Net				

5.6 Terrain semantic segmentation based on the RGB images generated from the HSI visualization and pixel-wise classification

The pixel-wise HSI image classification is the key component of this experimentation to evaluate its effectiveness as an image annotation tool to reduce the effort for labelling. A single image of 512x512pixels annotation takes more than an hour as an average for dataset preparation, making it laborious work.

Similar to the image classification in chapter three, the SS_CNN method is employed to classify the image pixels. The training data was captured from each hyperspectral image in the same fashion as explained in chapter three. Compared to RGB image semantic segmentation, the pixel-wise classification displays significant details in the output. According to Figure 24, the input RGB image is a complex terrain scene containing a water stream with muddy surroundings and floating algae on the water surface. The pixel-wise classification captured the classes in the image with higher contrast than RGB semantic segmentation.

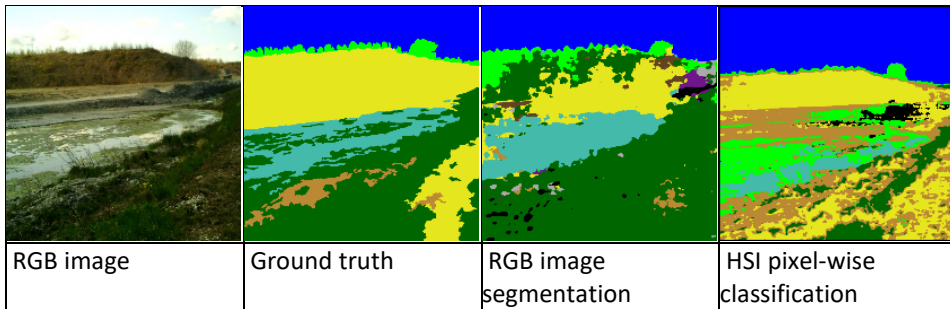


Figure 24. Pixel-wise classification and RGB semantic segmentation comparison.

However, the pixel-wise classification resulted in misclassifications as well. By applying careful manual corrections, the classification can be further refined to be used as an input for CNN training. As mentioned before, it is expected to reduce the human effort for dataset preparation by minimising labelling. The pixel-wise classification takes an average of 9.2 minutes to classify a 512x512px HSI datacube using the aforementioned SS_CNN method. And it is needed to denote the training pixel patches in the image, which takes approximately three minutes. After pixel-wise classification, it takes approximately six minutes for manual corrections. The proposed HSI-based method can complete image annotation in less than 20 minutes, which is 60% time saving for image annotation.

To validate the hypothesis, semantic segmentation CNN training and testing was performed in two ways. One of them by taking the pixel-wise classification result as image labels without refining them further. The second experiment was by using the manually refined pixel-wise classification result for CNN training. The RGB images were obtained from the manifold alignment method by using the HSI datacubes. The number of spectral bands for the manifold alignment method was nine, and the optimum bands were selected using the min-max pooling technique.

Even though RGB image generation was achieved with nine bands, the pixel-wise classification used 25 band HSI images. The same band selection method was used to create 25 band images, and it was observed that increasing the number of bands increased the pixel-wise classification accuracy.

Deeplab V3+ network with Resnet18 backbone used for semantic segmentation evaluation. The third row of Table 21 shows the pixel-wise classification results. The fourth row of the same table shows the semantic segmentation results when the pixel-wise classification is used as labels for CNN training. Because the pixel-wise classification result was used as training labels without further refinement for the semantic segmentation CNN, the overall segmentation output has been affected similarly. The last row of the same table shows the semantic segmentation outcome when it uses refined pixel-wise classifications. It is evident that from Table 22, the manual touch-up for pixel-wise classification could greatly enhance the semantic segmentation results. With this dataset, it was a 6% improvement over un-refined pixel-wise classifications.

As mentioned previously, this result was obtained from a small dataset. By increasing the number of images in the dataset, the result could be further improved.

Table 21. Sample images from the dataset and results.






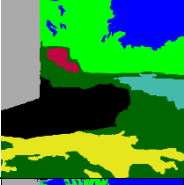
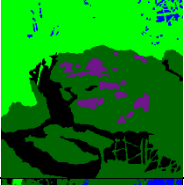
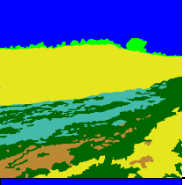
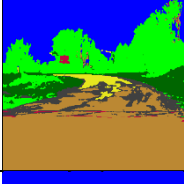
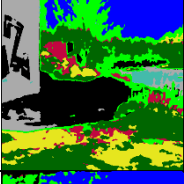
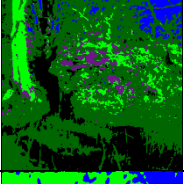
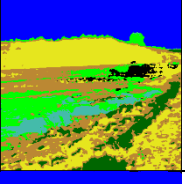
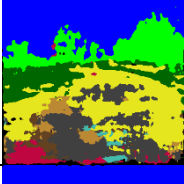
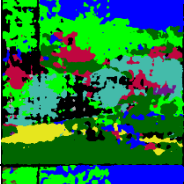
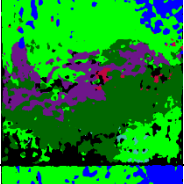

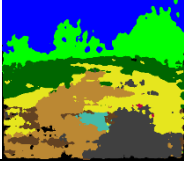
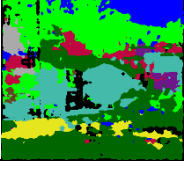
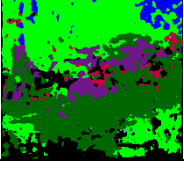

	Image 1	Image 2	Image 3	Image 4
Input RGB				
Ground truth				
Spectral-spatial CNN pixel-wise classification				
Deeplab V3+ Resnet18 segmentation – Before touch-up				
Deeplab V3+ Resnet18 segmentation – After touch-up				

Table 22. The classification performance of HSI generated images and annotations.

Model	Mean IoU	Mean F1 score	Mean Accuracy	Global accuracy
Without touch-up	0,2733	0,2968	0,4363	0,6206
After touch-up	0.3301	0,3073	0,4733	0,6688

5.7 Spectral-spatial image classification

In contrast to the spatial images for semantic segmentation, the spectral images contain more information, classifying the images with higher accuracy than the RGB input images. Since spectral-spatial combination can extract the spectral signature, object shapes, colours, etc., the classification could yield higher accuracy. Therefore, in this section, several spectral image datasets were used by selecting different numbers of spectral bands and used for DCNN semantic segmentation. The spectral band counts of the created datasets were 9, 16 and 25. The selection was according to the previously mentioned criteria, where the band count forms a square matrix. Such a square matrix

can help construct a multi-spectral sensor with band-pass filters deposited in a mosaic pattern for future development. The band selection method was the min-max pooling method proposed by this thesis.

5.7.1 Classification method

The U-Net architecture has been modified to accommodate the different spectral bands in the input layer, such as 9, 16 and 25. The chosen U-Net architecture has been developed for biomedical image classification with multi-layer images, appropriate for this task.







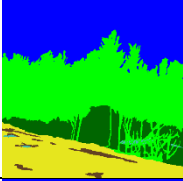

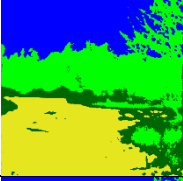
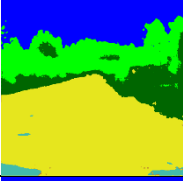
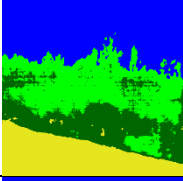
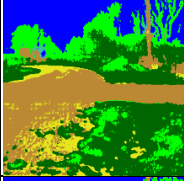
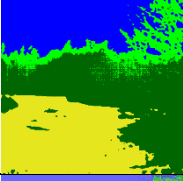
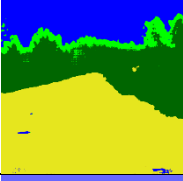
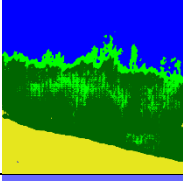
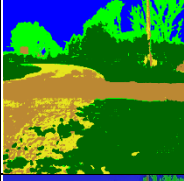


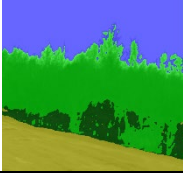

In contrast to the RGB image datasets, spectral image datasets are not widespread and not widely available as open-sourced datasets. Therefore spectral dataset used in this study has been collected and annotated by the author. The hyperspectral datasets containing off-road terrain scenes were snapped in multiple locations of Estonia. The dataset was captured under sunlight. In some instances, the sunlight intensity was considerably high, which caused image saturation. The reason was the integration time limitation of the camera, which is limited to a minimum of 1 ms, thus reducing further the integration was not an option. A neutral density filter was used to solve the light intensity issue. The used ND filter was ND3 – 400 variable neutral density filter. Colour casting appeared on the image due to this filter use, which was not corrected as it is an additional pre-processing step. The objective of the classification is to use DCNN to reduce additional pre-processing steps.

5.7.2 Results of spectral-spatial classification

Here are the classification results for 4 test datacubes selected from the test dataset. The hyperparameters for the DCNN training were, learning rate 0,001, mini-batch size of 4 and the stochastic gradient descent method was used as the optimizer. Due to the large size of the images and limited memory capacity of the processing hardware, smaller mini-batch size was selected.

The spectral image classification results are presented in Table 23 and Table 24. Validation accuracy for the nine bands image dataset was 70,85 %, 16 bands dataset with 73,2 % and 25 bands dataset classification accuracy resulted in 67,35 %.

Table 23. Sample images from three different datasets of 9, 16 and 25 bands HSI images with classification results.

	Image 1	Image 2	Image 3	Image 4
Input Image in RGB				
Ground truth				
HSI 9 Bands				
HSI 16 bands				
HSI 25 bands				

With U-Net architecture, the pixel classification accuracy is lower for the spectral images than the RGB images semantic segmentation using Deeplab v3+ ResNet18. The highest classification accuracy for the spectral dataset was 70,00 % for the nine bands dataset, while Deeplab v3+ Resnet18 for RGB images achieved 80,6 %. However, the mean IoU is higher for HSI images of 25 band DCNN, which achieved 62,6 % accuracy, and this is approximately 6 % higher than the next highest performing HSI DCNN model.

Table 24. Semantic segmentation accuracy for different numbers of image bands.

No of Bands	Mean IoU	Mean F1 score	Mean accuracy	Global accuracy
9	0,5613	0,4650	0,3739	0,7085
16	0,3065	0,4532	0,3911	0,7315
25	0,6226	0,4981	0,3875	0,7574

Hyperspectral imaging is highly vulnerable to environment light variations. In off-road conditions, the sunlight variation due to weather and the time of the day can result in poor image quality.

5.8 Chapter summary

In this chapter, various methods for HSI classification have been explored along with RGB segmentation. Comparing RGB segmentation using Deeplab V3+ ResNet18 to spectral-spatial classification with U-Net architecture demonstrates that HSI classification is 5 % higher accurate than RGB in terms of mean IoU. One of the challenges in DCNNs is gradient vanishing with deeper layers.

The segmentation experiments based on RGB images generated from HSI datacubes show that the classification accuracy is approximately 17 % inferior to the RGB counterpart.

However, RGB semantic segmentation still shows better pixel accuracy than HSI based methods. The demonstrated classification results were achieved using a few hundred training pixels from a few HSI data cubes. The HSI pixel classification takes very little time compared to the manual pixel-wise annotation of an RGB image. Thus the use of HSI pixel-wise classification results as labelled data shows significant improvement in dataset preparation. Some classes mixed up with others, such as grass detected on the top edges of the tree line and wet-fallen leaves mixed up with dry grass, which degraded the classification of those terrain classes. These artefacts can be removed by post-processing. However, for more complex terrains, the classification method needs to be improved in future. The obtained HSI classification result can be used as pixel-wise annotated label images for neural networks. The training images for the neural network can be obtained from HSI data cube as false-RGB images. Overall image classification results are summarised in Table 25. In conclusion, with the help of hyperspectral imaging, pixel-wise classification can be used to reduce the manual labelling process.

Table 25. Overall classification results comparison.

Dataset		Mean IoU	Mean F1 score	Mean accuracy	Global accuracy
RGB images + manual labels		0,5112	0,4832	0,6985	0,8068
Generated images + manual labels		0,3313	0,4013	0,4488	0,7057
Generated images + pixelwise classification	Without touch-up	0,2733	0,2968	0,4363	0,6206
	After touch up	0,3301	0,3073	0,4733	0,6688
HSI	HSI 9 bands	0,5613	0,4650	0,3739	0,7085
	HSI 16 bands	0,3065	0,4532	0,3911	0,7315
	HSI 25 bands	0,6226	0,4981	0,3875	0,7574

6 Conclusion

The overall objective of this work was to investigate the capabilities of hyperspectral imaging methods to enhance the performance of perception systems used in autonomous ground vehicles. The emphasis was on unstructured terrain scenarios, which is a highly active research area in mechatronics. The same methods could apply to any unknown terrain perception scenario as well. All the proposed methods were validated using hyperspectral data acquired in unstructured terrain environments in Estonia.

One of the main aims of the work was to develop a band selection method for unstructured terrain classification applications for autonomous vehicle perception. The min-max pooling band selection was introduced in Chapter three to select the most effective spectral bands from hyperspectral datacubes. This method takes significantly less time compared to information density-based methods. According to the experiment results, the proposed methods showed 8% better classification accuracy over comparable band selection methods. The proposed method is mathematically simpler to implement on low power computing hardware used in autonomous vehicles.

Another objective was to use the hyperspectral dataset to prepare image training datasets for RGB image semantic segmentation to minimize manual labelling. Manual semantic labelling of a single 512 x 512 px terrain image takes approximately 2 hours. The HSI classification was able to produce the classification result in a minute. HSI classification result with further post-processing, the image annotation was significantly reduced. Chapter 4 discusses the manifold alignment-based method for RGB image generation from HSI. These RGB images, generated from HSI datacubes, were used for DCNN training as a dataset. This HSI generated RGB image dataset resulted in approximately 17% less accurate than high resolution original RGB images. It was an expected result because the conversion process from HSI to RGB brings the loss of information and the addition of noise. However, increasing the pixel pairs from both HSI datacube and high-resolution RGB images taken from the same location could enhance the generated RGB image quality by reducing the noise and other artefacts.

The pixel-wise HSI classification has been presented in Chapter 5. HSI pixel-wise classification using spectral data is compared with RGB semantic segmentation. The results show that the HSI pixel-wise classification is more accurate in certain terrain classes than the RGB image semantic segmentation.

The unstructured terrain semantic segmentation is a much more complex task than that of structured terrains. Since one of the objectives was to demonstrate hyperspectral image segmentation accuracy for terrain classification in terms of spectral and spatial classification, several image datasets containing a different number of image bands were used to train classification DCNNs. The result from those classifications was compared against RGB image segmentation. The IoU performance matrix showed that the hyperspectral images yielded 11% better classification accuracy over RGB images. In this case, hypercubes with 25 bands yield the highest classification accuracy compared with 9 and 16 bands using the same DCNN model. The model was based on UNet architecture which is an encode-decoder model with skip connections.

Here is a summary of the results of this research.

- A simple band selection method developed and demonstrated the effectiveness in the classification of terrain HSI datacubes.
- Pixel-wise classification achieved high accuracy for the tested datasets.

- Use of classification as labels improved image annotation time. (Manual annotation 1 hrs per image, HSI classification takes ~9 min + corrections)
- RGB image generation based on the manifold alignment method resulted in a higher correlation to the RGB images taken from the same location.
- HSI spectral-spatial classification can achieve higher classification accuracy over RGB semantic segmentation for unstructured environments
- Unstructured terrain hyperspectral dataset can be used for further research.

This thesis work demonstrated the benefits of hyperspectral imaging methods to improve autonomous vehicle perception systems.

6.1 Future works

In this study, one of the outcomes was band selection for hyperspectral imagery in off-road terrain conditions. A push broom mobile hyperspectral camera was used for the experiments that used the line scan method for image acquisition. However, manufacturers can fabricate a custom hyperspectral imaging sensor based on specific wavelength filters with fewer bands. Therefore, one of the future directions could be a fabrication of a multi-spectral imaging sensor for AGV perception that can use the area scan method to acquire images. The proposed band selection method could be used to select a suitable number of spectral bands. Such a custom snapshot hyperspectral imaging sensor could provide all the necessary spectral information to classify off-road terrain scenarios. In this case, it is necessary to consider that geographic location influences the terrain classes in those regions, affecting classification.

Developing dedicated convolutional neural network models for efficient and accurate terrain spectral image classification is another research direction. Such improvements could bring spectral imaging into real-life UGV applications.

The results presented in this thesis work shows that the hyperspectral imaging-based classification performs better than the RGB imaging methods. Currently, some companies are working on developing video-rate hyperspectral cameras. With such a development, there is a possibility to use video-rate hyperspectral imaging in real-time scene segmentation and object detection for unmanned ground vehicle applications.

For short-range depth perception, stereo vision RGB is used as an additional visual input in some applications. Similarly, stereo hyperspectral imaging can enhance short-range depth perception better than stereo RGB imaging.

List of Figures

Figure 1. (a) Military UGV developed by Rheinmetall AG[3], (b) agriculture robot weeding on a farm land[4], (c) Starship parcel delivery robot running on snow[5], (d) driverless shuttle bus in Japan for rural public transport [6].	10
Figure 2. (a) A structured terrain where traversable and non-traversable areas can be easily distinguished. (Tallinn University of Technology parking lot) (b) Unstructured terrain where the traversable terrain is ambiguous (Location near Lasnamäe, Estonia).	11
Figure 3. U-Net architecture for biomedical image segmentation [60].	20
Figure 4. The RGB image of the terrain HSI dataset.	27
Figure 5. Spectral data cube unfolding.	28
Figure 6. Min-max pooling method.	29
Figure 7. Spectral characteristics of Tree and Water terrain classes.	31
Figure 8. Places of the image where training sample pixels captured.	32
Figure 9. Ground truth and classification results for each band set.	33
Figure 10. Classification results for different band selection methods.	35
Figure 11. The false colour representation of the Indian pines dataset	35
Figure 12. Ground truth of AVIRIS Indian Pines dataset from Purdue University [116].....	36
Figure 13. Spectral signatures of Indian Pines dataset classes	38
Figure 14. HSI classification as an input for the RGB semantic segmentation.	40
Figure 15 a. RGB image taken from the same site	41
Figure 16. Manifold alignment for HSI and RGB images.	42
Figure 17. Pixel-pair regions for both RGB and HSI.	45
Figure 18. RGB representation using manifold alignment method.	45
Figure 19. Generated RGB visualization images from various methods.	46
Figure 20. RGB visualization of HSI datacube in VNIR range bands	48
Figure 21. RGB image generation from the visible light range bands.	49
Figure 22. Pearson’s correlation coefficient for each colour channel between original RGB and generated RGB from HSI.	50
Figure 23. The composition of the terrain dataset.	55
Figure 24. Pixel-wise classification and RGB semantic segmentation comparison.	59

List of Tables

Table 1. Main differences between HSI, MSI, Spectroscopy and RGB imaging. Classification bullet rate (1-3) is a relative representation of the information available in each acquisition method [83].	21
Table 2. Selected spectral bands from various band selection methods.	32
Table 3. CNN classifier architecture for HSI pixel-wise classification.	32
Table 4. Classification pixel accuracy for different numbers of spectral bands.	34
Table 5. Classification pixel accuracy using different band selection methods.	34
Table 6. Indian Pines dataset classes and the number of samples for classification.	37
Table 7. Selected spectral bands from the Indian Pines dataset.	37
Table 8. The HSI classification with a different number of image bands.	37
Table 9. Classification pixel-wise accuracy for the Indian Pines dataset based on a different number of bands.	38
Table 10. RMSE for each image generation method.	47
Table 11. Pearson’s correlation coefficients for each colour channel.	47
Table 12. Pearson’s correlation coefficient of RGB images based on the number of spectral bands	48
Table 13. Pearson’s correlation coefficient for the RGB images generated by using visible light spectral bands.	49
Table 14. Unstructured terrain HSI datacubes visualized in RGB using manifold alignment.	51
Table 15. Correlation coefficients for above Table 10 image set.	51
Table 16. List of terrain classes	54
Table 17. Sample images from the RGB image dataset and classification results.	56
Table 18. Segmentation accuracy for higher resolution RGB images	57
Table 19. Segmentation accuracy matrices.	57
Table 20. Sample images from the dataset and results	58
Table 21. Sample images from the dataset and results.	60
Table 22. The classification performance of HSI generated images and annotations.	60
Table 23. Sample images from three different datasets of 9, 16 and 25 bands HSI images with classification results.	62
Table 24. Semantic segmentation accuracy for different numbers of image bands.	62
Table 25. Overall classification results comparison.	63

References

- [1] “Tesla Places Big Bet on Vision-Only Self-Driving - IEEE Spectrum.” <https://spectrum.ieee.org/tesla-places-big-bet-vision-only-self-driving> (accessed Oct. 01, 2021).
- [2] “Intel’s Mobileye demos autonomous car equipped only with cameras, no other sensors | Reuters.” <https://www.reuters.com/article/tech-ces-intel-idINKBN1Z60C0> (accessed Oct. 01, 2021).
- [3] “Rheinmetall Mission Master SP & XT A-UGVs | HICONSUMPTION.” <https://hiconsumption.com/rheinmetall-mission-master-sp-and-xt-a-ugvs/> (accessed Aug. 27, 2021).
- [4] “Autonomous weeding & agricultural robots.” <https://www.naio-technologies.com/en/home/> (accessed Aug. 27, 2021).
- [5] “A Day in the Life of a Starship Robot - Purdue University News.” <https://www.purdue.edu/newsroom/stories/2019/a-day-in-the-life-of-a-starship-robot.html> (accessed Aug. 27, 2021).
- [6] “Seniors-Shuttling Autonomous Cars: driverless shuttle bus.” <https://www.trendhunter.com/trends/driverless-shuttle-bus> (accessed Aug. 27, 2021).
- [7] A. P. Spring, “A History of Laser Scanning, Part 1: Space and Defense Applications,” *Photogramm. Eng. Remote Sens.*, vol. 86, no. 7, pp. 419–429, 2020, doi: 10.14358/pers.86.7.419.
- [8] S. D. Pendleton *et al.*, “Perception, planning, control, and coordination for autonomous vehicles,” *Machines*, vol. 5, no. 1, pp. 1–54, 2017, doi: 10.3390/machines5010006.
- [9] SAE, “SAE J3016_202104 - Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles.” SAE International, 2021.
- [10] ISO, “ISO/SAE PAS 22736 - Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles.” [Online]. Available: <https://www.iso.org/standard/73766.html>.
- [11] O. Liu, S. Yuan, and Z. Li, “A survey on sensor technologies for unmanned ground vehicles,” in *Proceedings of 2020 3rd International Conference on Unmanned Systems, ICUS 2020*, 2020, pp. 638–645, doi: 10.1109/ICUS50048.2020.9274845.
- [12] J. wen Hu *et al.*, “A survey on multi-sensor fusion based obstacle detection for intelligent ground vehicles in off-road environments,” *Front. Inf. Technol. Electron. Eng.*, vol. 21, no. 5, pp. 675–692, 2020, doi: 10.1631/FITEE.1900518.
- [13] E. Balestrieri, P. Daponte, L. De Vito, and F. Lamonaca, “Sensors and measurements for unmanned systems: An overview,” *Sensors*, vol. 21, no. 4, pp. 1–27, 2021, doi: 10.3390/s21041518.
- [14] C. Goodin, M. Doude, C. R. Hudson, and D. W. Carruth, “Enabling off-road autonomous navigation-simulation of LIDAR in dense vegetation,” *Electron.*, vol. 7, no. 9, 2018, doi: 10.3390/electronics7090154.
- [15] H. C. Kang, H. N. Han, H. C. Bae, M. G. Kim, J. Y. Son, and Y. K. Kim, “Hsv color-space-based automated object localization for robot grasping without prior knowledge,” *Appl. Sci.*, vol. 11, no. 16, 2021, doi: 10.3390/app11167593.
- [16] J. Cao, C. Song, S. Peng, F. Xiao, and S. Song, “Improved traffic sign detection and recognition algorithm for intelligent vehicles,” *Sensors (Switzerland)*, vol. 19, no. 18, 2019, doi: 10.3390/s19184021.

- [17] B. Y. Suprpto, A. Ghaida, H. Hikmarika, and S. Dwijayanti, "Road and Vehicles Detection System Using HSV Color Space for Autonomous Vehicle," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 6, no. 1, p. 42, Jul. 2020, doi: 10.26555/jiteki.v16i1.16949.
- [18] X. Dong, M. Jakobi, S. Wang, M. H. Köhler, X. Zhang, and A. W. Koch, "A review of hyperspectral imaging for nanoscale materials research," *Appl. Spectrosc. Rev.*, vol. 54, no. 4, pp. 285–305, 2019, doi: 10.1080/05704928.2018.1463235.
- [19] "Multispectral vs Hyperspectral Imagery Explained - GIS Geography." <https://gisgeography.com/multispectral-vs-hyperspectral-imagery-explained/> (accessed Aug. 18, 2021).
- [20] M. A. Wulder *et al.*, "Current status of Landsat program, science, and applications," *Remote Sens. Environ.*, vol. 225, pp. 127–147, May 2019, doi: 10.1016/j.rse.2019.02.015.
- [21] "Estonian robotics developer Milrem delivers its first UGV to Australia." <https://estonianworld.com/technology/estonian-robotics-developer-milrem-delivers-its-first-ugv-to-australia/> (accessed Aug. 18, 2021).
- [22] "Cleveron Unveils New Driverless Semi-Autonomous Last Mile Delivery Vehicle - Cleveron 701 — Cleveron." <https://cleveron.com/cleveron-mobility/cleveron-701> (accessed Aug. 26, 2021).
- [23] "Estonian delivery robots are transforming the world — Estonia." <https://estonia.ee/delivery-robots-created-by-estonian-engineers-are-transforming-the-world/> (accessed Aug. 18, 2021).
- [24] Z. Chen and X. Huang, "Pedestrian detection for autonomous vehicle using multi-spectral cameras," *IEEE Trans. Intell. Veh.*, vol. 4, no. 2, pp. 211–219, 2019, doi: 10.1109/TIV.2019.2904389.
- [25] C. Winkens, F. Sattler, and D. Paulus, "Hyperspectral terrain classification for ground vehicles," *VISIGRAPP 2017 - Proc. 12th Int. Jt. Conf. Comput. Vision, Imaging Comput. Graph. Theory Appl.*, vol. 5, no. Visigrapp, pp. 417–424, 2017, doi: 10.5220/0006275404170424.
- [26] A. L. Rankin and L. H. Matthies, "Passive sensor evaluation for unmanned ground vehicle mud detection," *J. F. Robot.*, vol. 27, no. 4, pp. 473–490, Apr. 2010, doi: 10.1002/rob.20341.
- [27] Ç. Kaymak and A. Uçar, "A Brief Survey and an Application of Semantic Image Segmentation for Autonomous Driving," in *Smart Innovation, Systems and Technologies*, vol. 136, 2019, pp. 161–200.
- [28] D. Feng *et al.*, "Deep Multi-Modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 3, pp. 1341–1360, Mar. 2021, doi: 10.1109/TITS.2020.2972974.
- [29] B. Li, Y. Shi, Z. Qi, and Z. Chen, "A survey on semantic segmentation," *IEEE Int. Conf. Data Min. Work. ICDMW*, vol. 2018-Novem, pp. 1233–1240, 2019, doi: 10.1109/ICDMW.2018.00176.
- [30] L. Sun, K. Yang, X. Hu, W. Hu, and K. Wang, "Real-Time fusion network for rgb-d semantic segmentation incorporating unexpected obstacle detection," *IEEE Robot. Autom. Lett.*, vol. 5, no. 4, pp. 5558–5565, 2020, doi: 10.1109/LRA.2020.3007457.
- [31] A. Signoroni, M. Savardi, A. Baronio, and S. Benini, "Deep learning meets hyperspectral image analysis: A multidisciplinary review," *J. Imaging*, vol. 5, no. 5, 2019, doi: 10.3390/jimaging5050052.

- [32] G. Konstantatos, "Current status and technological prospect of photodetectors based on two-dimensional materials," *Nat. Commun.*, vol. 9, no. 1, pp. 9–11, 2018, doi: 10.1038/s41467-018-07643-7.
- [33] "XIMEA - Hyperspectral cameras with USB3." <https://www.ximea.com/en/usb3-vision-camera/hyperspectral-usb3-cameras-mini> (accessed Aug. 23, 2021).
- [34] S. S. I. Ltd, "Specim IQ Tech Specs," 2019. <https://www.specim.fi/iq/tech-specs/> (accessed Nov. 14, 2019).
- [35] "Hyperspectral Imaging Systems | Hyperspectral Camera | Resonon." <https://resonon.com/> (accessed Aug. 23, 2021).
- [36] C. Fu, C. Mertz, and J. M. Dolan, "LIDAR and Monocular Camera Fusion: On-road Depth Completion for Autonomous Driving," *2019 IEEE Intell. Transp. Syst. Conf. ITSC 2019*, pp. 273–278, 2019, doi: 10.1109/ITSC.2019.8917201.
- [37] U. A. Muller, L. D. Jackel, Y. LeCun, and B. Flepp, "Real-time adaptive off-road vehicle navigation and terrain classification," *Unmanned Syst. Technol. XV*, vol. 8741, no. May 2013, p. 87410A, 2013, doi: 10.1117/12.2015533.
- [38] M. Suzuki, T. Saitoh, E. Terada, and Y. Kuroda, *Near-to-far self-supervised road estimation for complicated environments*, vol. 43, no. 18. IFAC, 2010.
- [39] D. Maturana, P.-W. Chou, M. Uenoyama, and S. Scherer, "Real-Time Semantic Mapping for Autonomous Off-Road Navigation," pp. 335–350, 2018, doi: 10.1007/978-3-319-67361-5_22.
- [40] C. J. Holder and T. P. Breckon, *Encoding Stereoscopic Depth Features for Scene Understanding in off-Road Environments*, vol. 10882 LNCS. Springer International Publishing, 2018.
- [41] S. Gupta, R. Girshick, P. Arbeláez, and J. Malik, "Learning Rich Features from RGB-D Images for Object Detection and Segmentation," in *Brain and Cognition*, vol. 60, no. 2, 2014, pp. 345–360.
- [42] A. N. Erkan, R. Hadsell, P. Sermanet, J. Ben, U. Muller, and Y. LeCun, "Adaptive long range vision in unstructured terrain," *IEEE Int. Conf. Intell. Robot. Syst.*, pp. 2421–2426, 2007, doi: 10.1109/IROS.2007.4399622.
- [43] C. J. Holder, T. P. Breckon, and X. Wei, "Depth Not Needed - An Evaluation of RGB-D Feature Encodings for Off-Road Scene Understanding by Convolutional Neural Network," pp. 6–9, 2018, [Online]. Available: <http://arxiv.org/abs/1801.01235>.
- [44] C. Caraffi, S. Cattani, and P. Grisleri, "Off-road path and obstacle detection using decision networks and stereo vision," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 4, pp. 607–618, 2007, doi: 10.1109/TITS.2007.908583.
- [45] A. Milella, G. Reina, and M. M. Foglia, "A multi-baseline stereo system for scene segmentation in natural environments," *IEEE Conf. Technol. Pract. Robot Appl. TePRA*, pp. 3–8, 2013, doi: 10.1109/TePRA.2013.6556370.
- [46] A. L. Rankin, A. Huertas, and L. H. Matthies, "Stereo-vision-based terrain mapping for off-road autonomous navigation," *Unmanned Syst. Technol. XI*, vol. 7332, no. April 2009, p. 733210, 2009, doi: 10.1117/12.819099.
- [47] A. Rankin *et al.*, "Unmanned ground vehicle perception using thermal infrared cameras," *Unmanned Syst. Technol. XIII*, vol. 8045, no. May 2011, p. 804503, 2011, doi: 10.1117/12.884349.
- [48] L. H. Matthies, P. Bellutta, and M. McHenry, "Detecting water hazards for autonomous off-road navigation," *Unmanned Gr. Veh. Technol. V*, vol. 5083, no. 1, p. 231, 2003, doi: 10.1117/12.496942.

- [49] D. F. Huber, L. J. Denes, M. Hebert, M. S. Gottlieb, B. Kaminsky, and P. Metes, "A Spectro-polarimetric Imager for Intelligent Transportation Systems," in *Intelligent Transportation Systems*, Jan. 1998, vol. 3207, pp. 94–102, doi: 10.1117/12.300844.
- [50] S. T. Namin and L. Petersson, "Classification of materials in natural scenes using multi-spectral images," *IEEE Int. Conf. Intell. Robot. Syst.*, pp. 1393–1398, 2012, doi: 10.1109/IROS.2012.6386074.
- [51] C. Winkens, V. Kobelt, and D. Paulus, "Robust Features for Snapshot Hyperspectral Terrain-Classification," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10424 LNCS, 2017, pp. 16–27.
- [52] C. Fuchs, F. Neuhaus, and D. Paulus, "High-resolution hyperspectral ground mapping for robotic vision," no. April 2018, p. 64, 2018, doi: 10.1117/12.2310066.
- [53] E. Shelhamer, J. Long, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017, doi: 10.1109/TPAMI.2016.2572683.
- [54] S. Hao, Y. Zhou, and Y. Guo, "A Brief Survey on Semantic Segmentation with Deep Learning," *Neurocomputing*, vol. 406, pp. 302–321, 2020, doi: 10.1016/j.neucom.2019.11.118.
- [55] G. Papandreou, L. C. Chen, K. P. Murphy, and A. L. Yuille, "Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2015 Inter, pp. 1742–1750, 2015, doi: 10.1109/ICCV.2015.203.
- [56] B. Baheti, S. Innani, S. Gajre, and S. Talbar, "Semantic scene segmentation in unstructured environment with modified DeepLabV3+," *Pattern Recognit. Lett.*, vol. 138, pp. 223–229, 2020, doi: 10.1016/j.patrec.2020.07.029.
- [57] B. Baheti, S. Innani, S. Gajre, and S. Talbar, "Eff-UNet: A novel architecture for semantic segmentation in unstructured environment," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2020-June, pp. 1473–1481, 2020, doi: 10.1109/CVPRW50498.2020.00187.
- [58] I. Sgibnev, A. Sorokin, B. Vishnyakov, and Y. Vizilter, "DEEP SEMANTIC SEGMENTATION FOR THE OFF-ROAD AUTONOMOUS DRIVING," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XLIII-B2-2, no. B2, pp. 617–622, Aug. 2020, doi: 10.5194/isprs-archives-XLIII-B2-2020-617-2020.
- [59] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," *36th Int. Conf. Mach. Learn. ICML 2019*, vol. 2019-June, pp. 10691–10700, 2019.
- [60] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *IEEE Access*, vol. 9, pp. 16591–16603, May 2015, [Online]. Available: <http://arxiv.org/abs/1505.04597>.
- [61] W. Zhou, J. S. Berrio, S. Worrall, and E. Nebot, "Automated Evaluation of Semantic Segmentation Robustness for Autonomous Driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 5, pp. 1951–1963, May 2020, doi: 10.1109/TITS.2019.2909066.
- [62] A. Wendel and J. Underwood, "Illumination compensation in ground based hyperspectral imaging," *ISPRS J. Photogramm. Remote Sens.*, vol. 129, pp. 162–178, 2017, doi: 10.1016/j.isprsjprs.2017.04.010.
- [63] "Supervisely - Web platform for computer vision. Annotation, training and deploy." <https://supervise.ly/> (accessed Jan. 29, 2020).

- [64] "SuperAnnotate," *SuperAnnotate AI, Inc.* <https://superannotate.com/> (accessed Apr. 29, 2021).
- [65] Mathworks, "Image Labeler." <https://www.mathworks.com/help/vision/ref/imagelabeler-app.html>.
- [66] "Labelbox: The leading training data solution." <https://labelbox.com/> (accessed Jan. 29, 2020).
- [67] K. Wada, "labelme: Image Polygonal Annotation with Python." 2016.
- [68] "Visual Object Tagging Tool (VoTT)," *Microsoft.* <https://github.com/Microsoft/VoTT> (accessed Apr. 29, 2021).
- [69] X. Qin, S. He, Z. Zhang, M. Dehghan, and M. Jagersand, "ByLabel: A boundary based semi-automatic image annotation tool," *Proc. - 2018 IEEE Winter Conf. Appl. Comput. Vision, WACV 2018*, vol. 2018-Janua, pp. 1804–1813, 2018, doi: 10.1109/WACV.2018.00200.
- [70] M. Suchi, T. Patten, D. Fischinger, and M. Vincze, "EasyLabel: A semi-automatic pixel-wise object annotation tool for creating robotic RGB-D datasets," *Proc. - IEEE Int. Conf. Robot. Autom.*, vol. 2019-May, pp. 6678–6684, 2019, doi: 10.1109/ICRA.2019.8793917.
- [71] M. A. Reza, A. U. Naik, K. Chen, and D. J. Crandall, "Automatic Annotation for Semantic Segmentation in Indoor Scenes," *IEEE Int. Conf. Intell. Robot. Syst.*, pp. 4970–4976, 2019, doi: 10.1109/IROS40897.2019.8968230.
- [72] P. Mishra, M. S. M. Asaari, A. Herrero-Langreo, S. Lohumi, B. Diezma, and P. Scheunders, "Close range hyperspectral imaging of plants: A review," *Biosyst. Eng.*, vol. 164, pp. 49–67, 2017, doi: 10.1016/j.biosystemseng.2017.09.009.
- [73] D. Caballero, R. Calvini, and J. M. Amigo, "Hyperspectral imaging in crop fields: precision agriculture," *Data Handl. Sci. Technol.*, vol. 32, pp. 453–473, 2020, doi: 10.1016/B978-0-444-63977-6.00018-3.
- [74] G. ElMasry, M. Kamruzzaman, D. W. Sun, and P. Allen, "Principles and Applications of Hyperspectral Imaging in Quality Evaluation of Agro-Food Products: A Review," *Crit. Rev. Food Sci. Nutr.*, vol. 52, no. 11, pp. 999–1023, 2012, doi: 10.1080/10408398.2010.543495.
- [75] D. P. Ariana, R. Lu, and D. E. Guyer, "Near-infrared hyperspectral reflectance imaging for detection of bruises on pickling cucumbers," *Comput. Electron. Agric.*, vol. 53, no. 1, pp. 60–70, 2006, doi: 10.1016/j.compag.2006.04.001.
- [76] R. Vejarano, R. Siche, and W. Tesfaye, "Evaluation of biological contaminants in foods by hyperspectral imaging: A review," *Int. J. Food Prop.*, vol. 20, no. 2, pp. 1264–1297, Dec. 2017, doi: 10.1080/10942912.2017.1338729.
- [77] L. M. Kandpal, J. Tewari, N. Gopinathan, P. Boulas, and B. K. Cho, "In-process control assay of pharmaceutical microtablets using hyperspectral imaging coupled with multivariate analysis," *Anal. Chem.*, vol. 88, no. 22, pp. 11055–11061, 2016, doi: 10.1021/acs.analchem.6b02969.
- [78] F. Franch-Lage, J. M. Amigo, E. Skibsted, S. Maspoch, and J. Coello, "Fast assessment of the surface distribution of API and excipients in tablets using NIR-hyperspectral imaging," *Int. J. Pharm.*, vol. 411, no. 1–2, pp. 27–35, 2011, doi: 10.1016/j.ijpharm.2011.03.012.
- [79] G. Edelman, T. G. van Leeuwen, and M. C. G. Aalders, "Hyperspectral imaging for the age estimation of blood stains at the crime scene," *Forensic Sci. Int.*, vol. 223, no. 1–3, pp. 72–77, 2012, doi: 10.1016/j.forsciint.2012.08.003.

- [80] G. J. Edelman, E. Gaston, T. G. van Leeuwen, P. J. Cullen, and M. C. G. Aalders, "Hyperspectral imaging for non-contact analysis of forensic traces," *Forensic Sci. Int.*, vol. 223, no. 1–3, pp. 28–39, 2012, doi: 10.1016/j.forsciint.2012.09.012.
- [81] M. Moroni, A. Mei, A. Leonardi, E. Lupo, and F. La Marca, "PET and PVC separation with hyperspectral imagery," *Sensors (Switzerland)*, vol. 15, no. 1, pp. 2205–2227, 2015, doi: 10.3390/s150102205.
- [82] A. Pölder, M. Jurma, and M. Tamre, "Waste Paper Sorting Using Imaging Spectroscopy," pp. 283–284.
- [83] T. Adão *et al.*, "Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry," *Remote Sens.*, vol. 9, no. 11, 2017, doi: 10.3390/rs9111110.
- [84] "Multispectral imaging technology," *Spectral Devices Inc.*, 2019. <https://www.spectraldevices.com/content/multispectral-imaging-technology%0Ahttp://files/262/multispectral-imaging-technology.html> (accessed Oct. 06, 2021).
- [85] "Custom and Off-the-shelf Multispectral Imaging | SILIOS Technologies." <https://www.silios.com/multispectral-imaging> (accessed Oct. 06, 2021).
- [86] Parrot, "Datasheet Specs sheet Sequoia : the multi-band sensor designed for agriculture," 2018, [Online]. Available: www.parrot.com.
- [87] Q. Li, X. He, Y. Wang, H. Liu, D. Xu, and F. Guo, "Review of spectral imaging technology in biomedical engineering: achievements and challenges," *J. Biomed. Opt.*, vol. 18, no. 10, p. 100901, 2013, doi: 10.1117/1.jbo.18.10.100901.
- [88] F. Luo, B. Du, L. Zhang, L. Zhang, and D. Tao, "Feature learning using spatial-spectral hypergraph discriminant analysis for hyperspectral image," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2406–2419, 2019, doi: 10.1109/TCYB.2018.2810806.
- [89] W. Ma, C. Gong, Y. Hu, P. Meng, and F. Xu, "The Hughes phenomenon in hyperspectral classification based on the ground spectrum of grasslands in the region around Qinghai Lake," *Int. Symp. Photoelectron. Detect. Imaging 2013 Imaging Spectrom. Technol. Appl.*, vol. 8910, no. August 2013, p. 89101G, 2013, doi: 10.1117/12.2034457.
- [90] R. Cai, Y. Yuan, and X. L. B, *Hyperspectral Band Selection with Convolutional Neural Network*, vol. 1. Springer International Publishing, 2018.
- [91] Y. Zhan, D. Hu, H. Xing, and X. Yu, "Hyperspectral Band Selection Based on Deep Convolutional Neural Network and Distance Density," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2365–2369, 2017, doi: 10.1109/LGRS.2017.2765339.
- [92] Q. Wang, F. Zhang, and X. Li, "Optimal Clustering Framework for Hyperspectral Band Selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5910–5922, 2018, doi: 10.1109/TGRS.2018.2828161.
- [93] W. Sun and Q. Du, "Hyperspectral band selection: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 118–139, 2019, doi: 10.1109/MGRS.2019.2911100.
- [94] X. Cao, B. Wu, D. Tao, and L. Jiao, "Automatic Band Selection Using Spatial-Structure Information and Classifier-Based Clustering," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 9, no. 9, pp. 4352–4360, 2016, doi: 10.1109/JSTARS.2015.2509461.
- [95] M. F. Baumgardner, L. L. Biehl, and D. A. Landgrebe, "220 Band AVIRIS Hyperspectral Image Data Set: June 12, 1992 Indian Pine Test Site 3." Sep. 2015, doi: doi:/10.4231/R7RX991C.
- [96] "AVIRIS." <https://aviris.jpl.nasa.gov/>.

- [97] "Hyperspectral Remote Sensing Scenes." http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes.
- [98] A. Datta, S. Ghosh, and A. Ghosh, "Unsupervised band extraction for hyperspectral images using clustering and kernel principal component analysis," *Int. J. Remote Sens.*, vol. 38, no. 3, pp. 850–873, 2017, doi: 10.1080/01431161.2016.1271470.
- [99] H. Huang, G. Shi, H. He, Y. Duan, and F. Luo, "Dimensionality Reduction of Hyperspectral Imagery Based on Spatial-Spectral Manifold Learning," *IEEE Trans. Cybern.*, vol. PP, pp. 1–13, 2019, doi: 10.1109/tcyb.2019.2905793.
- [100] J. Feng *et al.*, "Convolutional Neural Network Based on Bandwise-Independent Convolution and Hard Thresholding for Hyperspectral Band Selection," *IEEE Trans. Cybern.*, pp. 1–15, 2020, doi: 10.1109/tcyb.2020.3000725.
- [101] P. R. Lorenzo, L. Tulczyjew, M. Marcinkiewicz, and J. Nalepa, "Band Selection from Hyperspectral Images Using Attention-based Convolutional Neural Networks," no. Section III, pp. 1–7, 2018, [Online]. Available: <http://arxiv.org/abs/1811.02667>.
- [102] N. P. Jacobson, S. Member, and M. R. Gupta, "Design Goals and Solutions for Display of Hyperspectral Images," vol. 43, no. 11, pp. 2684–2692, 2005.
- [103] D. Liao, Y. Qian, J. Zhou, and Y. Y. Tang, "A Manifold Alignment Approach for Hyperspectral Image Visualization with Natural Color," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3151–3162, 2016, doi: 10.1109/TGRS.2015.2512659.
- [104] H. Su, Q. Du, and P. Du, "Hyperspectral image visualization using band selection," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 6, pp. 2647–2658, 2014, doi: 10.1109/JSTARS.2013.2272654.
- [105] X. Kang, P. Duan, and S. Li, "Hyperspectral image visualization with edge-preserving filtering and principal component analysis," *Inf. Fusion*, vol. 57, no. July 2019, pp. 130–143, 2020, doi: 10.1016/j.inffus.2019.12.003.
- [106] K. Kotwal and S. Chaudhuri, "Visualization of hyperspectral images using bilateral filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2308–2316, 2010, doi: 10.1109/TGRS.2009.2037950.
- [107] K. V. Kale, M. M. Solankar, and D. B. Nalawade, "Hyperspectral Endmember Extraction Techniques," *Process. Anal. Hyperspectral Data*, pp. 1–20, 2020, doi: 10.5772/intechopen.88910.
- [108] K. Książek, M. Romaszewski, P. Głomb, B. Grabowski, and M. Cholewa, "Blood stain classification with hyperspectral imaging and deep neural networks," *Sensors (Switzerland)*, vol. 20, no. 22, pp. 1–24, 2020, doi: 10.3390/s20226666.
- [109] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016, doi: 10.1109/TGRS.2016.2584107.
- [110] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep convolutional neural networks for hyperspectral image classification," *J. Sensors*, vol. 2015, 2015, doi: 10.1155/2015/258619.
- [111] Z. Meng, L. Li, X. Tang, Z. Feng, L. Jiao, and M. Liang, "Multipath residual network for spectral-spatial hyperspectral image classification," *Remote Sens.*, vol. 11, no. 16, pp. 1–19, 2019, doi: 10.3390/rs11161896.

- [112] C. B. Chittineni, "Utilization of spectral-Spatial information in the classification of imagery data," *Comput. Graph. Image Process.*, vol. 16, no. 4, pp. 305–340, 1981, doi: 10.1016/0146-664X(81)90044-7.
- [113] Y. Guo, H. Cao, J. Bai, and Y. Bai, "High Efficient Deep Feature Extraction and Classification of Spectral-Spatial Hyperspectral Image Using Cross Domain Convolutional Neural Networks," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, no. 1, pp. 345–356, 2019, doi: 10.1109/JSTARS.2018.2888808.
- [114] Y. Li, H. Zhang, and Q. Shen, "Spectral–Spatial Classification of Hyperspectral Imagery with 3D Convolutional Neural Network," *Remote Sens.*, vol. 9, no. 1, p. 67, Jan. 2017, doi: 10.3390/rs9010067.
- [115] B. Liu *et al.*, "Spectral – spatial classification of hyperspectral image using three-dimensional convolution network," vol. 12, no. 1, 2020, doi: 10.1117/1.JRS.12.
- [116] Acerca de Grupo de Inteligencia Computacional (GIC), "Hyperspectral Remote Sensing Scenes," 2021. http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes.
- [117] Xiaofei He and Partha Niyogi, "Locality preserving projections," *Neural Inf. Process. Syst.*, vol. 16, p. 153, 2010, [Online]. Available: <https://papers.nips.cc/paper/2359-locality-preserving-projections.pdf>.
- [118] Y. Zhu, P. K. Varshney, and H. Chen, "Evaluation of ICA based fusion of hyperspectral images for color display," *FUSION 2007 - 2007 10th Int. Conf. Inf. Fusion*, 2007, doi: 10.1109/ICIF.2007.4408068.
- [119] G. Lin, A. Milan, C. Shen, and I. Reid, "RefineNet: Multi-path refinement networks for high-resolution semantic segmentation," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-January, pp. 5168–5177, 2017, doi: 10.1109/CVPR.2017.549.
- [120] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11211 LNCS, pp. 833–851, 2018, doi: 10.1007/978-3-030-01234-2_49.

Acknowledgements

I would like to thank the Department of Electrical Power Engineering and Mechatronics, Tallinn University of Technology, Estonia, for giving me an opportunity to further research towards a PhD in mechatronics. Moreover, I would like to thank my main supervisor, professor Mart Tamre and co-supervisor Dr Robert Hudjakov, for their guidance and support throughout the studies. I especially thank to following colleagues at the department who assisted and motivated me, Even Sekhri, Ali Zahavi and Svetlana Gromova, Merle Kutsar. I am fortunate to have my family and friends constantly encouraging me to follow my dreams and ambitions.

Furthermore, I would like to express my gratitude to the colleagues at the department for their support and well wishes to culminate my academic dream.

Without financial support from the DoRa program, the article publications could not be possible, and I appreciate the financial assistance I received during the PhD program.

Dhanushka Chamara Liyanage

Abstract

Smart Terrain Perception Using Hyperspectral Imaging

Hyperspectral imaging gives a huge advantage over RGB images in terms of information abundance. This technology has yielded higher success in various application domains of machine vision. However, its uses in autonomous vehicle perception are rather unexplored. This thesis investigates various possibilities of HSI offering for perception improvement in autonomous vehicles. The emphasis was on unstructured terrain conditions where there has been limited research conducted.

Even though hyperspectral images contain a large amount of data, it has both advantages and drawbacks too. The computing power needed to process those images are enormous as one image could contain hundreds of image bands. Various feature selection and feature extraction methods have been developed to reduce the computing burden while maximising the classification outcome. However, in most cases, they have been developed for certain applications. In this thesis, the prime focus is on unstructured or off-road terrain segmentation, a simplified band selection method is proposed. The accuracy of the band selection method has been compared with other comparable bands selection methods, where it showed approximately 3% better classification accuracy with experimental results.

Semantic segmentation models need images with labels to train them. When multiple imaging technologies are involved with different resolutions, they need to prepare multiple image datasets. Such a demand for too many datasets increases the effort needed to prepare the datasets. In the case of HSI and RGB, if it is possible to share the same image labels with RGB and HSI, it could help to reduce the labelling effort needed. Since the RGB images contain lesser data compared to HSI datacubes, the RGB images processing is faster. The cost-effectiveness of RGB imaging for scene understanding makes it an essential part of the perception system. Therefore, HSI could only enhance the perception system with its capabilities rather than replacing RGB imaging. The RGB images generated from HSI datacubes were used for the semantic segmentation experiments to evaluate the possibilities to share some part of the dataset with HSI.

As previously mentioned, the RGB image generation method from HSI was introduced. Three different approaches were investigated to find the optimal correlation between the RGB images generated from HSI and original RGB images captured from RGB imaging cameras. The methods are bilateral filtering, selecting three image bands from each red, green and blue region of visible light wavelength range and manifold alignment method. The manifold alignment proved to be the optimal method for the RGB image generation from HSI datacubes. The highest correlation to the original RGB images was achieved with nine band HSI datacubes. With deep convolutional neural networks, semantic segmentation accuracy of the image dataset created from the RGB images generated from HSI was compared to the original RGB images captured using RGB cameras. The semantic segmentation accuracy of RGB images generated from HSI datacubes is slightly lower compared to original RGB images.

Spectral images are typically classified using spectral data in pixel-wise. However, combining spectral-spatial features could achieve better classification accuracy compared to pixel-wise HSI classification and three-channel RGB classification. Therefore, the HSI dataset has been used for semantic segmentation with the spectral-spatial combination. The results showed that the segmentation is much higher with spectral-spatial combination, which stands at 11% higher than RGB semantic

segmentation with state-of-the-art classification networks. The segmentation accuracy was evaluated with three different datasets prepared based on the number of spectral bands. They were 9, 16 and 25 bands. The 25-band dataset yielded the highest classification accuracy. Even though the nine-band HSI classification accuracy was lower than the 25-band dataset, still nine bands classification showed higher accuracy than RGB semantic segmentation.

All in all, the hyperspectral imaging method could enhance the perception system accuracy for autonomous vehicles running on unstructured terrains or off-road conditions. In terms of real-life implementation of a spectral imaging method for autonomous driving vehicles, it is possible to develop a multispectral imaging sensor with fewer spectral bands, which is most efficient for scene classification. Together with such an optimized spectral imaging camera, the images can capture in the area scan mode. Developing DCNN models tailored for spectral imaging-based scene understanding, spectral imaging for perception can become a reality.

Lühikokkuvõte

Hüperspektraal-pilditehnika maastiku nutikaks tajumiseks

Hüperspektraalne pildistamine annab RGB-kujutistele tohutu eelise teabe rohkuse osas. See tehnoloogia on andnud suuremat edu erinevates masinanägemise rakendusvaldkondades. Siiski on selle kasutamine autonoomse tajumise puhul üsna uurimata. Selles väitekirjas uuritakse erinevaid võimalusi HSI pakub taju parandamiseks autonoomsete sõidukite. Rõhuasetus oli struktureerimata maastikutingimustel, kus uuringud on olnud piiratud.

Kuigi hüperspektraalsed pildid sisaldavad suurt hulka andmeid, on sellel nii eelised kui ka puudused. Nende piltide töötlemiseks vajalik andmetöötlusvõimsus on tohutu, sest üks pilt võib sisaldada sadu pildiribasid. Andmetöötluskoormuse vähendamiseks on välja töötatud erinevad funktsioonide valimise ja eraldamise meetodid, maksimeerides samal ajal klassifitseerimise tulemust. Enamikul juhtudel on need siiski välja töötatud teatavate rakenduste jaoks. Selles väitekirjas keskendutakse peamiselt struktureerimata või maastikulõigule, tehakse ettepanek kasutada lihtsustatud ribavaliku meetodit. Ribavaliku meetodi täpsust on võrreldud teiste võrreldavate ribade valikumeetoditega, kus see näitas ligikaudu 3 % paremat klassifitseerimistäpsust katsetulemustega.

Semantiline segmenteerimismudelid vajavad pilte siltidega, et neid trennida. Kui mitu pilditöötlustehnoloogiat on seotud erinevate resolutsioonidega, peavad nad ette valmistama mitu pildiandmestikku. Selline nõudlus liiga paljude andmekogumite järele suurendab andmekogumite ettevalmistamiseks vajalikke jõupingutusi. Kui HSI ja RGB puhul on võimalik jagada samu kujutismärke RGB ja HSI-ga, võib see aidata vähendada vajalikku märgistamiskoormust. Kuna RGB pildid sisaldavad vähem andmeid kui HSI-andmekuubikud, siis RGB piltide töötlemine on kiirem. RGB pildistamise kulutõhusus stseeni mõistmiseks muudab selle tajusüsteemi oluliseks osaks. Seetõttu võis HSI vaid parandada tajusüsteemi oma võimetega, selle asemel et asendada RGB pilditöötlust. HSI andmekuubikutest saadud RGB-pilte kasutati semantilistes segmenteerimiskatsetes, et hinnata võimalusi jagada osa andmekogumist HSI-ga.

Nagu eespool mainitud, võeti kasutusele RGB pildi genereerimise meetod HSI-st. Uuriti kolme erinevat lähenemisi, et leida optimaalne korrelatsioon HSI-st saadud RGB piltide ja RGB-kaamerate pildistatud algsete RGB piltide vahel. Meetodid on kahepoolne filtreerimine, valides igast nähtava valguse lainepikkuse vahemikus olevast punasest, rohelisest ja sinisest piirkonnast kolm kujutisriba ja kollektori joondamise meetod. Kollektori joondamine osutus optimaalseks meetodiks RGB pildi genereerimiseks HSI-andmekuubikutest. Suurim korrelatsioon algsete RGB piltidega saavutati üheksa sagedusala HSI andmekuubiga. Sügavate konvolutsiooniliste närvivõrkude puhul võrreldi HSI loodud RGB piltide semantilist segmenteerimistäpsust RGB kaamerate abil salvestatud algsete RGB piltidega. HSI andmekuubikutest genereeritud RGB-kujutiste semantiline segmenteerimistäpsus on võrreldes algsete RGB piltidega veidi väiksem.

Spektraalsed kujutised klassifitseeritakse tavaliselt spektraalandmete abil piksli järgi. Spektraalsete ja ruumiliste omaduste kombineerimine võib aga saavutada parema klassifitseerimistäpsuse võrreldes pikslite HSI klassifikatsiooniga ja kolmekanalilise RGB klassifikatsiooniga. Seetõttu on HSI andmekogumit kasutatud semantiliseks segmenteerimiseks spektraal-ruumilise kombinatsiooniga. Tulemused näitasid, et segmenteerimine on palju suurem spektraal-ruumilise kombinatsiooni puhul, mis on 11 % suurem kui RGB semantiline segmenteerimine tiptasemel klassifikatsioonivõrkudega.

Segmenteerimistäpsust hinnati kolme erineva andmekogumiga, mis koostati spektriribade arvu põhjal. Nad olid 9, 16 ja 25 bändi. 25-ribaline andmestik andis kõrgeima klassifitseerimistäpsuse. Kuigi üheksaribaline HSI klassifikatsiooni täpsus oli väiksem kui 25-ribaline andmestik, näitas üheksa sagedusala klassifikatsioon siiski suuremat täpsust kui RGB semantiline segmenteerimine.

Kokkuvõttes võib hüperspektraalne pildistamise meetod suurendada autonoomsete sõidukite tajusüsteemi täpsust, mis töötab struktureerimata maastikul või maastikul. Autonoomsete sõidukite spektraalse pildistamise meetodi tegeliku rakendamise seisukohast on võimalik välja töötada vähem spektriribadega multispektraalne pildiandur, mis on stseeni liigitamiseks kõige tõhusam. Koos sellise optimeeritud spektraalkaameraga võivad pildid jäädvustada ala skaneerimisrežiimis. DCNN-mudelite väljatöötamine, mis on kohandatud spektraalse pildistamise alusel stseeni mõistmiseks, võib muutuda reaalsuseks.

Appendix

Publication I

D. C. Liyanage, R. Hudjakov, M. Tamre, "Hyperspectral / Multispectral imaging methods for quality control" as a chapter of the book titled "Handbook of Research on New Investigations in Artificial Life, AI, and Machine Learning" by IGI global publisher, pp. 1–33, 2021, doi.org/10.4018/978-1-7998-8686-0.



www.igi-global.com

Hyperspectral / Multispectral Imaging Methods for Quality Control

Dhanushka Chamara Liyanage, Robert Hudjakov, Mart Tamre
Tallinn University of Technology, Estonia

ABSTRACT

Product quality assurance is a vital component in any manufacturing process. With the advancement of machine vision, product quality inspection has been vastly improved. It could not be achieved with human inspection otherwise regarding consistency, accuracy, and speed. Emerging modern sensor technologies and image processing algorithms ensure product and process quality in various modern industries, including pharmaceutical manufacturing, food production, agriculture, and waste sorting. Hyperspectral and multispectral imaging are among those novel technologies which find their uses in the quality inspection domain. Due to a spectral image comprising numerous spectral bands, a spectral image contains more information than an RGB image. With the help of spectral and spatial information in spectral images, it is possible to discriminate the quality indices of various products with higher accuracy than RGB imaging methods. This chapter looks into recent developments in product quality evaluation using spectral imaging methods. The literature review covers food production, the textile industry, pharmaceuticals manufacturing applications.

Keywords: Inspection, Manufacturing, Machine Vision, Product Quality, Spectral Imaging, Spectroscopy, Food Quality, Agriculture, Textiles, Pharmaceuticals

INTRODUCTION

According to Crosby, the quality of a product is defined as its conformity to the specification (Crosby, 1979). There are various processes employed in production environments to assure product compliance to its specifications. Visual product inspection is one of such methods used in production environments to assure the absence of defects in products. Mostly, the product quality inspection is done by human operators visually on factory floors. However, human inspection is unreliable as there are various factors affecting product quality determination. Different human operators may have different judgements in their product quality inspection due to work experience, conditions at the work environment, level of understanding of the product, psychological factors, fatigue, biological factors of the worker. Products with more than one type of defect take a longer time for visual inspection. When the acceptance criterion of defects changes, it complicates visual inspection. The human inspection has more downfalls with qualitative measurements as results for the qualitative measurements are varying and difficult to compare. Moreover, human perception can easily accept false positives (Kerkeni et al., 2016). As a result, human inspection is inconsistent, subjective, and slow.

Furthermore, the visual inspection is incapable of determining physicochemical characteristics of the product, such as moisture content, presence of various microorganisms, texture, etc. The physicochemical characteristics estimations are vital to ensure the quality of different products such as pharmaceuticals, food, and beverages. There are various analytical methods used in production floors to estimate those physicochemical parameters of products.

The process control methods can be distinguished into four categories: in-line, on-line, at-line, and off-line methods. In-line methods are directly immersed into the process flow, while on-line methods use a bypass channel from the main process flow. The at-line methods analyse the samples next to the process flow by sample extraction and off-line methods process analysis separately from the process flow by withdrawing some samples (Boldrini et al., 2012). Even though some manufacturing processes can employ off-line quality evaluation methods by taking random samples, they are not suitable for quality critical production processes. The food and beverages manufacturing, pharmaceuticals manufacturing industries require a complete quality inspection prior to human

consumption, requiring in-line quality checks. Again, in some cases, random sample testing is not suitable as it may destroy the product. In such cases, non-destructive in-line process control is a vital requirement in ensuring product quality.

Therefore, it is proved that the industry needs consistent, accurate, fast, cost-effective and in-line automated inspection methods to increase the effectiveness of visual inspection. With the recent developments in imaging and computer technology, machine vision solves most quality inspection and control needs in production environments.

Machine vision for quality inspection and control

A machine vision system combines image acquisition cameras, illumination sources, processing computers and analysing algorithms to produce accurate decisions to assure various visible characteristics of the products (Du & Sun, 2006). These systems are used to detect product defects, make visual measurements and check various product markings as part of the quality control process in various industries. Detection of scratches on the visible area of aerosol metal-can production, dimension check of mechanical parts in the automobile industry, identifying the electronic components' location and orientation, identifying the types of the components in electronics assembly lines are a few examples of machine vision-based quality inspection using RGB (Red, Green, Blue) and monochrome imaging cameras.

Even though machine vision based on RGB sensors solves a significant part of visual quality inspection problems, it has some drawbacks too. The RGB and monochrome vision systems can only detect the objects' colour and shape properties. It does not help industrial problems where product quality cannot be assured using colour, shape, detecting specific patterns or checking dimensions, i.e. meat freshness, quantitative determination of various chemical presence in fruits and vegetables. Moreover, the products like fruits and vegetables of the same kind and variety do not have the same shape and colour in all the items, making it more complex to analyse using RGB imaging methods. In order to cover those needs, there should be a method that can acquire more information than RGB sensors can do. Spectral imaging is an advanced imaging technology that captures spectral and spatial information of the product under inspection.

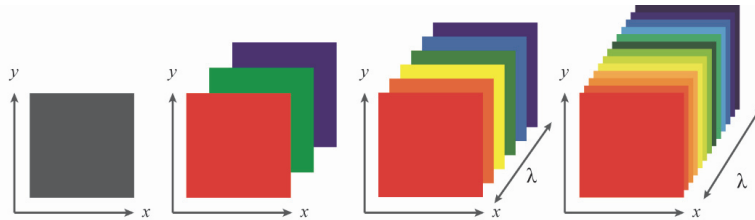
Spectral imaging

Spectroscopy gives light absorption or reflectance characteristics of a point on the object under inspection (Wallace et al., 2009). The imaging spectroscopy or spectral imaging method combines spectroscopy and imaging, first introduced in the 1970s for remote sensing applications with the Landsat program (Wulder et al., 2019). The technology has evolved into other application domains such as agriculture, food, pharmaceuticals, etc. Over recent decades, steady development in spectral imaging technologies has been happening in image acquisition hardware, software, and analysis methods.

The computer vision-based inspection systems have advanced with the introduction of imaging sensors capable of seeing beyond the visible light range, which is 400 – 700 nm wavelength range in the electromagnetic spectrum. Now they can take advantage of ultraviolet (UV) and infrared (IR) regions of the electromagnetic spectrum for machine vision applications.

Unlike monochrome or RGB colour imaging methods, spectral imaging or imaging spectroscopy is a technology that acquires tens or hundreds of images in various contiguous wavelengths in the electromagnetic spectrum. In the monochrome imaging method, the image contains a single grayscale

colour channel. In contrast, RGB colour images contain Red, Green and Blue as three separate channels within the visible light spectrum. Spectral imaging is classified into two categories according to the number of images bands contained in a spectral image cube. Multispectral Imaging (MSI) and Hyperspectral Imaging (HSI) are the two main branches. The difference between those two branches is that the multispectral image consists of tens of spectral bands, where the hyperspectral image contains hundreds of spectral bands. Below, Figure 1 illustrates the differences between each imaging technology.



Imaging Method	Monochrome	RGB	Multispectral	Hyperspectral
Color Channels	1	3	3 - ~10	~10 - ~100
Information Depth	Spatial	Spatial	Spatial + spectral	Spatial + spectral

Figure 1. Different imaging methods and their characteristics

Due to their unique design in sensor construction, spectral imaging technology uses various image acquisition techniques to capture image data. These methods are whisk-broom, push-broom and tunable filter techniques, where each method is used in point-scanning, line-scanning and area-scanning, respectively (Iqbal et al., 2014). The point scan method, which is the whisk broom technique, can acquire spectral data of one pixel at a time. Hence, the whisk broom method scans the object in lateral (X) and longitude (Y) directions. With line scanning, the imaging sensor can acquire a wider strip of pixels. Therefore, it is needed to sweep the sensor across the object under inspection to record a complete image. Most hyperspectral imaging sensors are line scan sensors, as it is technically challenging to embed hundreds of light filters on top of neighbouring pixels to create an area scanning sensor. Area scan captures the image in either one wavelength or a few wavelengths simultaneously. Therefore, the area-scan method is standard with multispectral sensors. In multispectral imaging, either a few light bandpass filters are deposited on top of the sensor or external filters use to capture the images in several wavelength bands. More about the construction of spectral imaging sensors and other relevant hardware is discussed in “Hyperspectral and multispectral imaging for evaluating safety and quality” journal article (Qin et al., 2013).

Together with different image acquisition methods, spectral imaging uses different electromagnetic wavelength ranges for machine vision applications. They are mainly Visible Near Infrared (VNIR), Short-Wave Infrared (SWIR), Medium-Wave InfraRed (MWIR) and Long-Wave Infrared (LWIR). The VNIR is in the 400 – 1000 nm wavelength range, SWIR spans 1000 – 2500 nm, MWIR spans 3 – 5 μm and LWIR spans 8 – 12 μm (Edmund Optics Inc, n.d.). Figure 2 illustrates the electromagnetic wavelength spectrum with all the above wavelength range classifications.

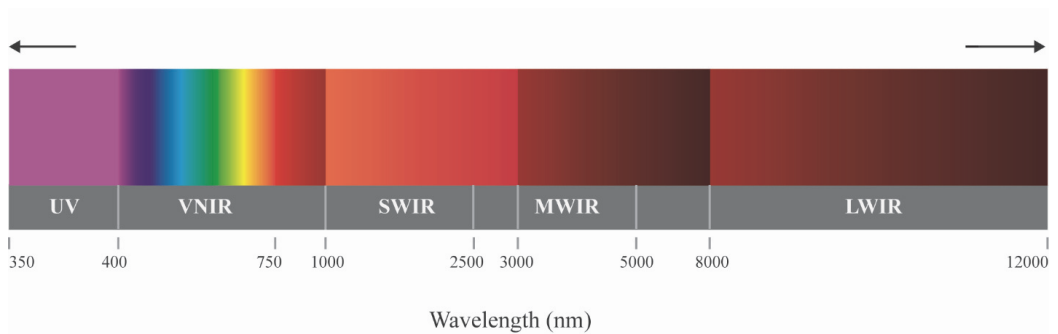


Figure 2. Electromagnetic spectrum

A Hyperspectral or multispectral image is a series of images captured at different wavelength bands in a continuous spectrum that forms an image data cube (Amigo, 2020). These data cubes are also called hypercubes. Spectral signatures of the subjects are the essential characteristics to extract from the hypercube. Every material, substance, or compound has unique spectral reflectance, transmittance characteristics which is the spectral signature of the same. Below Figure 3 shows an example of spectral reflectance characteristics or spectral signature of a plant leaf. With the help of hyperspectral signature, it has been possible to detect various infections, diseases to crops (Yuan et al., 2019), (ElMasry et al., 2012), and defects in vegetables and fruits (ElMasry et al., 2007), (ElMasry et al., 2008), (Xing et al., 2005). There has also been a considerable number of researches on the use of HSI for meat production industry for meat quality analysis (Qiao et al., 2007), (ElMasry, Barbin, et al., 2012), (Kamruzzaman et al., 2012). Waste sorting and recycling is also one of the research areas where HSI is used (Serranti et al., 2015).

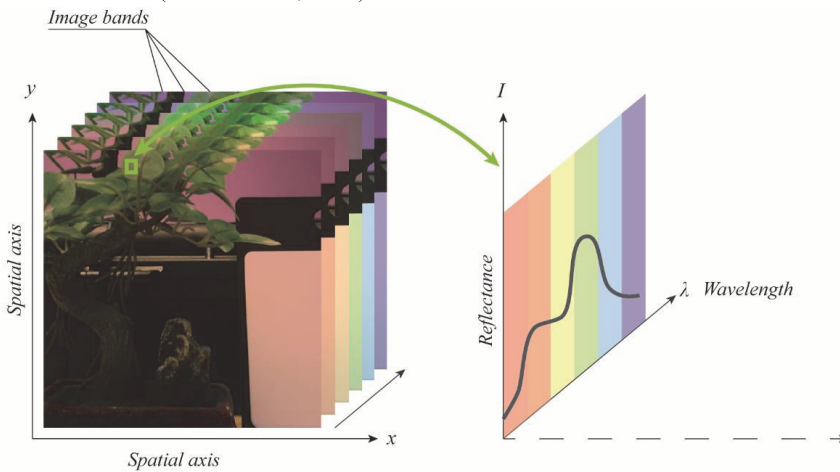


Figure 3. The reflectance spectrum of one pixel is the spectral signature of the object in the pixel. X and Y are spatial dimensions, while λ is the wavelength

Hyperspectral image acquisition hardware set-up

Typical hyperspectral imaging camera imaging acquisition follows push-broom technique as those cameras are mostly line scan cameras. Thus, it uses a linear or rotary motion mechanism for scanning the object under inspection. These systems use special lighting to illuminate the object depending on the wavelength range of the camera. The most suitable lighting option in the visible light (VNIR) range is incandescent lights (Zahavi et al., 2019). In HSI, the image acquisition control, storing and analysing are done by a computer in the same set-up. A typical set-up for hyperspectral imaging is shown in Figure 4. Unlike hyperspectral cameras, the simultaneous spatial and spectral data acquisition ability eliminates the need for a moving table. Therefore, multispectral cameras can grab images faster than hyperspectral counterparts.

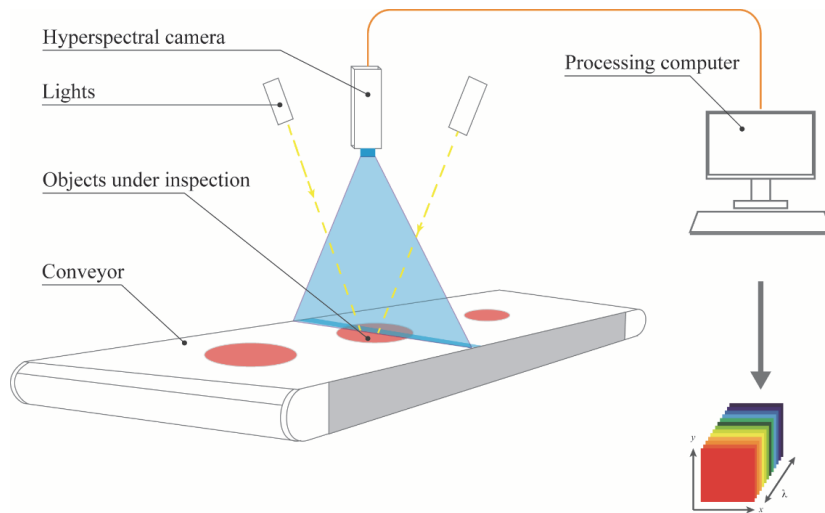


Figure 4. A typical hyperspectral imaging system

This chapter surveys product quality assessment methods based on HSI/MSI technologies in various industrial sectors. The article covers the research literature published in recent years related to quality assurance. Moreover, various image processing, analysis methods are presented as well.

INDUSTRIAL APPLICATIONS OF HYPERSPECTRAL / MULTISPECTRAL IMAGING METHODS FOR QUALITY EVALUATION

Hyperspectral and multispectral imaging methods have a wide range of applications in the food, agriculture, pharmaceuticals, material inspection and waste recycling industries. However, product quality inspection in the food and agriculture industry is the most dominant application area compared to other industrial sectors. In this section, the applications and analytical methods are discussed according to those industrial sectors.

Textile Industry

The textile industry finds various applications for hyperspectral imaging as it is possible to discriminate various textile materials and chemical coatings based on their spectral characteristics. Evaluations of several textile quality indices have been published by Mirschel et al., related to the application of various chemicals on textiles. One of their studies proposed a method to monitor the thickness and homogeneity of hot melt adhesive layers in the laminates made of black polyester textiles. The authors have used partial least squares regression (PLSR) for thickness estimation of the adhesive layer in the near-infrared (NIR) spectral range. The importance of this study is that the laminate adhesive layer covered by the top black textile layer and NIR wavelength range penetrates to a sufficient depth until the adhesive layer. This research revealed that it could estimate the thickness and homogeneity of the adhesive layer with sufficient precision to carry out process control. The proposed method showed the Root Mean Square Error of Prediction (RMSEP) 6 gm^{-2} (Mirschel et al., 2019). The same authors proposed a method to quantitatively estimate the application weight and homogeneity of finishing chemicals in textiles. In this study, the researchers have used the NIR wavelength range (1320 – 1900 nm). The chemical agents used in this research were colourless, which cannot be detected with machine vision methods in the visible light wavelength range. PLS model has been employed to determine the application weight, which yielded results with predicted precision of the chemical for a flame retardant 2 gm^{-1} and stiffening agent with RMSEP $1,6 \text{ gm}^{-1}$ (Mirschel et al., 2018) (Mirschel et al., 2017).

Another crucial task in the textile industry is Fibre fabric identification, which demands a fast, non-destructive method to identify the fibres. The classification of eight different such woven fabric fibres was achieved in the SWIR wavelength range. The k-nearest neighbour (kNN) method yielded 80% or more accuracy in classification for cotton, polyester, polyethylene, wool, PVC, nylon and linen fabrics where Locally Preserving Projection (LPP) method used for band reduction (Li et al., 2019). Jin and others also proposed a method to discriminate synthetic fibres in textile production using NIR spectral range. This study covers six types of synthetic fibres such as polyethylene, para-aramid, polypropylene, polyester, polyamide, acrylic fibres. The hyperspectral images were pre-processed before performing classification by averaging spectra pixels and applying the Savitzky-Golay filtering method. Pre-processed images of fabric samples were analysed using Principal Component Analysis – Linear Discriminant Analysis (PCA-LDA) model with discrimination accuracy of 100% (Jin et al., 2017). In contrast to the first research on fibre identification which uses a mix of natural and synthetic fibres, the second research was entirely based on synthetic fibres. Another remark is that the PCA-LDA method can achieve better classification accuracy over kNN for synthetic fibres.

Detection of foreign objects in products or contamination is an essential part of the manufacturing sector. In cotton production, there are possibilities to contaminate the cotton lint from various sources. During the ginning process of cotton, the cotton seeds and other foreign objects filtered out. However, there can be some foreign matter that remains with the lint affecting the cotton quality. Detection of

foreign matter in cotton using fluorescence hyperspectral imaging has been explored by Mustafic et al. They were using UV lights of 365 nm at the peak intensity, and the chosen wavelength range was 425 – 700 nm in visible light range. The authors have used 19 spectral bands for the classification, where the dimensionality reduction has been done with PCA. A 90% average classification rate has been achieved using linear discriminant analysis, and they claim that the method is suitable to detect foreign matter with a higher amount of fluorophores (Mustafic et al., 2016).

In waste sorting applications, the objective is to classify various materials into separate groups for recycling. For textile waste sorting, E. Herrala et al. researched the possibilities of using NIR hyperspectral imaging methods for various natural and synthetic textile materials sorting. The research showed successful classification of cotton, polyester, linen, wool materials. However, the mixed fabric materials with tiny amounts of natural or synthetic material content were difficult to classify (Herrala & Oy, 2020). Similarly, Makela and others proposed a method to determine the polyester content of man-made and natural cellulose and polyester blends. They have used Principal Component Analysis (PCA) for the identification of characteristic wavelengths, and the estimation of polyester content was achieved by using the PLSR method with a prediction error of 4.5% within a range of 0 -100 % polyester (Mäkelä et al., 2020).

Below Table 1 summarises hyperspectral imaging-based quality indices evaluation methods in the textile industry.

Table 1. Textile quality parameters evaluation using hyperspectral imaging methods

Application	Spectral range	Methods	Reference
Polyester content estimation in natural and man-made cellulose and polyester blends	NIR	PCA, PLSR	(Mäkelä et al., 2020)
Fibre discrimination in woven fabrics		SWIR LPP, kNN	(Li et al., 2019)
Synthetic fibres discrimination	SWIR (900 – 2500 nm)	PCA-LDA	(Jin et al., 2017)
Adhesive layer thickness and homogeneity estimation of textile laminates	NIR	PLSR	(Mirschel et al., 2019)
Quantitative analysis of application weights of chemical additives in textile production. (Flame retardants and stiffening agents)	NIR (1320 – 1900 nm)	PLS, PCA	(Mirschel et al., 2018), (Mirschel et al., 2017)
Detection of foreign matter in cotton	VIS (425 – 700 nm)	LDA	(Mustafic et al., 2016)

Impact of HSI in Food Production Quality Assurance

Most of the HSI and MSI based quality inspection efforts have been made into food and agriculture-related areas. In food quality assurance, it is essential to use non-destructive methods that need a thorough quality inspection for the entire production. However, the non-destructive tests are not that important for batch quality checking scenarios as it is only required to test a small sample randomly taken from a batch. In agriculture, in-situ monitoring of the plants is essential to minimise the risk of

poor harvest caused by the poor nutrient in plants. Timely estimation of plant nutrients can significantly enhance the yield. Therefore, it is obvious the necessity of optical methods for quality assurance over any other methods.

Agriculture Industry

As most of the applications for spectral imaging can be found in the agricultural sector, a detailed description of the vital quality indices evaluation for each crop can provide a greater understanding of the uses of the technology.

Strawberry serves as one of the main ingredients for various food products, including cakes, jams, flavoured juices, milk products, etc. The internal quality attributes of strawberries which are moisture content (MC), total soluble solids (TSS), acidity estimation of strawberries, have been done using HSI methods. Visible Near Infrared (VNIR) wavelength range has been used for capturing the hyperspectral data cubes. The data cubes were pre-processed using mean centring and automatic baseline correction before using the Partial Least Squares (PLS) method for the analysis. The optimal wavelength bands have been selected using the highest absolute values of regression coefficients - β from the PLS model (ElMasry et al., 2007). Visual detection of bruises on McIntosh apples is a difficult task for colour imaging methods. Mainly due to different background colours where imaging takes place and different ages of bruises. However, using the HSI method, the early detection (< 12 h) of apple bruises has been achieved using spectral imaging in the VNIR wavelength range. Three spectral bands in NIR region 750, 820, and 960 nm have been identified as the most influential spectral bands based on stepwise discrimination analysis along with VIP scores from the PLS method. With a limited number of spectral bands, it is possible to implement a multispectral imaging system to in-line monitoring of apple bruises on the field. The multilevel adaptive thresholding method proved that the apple bruises could be detected, from as new as 1 hour to as old as 3 days (ElMasry et al., 2008). Like those strawberry internal quality parameters evaluation, mango fruit internal quality parameters prediction has been experimented with using HSI. The selected quality parameters were firmness, total soluble solids, and titrable acidity. Even though the PLSR prediction model has been developed for VNIR spectral range imaging, excessive prediction errors in the PLSR model restricted its use only for early screening. As a result, the authors claim that titrable acidity and total soluble solids estimation need further improvements (Rungpichayapichet et al., 2017). An on-line quality assessment method for pomegranate fruits using the MSI system has been proposed by Khodabakhshian et al. to predict pH, total soluble solids, and titrable acidity. In this application, four significant wavelength bands were selected as 700, 800, 900, 1000 nm, and multiple linear regression models have been used to estimate the internal quality parameters (Khodabakhshian et al., 2017). Bruise detection or mechanical damage detection method for pickling cucumber has been proposed using HSI. In this research, the hyperspectral images were acquired within NIR / SWIR spectral range. The research suggests that the best wavelength range for mechanical damage detection lies within 950 - 1350 nm. Band ratio and band differences in selected wavelength bands were used as classification methods where 988, 1085 nm was used for the band ratio, and 1346 and 1425 nm were used for the band difference method. The results show that both methods achieved more than 80% classification accuracy (Ariana et al., 2006).

Crop growth status is an essential factor in industrial agriculture to reduce the risk of poor harvest. In rice cultivation, the nitrogen content estimation of the rice plant during the panicle initiation stage helps to calculate optimum nitrogen fertiliser requirement. Using VNIR hyperspectral imaging, Onoyama and others estimated the nitrogen content of rice plants in their panicle initiation stage. The

reflectance intensities have been extracted using the difference of GreenNDVI-NDVI with three selected spectral bands in near-infrared, green and red bands at 845 nm, 564 nm and 668 nm, respectively. The nitrogen estimation models have been developed using the partial least squares regression method with 0.95 gm-2 of RMSE (Onoyama et al., 2013).

Food quality can deteriorate during its storage under different climate conditions. Freeze damages during storage of white button mushrooms before they are visually evident could be identified using HSI. Their research suggests that ice formation causes structural damage in mushrooms, which results in reflectance characteristics change. The research has been conducted within the 400 – 1000 nm wavelength range and was analysed using PCA and LDA methods. According to the authors, Gowen et al., the classification results yielded 97.9 % accuracy of freeze damage detection (Aoife A. Gowen et al., 2009).

HSI has been used to detect mycotoxins and mycotoxigenic fungi in cereal grain sorting such as wheat, maize and barley. One such application is Deoxynivalenol (DON) detection in cereals. Since prolonged DON exposure can cause health hazards to humans and farm animals, it is vital to identify such mycotoxins. The most significant wavelength bands for this task were found in 750 – 1650 nm, the near-infrared wavelength range. Various studies have been conducted about DON detection using different analytical / classification methods such as PLSR, PLS-DA, SVM to determine the concentration. A comprehensive review of fungi assessment on cereals has been published by Femenias et al. (Femenias et al., 2020).

Dried vegetable quality inspection is possible with multispectral imaging, according to numerous studies. Primarily it has a large number of use cases in moisture content prediction. The real-time quality inspection is possible due to the limited number of wavelength bands used from multispectral imaging in the range of 675-975 nm with 25 bands. Prediction of moisture content and shrinkage ratio has been achieved using PLSR and LS-SVM in dried carrot slices with a coefficient of determination for the shrinkage ratio as 0.942 while 0.953 for the moisture content (Yu et al., 2020).

Tomato paste is a highly consumed vegetable product globally. The viscosity of tomato paste can vary according to Sucrose content, affecting the taste and negatively impacting the industry. An MSI based method for sucrose adulteration detection in tomato paste has been proposed by Liu et al. The research suggests PLSR, LS-SVM, BPNN models for quantitative and qualitative assessments within the wavelength range of 405 – 970 nm. Their study reveals LS-SVM model has achieved a 1% accuracy level in the quantitative estimation of adulterated sucrose content (Liu et al., 2017).

HSI and MSI have numerous applications in the agricultural sector for product quality assurance as it provides a non-destructive solution to the industry. Agriculture crops related quality parameters evaluation and the used methods are summarised in Table 2.

Table 2. Summary of agriculture crops/products quality parameters prediction using HSI

Product	Quality Indices	Wavelength Range	Analysis Methods	Reference
Strawberry	MC, TSS, Acidity	VNIR (400 – 1000 nm)	PLS	(ElMasry et al., 2007)
Rice	Nitrogen	VNIR (400 – 1000 nm)	PLSR	(Onoyama et al., 2013)
Apple	Bruise detection	VNIR (400 – 1000 nm)	Multilevel adaptive thresholding of selected bands	(ElMasry et al., 2008)

White button Mushrooms	Freeze damage detection	VNIR (400 – 1000 nm)	PCA, LDA	(Aoife A. Gowen et al., 2009)
Cereals (Wheat, Maise, Barley)	Mycotoxins detection	NIR		(Femenias et al., 2020)
Pickling cucumber	Bruise detection	NIR (900 – 1700 nm)	PCA, Band ratio, Band difference	(Ariana et al., 2006)
Dried carrot		VNIR (675-975 nm) Multispectral – 25 bands	PLSR and LS-SVM	(Yu et al., 2020)
Mango	Firmness, Total Soluble Solids (TSS), Titrable Acidity (TA)	VNIR (450-998 nm)	PLSR	(Rungpichay apichet et al., 2017)
Pomegranate fruits	Total Soluble Solids (TSS), Titrable Acidity (TA), pH	VNIR (400-1100 nm)	MLR	(Khodabakhshian et al., 2017)

Seafood Production

The quality evaluation in fish and other seafood is primarily about the consistency of the meat, visual aspects, and odour. An instrumental method for fish quality estimation is freshness based on chemical spoilage assessment caused by microbiological bacteria presence (Menesatti et al., 2010).

The moisture content is one of the quality parameters which determines prawns' taste, shelf life and price. Prawns are dehydrated to extend the shelf life. However, it is essential to maintain a certain amount of moisture within prawns as too high moisture can cause microorganism growth while too low moisture can destroy nutrition content. Dehydrated prawns' moisture content and distribution were determined in the spectral range of 380 – 1100 nm by Di Wu et al. The research suggests the Successive Projections Algorithm (SPA) as hyperspectral band selection algorithm, which selected 12 spectral bands for moisture determinations in the VNIR range (428 – 999 nm). The proposed MLR algorithm yielded a coefficient of determination of 0.962, proving the suitability of the method for moisture estimation (Wu et al., 2012).

The freshness and safety of fish are determined by quality parameters such as Thiobarbituric Acid Reactive Substances (TBARS), Total Volatile Basic Nitrogen (TVB_N) and total viable counts (TVC). These parameters prediction for rainbow trout fillets have been developed using multispectral imaging in the spectral range of 430- 1010 nm. The most significant six wavelength bands have been identified to predict the above quality indices. The prediction models have been developed using the PLSR method, and the best prediction results were achieved for TVC while TBARS prediction was poor (Khoshnoudi-Nia & Moosavi-Nasab, 2019a). The same authors have extended the research to optimise the method to conduct a quality evaluation using multispectral imaging for the additional parameters such as Psychrotrophic Plate Count (PPC) and sensory score. The authors have identified nine optimal spectral bands using a genetic algorithm for their proposed method (Khoshnoudi-Nia & Moosavi-Nasab, 2019b).

Fish product freshness assessment was conducted to find the possibilities to assess the freshness as storage of days for fresh and frozen-thawed cod fillets using HSI. The authors claim that their

proposed VNIR hyperspectral imaging method can determine the storage duration of cod fillets on ice. They achieved a prediction accuracy of 1.6 days as the product storage duration on ice. Also, this method can be used as an on-line quality inspection technique as it can process one fillet per second (Sivertsen et al., 2011).

Meat Production

The quality of fresh meat largely depends on its water holding capacity. For beef water holding capacity, estimation has been investigated by using NIR hyperspectral imaging. In this research, six wavelengths were identified as crucial wavelengths for the PLS model, and they were 940, 997, 1144, 1214, 1342, and 1443 nm. The authors, ElMasry et al., have used PCA and PLS methods to quantify the water holding capacity, producing prediction accuracy with 0.87 as the coefficient of determination (ElMasry et al., 2011). Tenderness prediction of cooked beef has been proposed using VNIR (400 – 1000 nm) hyperspectral imaging method by Govindarajan et al. The proposed model predicts tenderness of cooked beef in three tenderness categories with 96.4 % accuracy (Naganathan et al., 2008). The fresh beef quality indices prediction ElMasry, G et al. have proposed a PLSR model to predict colour, pH and tenderness in a separate study. This method has been done in the NIR wavelength range of 900 – 1700 nm (Elmasry, Sun, et al., 2012).

Pork quality evaluation in terms of colour, texture (firmness), and exudation (drip loss) characteristics have been carried out by Qiao, J et al. in the spectral range of 430 – 1000 nm. In the proposed method, they have used PCA for dimensionality reduction. Artificial Neural Networks (ANN) based classifier has been able to classify the samples by 85 % with 10 PCs (Qiao et al., 2007). Water holding capacity estimation and tenderness sensing method of pork has been proposed using three wavebands using NIR multispectral imaging. The three wavebands used to estimate those parameters were 1280, 1440 and 1660 nm, where a backpropagation neural network has been used as an analytical model (Huang et al., 2015). Likewise, lamb meat quality evaluation for four different sheep breeds has been investigated by Kamruzzaman et al. using a NIR spectral imaging-based method. They have developed a prediction model for quality indices evaluations. Their approach aimed to estimate the lamb meat's pH, colour and drip loss (Kamruzzaman et al., 2012).

The bacteria such as *Listeria monocytogenes* growing in ready-to-eat meat and fish can be a severe health threat, especially for pregnant women, newborns, and adults with weak immune systems (Centre for Disease Prevention & Food Safety Authority, 2019). Some studies have used VNIR HSI methods for the rapid identification of bacteria. Several bacteria strains of *Cronobacter*, *Salmonella*, *Escherichia Coli*, *Staphylococcus*, and *Listeria monocytogenes*, have been analysed using PCA and K-Nearest Neighbor (k-NN) classification method. The *Listeria* was detected with 100% accuracy (Michael et al., 2019). A detailed review of microorganisms on food until 2017 was published by Wang et al. (K. Wang et al., 2018). Agglutination detection in fried minced meat during initial frying is important for ready-to-eat fast food production, and this has been investigated in minced beef and diced turkey. The proposed classification method is canonical discriminant analysis or Fisher's discriminant analysis. Also, the most informative spectral bands have been identified as 470, 700, 850 and 970nm for the classification (Daugaard et al., 2010).

The Aerobic Plate Count (APC) is considered a microbiological indicator for porks' sanitary quality and food safety. The APC count of cooked pork sausages has been investigated in VNIR spectral range using a multispectral camera with 19 bands. The primary wavelengths which are significant for APC detection have been identified within 570 – 850 nm. A PLSR model has been proposed to detect APC with a coefficient of determination of 0.89 (Ma et al., 2014). Multispectral imaging system for

industrial-scale poultry quality assessment has been suggested for the microbial level, which can be used to obtain Total Viable Counts (TVC) and *Pseudomonas* spp. Chicken breast fillets, thigh fillets, marinated souvlaki and burgers were used in the experiments (Spyrelli et al., 2020). Below Table 3 summarises the quality indices estimation and prediction methods proposed by various researchers on meat quality.

Table 3. Summary of meat quality parameters prediction using HSI

Product	Quality Indices	Wavelength Range	Analysis Methods	Reference
Fresh beef	pH, Tenderness, Color	NIR (900–1700 nm)	PLSR	(Elmasry, Sun, et al., 2012)
Cooked beef	Tenderness	VNIR (400–1000 nm)		(Naganathan et al., 2008)
Pork	Colour, Texture, Exudation	VNIR (430–1000 nm)		(Qiao et al., 2007)
Pork	Water holding capacity, Tenderness	1280 nm, 1440 nm and 1660 nm MSI	BPNN	(Huang et al., 2015)
Lamb	pH, Tenderness, Color	NIR (900 – 1700 nm)	PLSR	(Kamruzzaman et al., 2012)
Ready-to-eat meat and fish	Listeria bacteria	VNIR (400 – 1000 nm)	k-NN	(Michael et al., 2019),
Ready-to-eat meat	Agglutination on minced meat frying	VNIR (430-970 nm) multispectral – 18 bands	FDA	(Daugaard et al., 2010)
Cooked pork sausages	Aerobic Plate Count	VNIR (400-970 nm) Multispectral – 19 bands		(Ma et al., 2014)
Poultry products	Total Viable Counts, <i>Pseudomonas</i> spp (microbial level assessment)	VNIR 18 bands (405 – 970 nm)	PLSR	(Spyrelli et al., 2020)

Beverage Production Industry

Hyperspectral imaging has been in use for foliar disease detection in tea plants which is known as anthracnose. The researchers have identified three spectral bands which are instrumental in detecting afore mentioned foliar disease as 542, 686 and 754 nm in the visible light wavelength range. Their classification method was based on unsupervised learning and adaptive two-dimensional thresholding. The results of the proposed method proved that they could detect the disease with 98 % accuracy (Yuan et al., 2019). Mishra et al. suggest that near-infrared hyperspectral imaging could be used for green teas classification according to the country of origin. The study has been conducted wavelength

range is 950 – 1760 nm. Even though there are some misclassifications, the authors suggest that the HSI and machine learning could achieve successful results for tea quality assessment (Mishra & Nordon, 2020). The amount of Free Amino Acids (FAA) indicates the freshness, taste and aroma of yellow tea. Yang et al. have used spectral image pre-processing and band selection to identify the most significant wavelength bands for FAAcontent detection in yellow tea. The pre-processing method was Savitzky-Golay filtering, while band selection was done using SPA. The authors have proposed a Genetic Algorithm-Support Vector Regression (GA_SVR) algorithm for predicting FAA with the selected five characteristic wavelength bands in the NIR region of the electromagnetic spectrum(Yang et al., 2019).

In dairy production, melamine is used to boost protein content which causes health problems. Therefore, a research group has investigated the detection of melamine concentration using HSI in milk powders using different approaches. One of their research methods used NIR spectral range HSI of 990 – 1700 nm. Their PLSR model was successfully used to evaluate melamine-milk samples of concentration 0.02 % - 1% (Lim et al., 2016). For the other approach, they have employed spectral similarity analysis (Fu et al., 2014).

Pharmaceutical Industry

The pharmaceutical production process requires an in-line monitoring method for Active Pharmaceutical Ingredients (API) in micro tablets for quality assessment. The API could monitor within VNIR and SWIR wavelengths regions of the electromagnetic spectrum. The prediction model developed using the PLSR achieved a coefficient of determination of more than 0.90, suggesting that the proposed HSI-based rapid in-line quality inspection method is suitable for pharmaceutical production (Kandpal et al., 2016). In earlier research, Franch-Lage et al. have proposed a Multivariate Curve Resolution (MCR) technique for Lorazepam surface homogeneity assessment (Franch-Lage et al., 2011). Franch-Lage et al. suggested that the VNIR spectral range is sufficient for the pharmaceutical surface homogeneity inspection. Moreover, some researchers suggest that UV multispectral imaging could be used for pharmaceutical tablet coating defect identification (Klukkert et al., 2016).

Other Application Areas

Apart from the above applications, there are a few other areas where hyperspectral imaging is used for quality assurance. Confectionaries, food ingredients, various chemical coatings, and coated materials production are some of those applications.

The quality assessment of butter cookies based on surface browning and water content has been proposed by Andresen et al. using multispectral imaging. In this study, the authors found that surface browning can be detected within the visible light wavelength range while water content can be determined in the near-infrared spectrum. The authors have identified two wavelengths to determine the browning score of cookies as 395 and 525 nm. Moreover, the authors suggest that the NIR region is suitable for water content determination and evaluated using a PLSR model (Andresen et al., 2013). The HSI could use of as a replacement for analytical methods to classify flavoured and unflavored olive oil. According to Romaniello et al.'s findings, the 400 – 570 nm and 695 nm spectral bands gives characteristic signatures for oils (Romaniello & Baiano, 2018).

Assessment of heavy metal contamination of water sources is usually done in laboratories. However, Rostom et al. suggest an HSI based method for heavy metal concentration estimation in water bodies

such as Mn, Co, Cu, Cd, Pb, Ni, Cr, Fe, and Zn using a portable spectroradiometer. They have conducted their experiments in Mariut Lake in Egypt. Moreover, the authors have used Chlorophyll absorption wavelengths which are 450 nm and 675 nm, for the detection of algae. For heavy metal detection, they have proposed spectral bands ranging from 350 nm to 1200 nm. The authors have used Linear Regression Analysis for developing the heavy metal estimation model (Rostom et al., 2017). Wang et al. have used an airborne panchromatic-multispectral imaging system to detect pollutants on large water bodies. The work was focused on detecting aquatic vegetation, algae and other pollutants such as garbage, sewage, etc. (Z. Wang et al., 2019).

Photographic papers are coated with a light-sensitive chemical layer. A nine-band LED multispectral imaging method presented by Lalonde et al. suggests density measurements of colourants in photographic paper manufacturing is feasible (Lalonde et al., 2015). However, the authors claim their current results are inferior for certain photographic paper products, which could be improved by fine-tuning the inspection apparatus.

A coefficient independent scattering model could be used for coating thickness determination applications using VIS-NIR HSI. Dingemans et al.'s novel approach could measure coating of 250 μm with 11 μm accuracy according to the published results. In this study, the coating material was a semitransparent film-forming low-gloss wood lacquer for outdoor usage (Dingemans et al., 2017). Thickness estimation of thin aluminium oxide (Al_2O_3) layers on stainless steel foil was an example presented by Gruber et al. for HSI thin film imaging. The research has been conducted in VIS-NIR spectral range, where principal component regression (PCR) with a PLSR method has been used for layer thickness prediction. The alumina coating process is a part of single-wall carbon nanotubes (SW-CNT) production for battery applications (Gruber et al., 2016).

Polymer packaging materials hinder the quality inspection of food and pharmaceuticals. Even though they are transparent, the quality inspection can only be possible before packaging or after removing the packaging. The feasibility of HSI for packaged goods inspection and the influences of PVC and PET packaging for hyperspectral imaging have been published by Gowen et al. Their research confirms that combined with image and spectral processing of HSI data, it is possible to inspect packaged products (A. A. Gowen et al., 2010).

DISCUSSION

The literature survey was conducted by referring to the widely popular and high impact scientific research journals. The survey covered significant research publications during the past 15 years. However, this may not contain all the research articles published during the period. It is a prohibitively tedious task to cover all the articles published, and there are restrictions to access all the journals published worldwide. Figure 5 shows the distribution of the articles in this survey from each industrial sector from 2006 to 2020. The majority of the HSI based quality assessment has happened in food production during the above mentioned period.

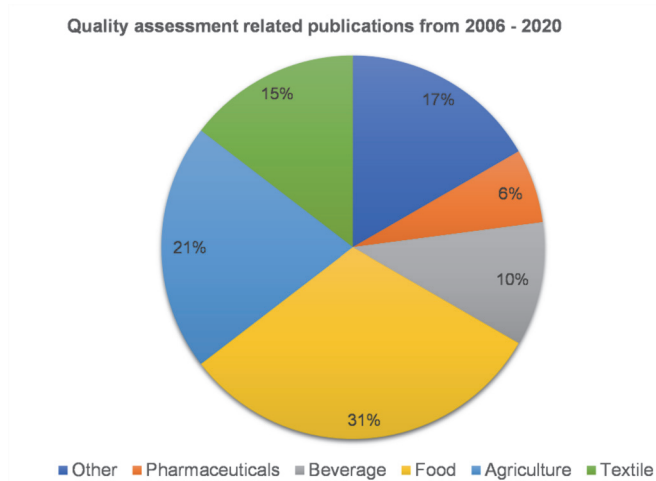


Figure 5. Selected article from 2006 to 2020 according to the industry

The number of publications is growing in hyperspectral imaging methods for quality assessment for the surveyed period. Even though early research literature related to the agriculture sector, the other industrial sectors find spectral imaging methods provide feasible solutions for quality evaluations. According to Figure 6, the food industry has a broader spread of research with spectral imaging. All the quality evaluation methods were conducted in VNIR, NIR or SWIR wavelength regions.

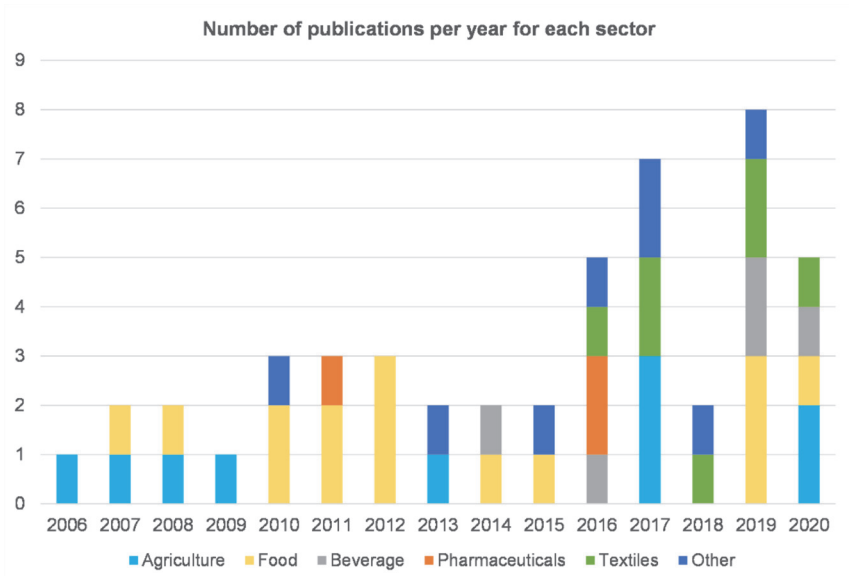


Figure 6. Number of publications per each year throughout the survey

CONCLUSION

The objective of the chapter was to present an in-depth overview of the current developments in hyperspectral and multispectral imaging for product quality assessments. Spectral imaging is still an emerging technology, where novel processing methods, hardware and applications rapidly developing.

The article presented numerous examples of hyperspectral imaging use cases in agriculture, pharmaceutical production, the textile industry, etc. Some studies suggest that the product quality inspection could be feasible using hyperspectral imaging, even the goods are packaged using polymer materials. At the same time, research evidence shows that hyperspectral remote sensing technology is a suitable method to determine the water quality on large water bodies. Such examples prove that spectral imaging could perform machine inspection, which could not do otherwise.

The survey revealed that most spectral imaging inspection methods are conducted within visible and near-infrared wavelength regions.

As a non-destructive, fast, in-situ quality inspection method, spectral imaging could benefit many industries.

REFERENCES

- Amigo, J. M. (2020). Hyperspectral and multispectral imaging: setting the scene. *Data Handling in Science and Technology*, 32, 3–16. <https://doi.org/10.1016/B978-0-444-63977-6.00001-8>
- Andresen, M. S., Dissing, B. S., & Løje, H. (2013). Quality assessment of butter cookies applying multispectral imaging. *Food Science & Nutrition*, 1(4), 315–323. <https://doi.org/10.1002/fsn3.46>
- Araújo, M. C. U., Saldanha, T. C. B., Galvão, R. K. H., Yoneyama, T., Chame, H. C., & Visani, V. (2001). The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometrics and Intelligent Laboratory Systems*, 57(2), 65–73. [https://doi.org/10.1016/S0169-7439\(01\)00119-8](https://doi.org/10.1016/S0169-7439(01)00119-8)
- Ariana, D. P., Lu, R., & Guyer, D. E. (2006). Near-infrared hyperspectral reflectance imaging for detection of bruises on pickling cucumbers. *Computers and Electronics in Agriculture*, 53(1), 60–70. <https://doi.org/10.1016/j.compag.2006.04.001>
- Boldrini, B., Kessler, W., Rebner, K., & Kessler, R. (2012). Hyperspectral imaging: a review of best practice, performance and pitfalls for in-line and on-line applications. *Journal of Near Infrared Spectroscopy*, 20(5), 438. <https://doi.org/10.1255/jnirs.1003>
- Centre for Disease Prevention, E., & Food Safety Authority, E. (2019). Multi-country outbreak of *Listeria monocytogenes* sequence type 6 infections linked to ready-to-eat meat products – November 25 2019. *EFSA Supporting Publications*, 16(12). <https://doi.org/10.2903/sp.efsa.2019.en-1745>
- Cheng, J. H., & Sun, D. W. (2014). Hyperspectral imaging as an effective tool for quality analysis and control of fish and other seafoods: Current research and potential applications. *Trends in Food Science and Technology*, 37(2), 78–91. <https://doi.org/10.1016/j.tifs.2014.03.006>
- Crosby, P. (1979). *Quality Is Free: The Art of Making Quality Certain*. New York: McGraw-Hill.
- Daugaard, S. B., Adler-Nissen, J., & Carstensen, J. M. (2010). New vision technology for multidimensional quality monitoring of continuous frying of meat. *Food Control*, 21(5), 626–632. <https://doi.org/10.1016/j.foodcont.2009.09.007>
- Dingemans, L. M., Papadakis, V. M., Liu, P., Adam, A. J. L., & Groves, R. M. (2017). Quantitative coating thickness determination using a coefficient-independent hyperspectral scattering model. *Journal of the European Optical Society*, 13(1). <https://doi.org/10.1186/s41476-017-0068-2>
- Du, C. J., & Sun, D. W. (2006). Learning techniques used in computer vision for food quality evaluation: A review. *Journal of Food Engineering*, 72(1), 39–55. <https://doi.org/10.1016/j.jfoodeng.2004.11.017>
- Edmund Optics Inc. (n.d.). What is SWIR? | Edmund Optics. Retrieved November 30, 2019, from <https://www.edmundoptics.com/knowledge-center/application-notes/imaging/what-is-swir/>
- Elmasry, G., Barbin, D. F., Sun, D. W., & Allen, P. (2012). Meat Quality Evaluation by Hyperspectral Imaging Technique: An Overview. *Critical Reviews in Food Science and Nutrition*, 52(8), 689–711. <https://doi.org/10.1080/10408398.2010.507908>
- ElMasry, G., Kamruzzaman, M., Sun, D. W., & Allen, P. (2012). Principles and Applications of Hyperspectral Imaging in Quality Evaluation of Agro-Food Products: A Review. *Critical Reviews in Food Science and Nutrition*, 52(11), 999–1023. <https://doi.org/10.1080/10408398.2010.543495>
- Elmasry, G., Sun, D. W., & Allen, P. (2012). Near-infrared hyperspectral imaging for predicting colour, pH and tenderness of fresh beef. *Journal of Food Engineering*, 110(1), 127–140. <https://doi.org/10.1016/j.jfoodeng.2011.11.028>

- EIMasry, G., Sun, D. W., & Allen, P. (2011). Non-destructive determination of water-holding capacity in fresh beef by using NIR hyperspectral imaging. *Food Research International*, *44*(9), 2624–2633. <https://doi.org/10.1016/j.foodres.2011.05.001>
- EIMasry, G., Wang, N., ElSayed, A., & Ngadi, M. (2007). Hyperspectral imaging for non-destructive determination of some quality attributes for strawberry. *Journal of Food Engineering*, *81*(1), 98–107. <https://doi.org/10.1016/j.jfoodeng.2006.10.016>
- EIMasry, G., Wang, N., Vigneault, C., Qiao, J., & ElSayed, A. (2008). Early detection of apple bruises on different background colors using hyperspectral imaging. *LWT - Food Science and Technology*, *41*(2), 337–345. <https://doi.org/10.1016/j.lwt.2007.02.022>
- Femenias, A., Gatius, F., Ramos, A. J., Sanchis, V., & Marín, S. (2020). Use of hyperspectral imaging as a tool for Fusarium and deoxynivalenol risk management in cereals: A review. *Food Control*, *108*, 106819. <https://doi.org/10.1016/j.foodcont.2019.106819>
- Franch-Lage, F., Amigo, J. M., Skibsted, E., MasPOCH, S., & Coello, J. (2011). Fast assessment of the surface distribution of API and excipients in tablets using NIR-hyperspectral imaging. *International Journal of Pharmaceutics*, *411*(1–2), 27–35. <https://doi.org/10.1016/j.ijpharm.2011.03.012>
- Gowen, A. A., O'Donnell, C. P., Esquerre, C., & Downey, G. (2010). Influence of polymer packaging films on hyperspectral imaging data in the visible-near-infrared (450–950 nm) wavelength range. *Applied Spectroscopy*, *64*(3), 304–312. <https://doi.org/10.1366/000370210790918337>
- Gowen, Aoife A., Taghizadeh, M., & O'Donnell, C. P. (2009). Identification of mushrooms subjected to freeze damage using hyperspectral imaging. *Journal of Food Engineering*, *93*(1), 7–12. <https://doi.org/10.1016/j.jfoodeng.2008.12.021>
- Gruber, F., Wollmann, P., Schumm, B., Grählert, W., & Kaskel, S. (2016). Quality control of slot-die coated aluminum oxide layers for battery applications using hyperspectral imaging. *Journal of Imaging*, *2*(2). <https://doi.org/10.3390/jimaging2020012>
- Herrala, E., & Oy, S. I. (2020). *Textile identification and sorting using hyperspectral imaging*. (April).
- Huang, Q., Li, H., Zhao, J., Huang, G., & Chen, Q. (2015). Non-destructively sensing pork quality using near infrared multispectral imaging technique. *RSC Advances*, *5*(116), 95903–95910. <https://doi.org/10.1039/c5ra18872e>
- Iqbal, A., Sun, D. W., & Allen, P. (2014). An overview on principle, techniques and application of hyperspectral imaging with special reference to ham quality evaluation and control. *Food Control*. <https://doi.org/10.1016/j.foodcont.2014.05.024>
- Jin, X., Memon, H., Tian, W., Yin, Q., Zhan, X., & Zhu, C. (2017). Spectral characterisation and discrimination of synthetic fibers with near-infrared hyperspectral imaging system. *Applied Optics*, *56*(12), 3570. <https://doi.org/10.1364/ao.56.003570>
- Kammies, T. L., Manley, M., Gouws, P. A., & Williams, P. J. (2016). Differentiation of foodborne bacteria using NIR hyperspectral imaging and multivariate data analysis. *Applied Microbiology and Biotechnology*, *100*(21), 9305–9320. <https://doi.org/10.1007/s00253-016-7801-4>
- Kamruzzaman, M., EIMasry, G., Sun, D. W., & Allen, P. (2012). Prediction of some quality attributes of lamb meat using near-infrared hyperspectral imaging and multivariate analysis. *Analytica Chimica Acta*, *714*, 57–67. <https://doi.org/10.1016/j.aca.2011.11.037>
- Kandpal, L. M., Tewari, J., Gopinathan, N., Boulas, P., & Cho, B. K. (2016). In-process control assay of pharmaceutical microtablets using hyperspectral imaging coupled with multivariate analysis. *Analytical*

Chemistry, 88(22), 11055–11061. <https://doi.org/10.1021/acs.analchem.6b02969>

- Kerkeni, L., Ruano, P., Delgado, L. L., Picco, S., Villegas, L., Tonelli, F., ... Masuelli, M. (2016). Understanding Color Image Processing by Machine Vision for Biological Materials. *Intech*, (tourism), 13. Retrieved from <https://www.intechopen.com/books/advanced-biometric-technologies/liveness-detection-in-biometrics>
- Khodabakhshian, R., Emadi, B., Khojastehpour, M., Golzarian, M. R., & Sazgarnia, A. (2017). Development of a multispectral imaging system for on-line quality assessment of pomegranate fruit. *International Journal of Food Properties*, 20(1), 107–118. <https://doi.org/10.1080/10942912.2016.1144200>
- Khoshnoudi-Nia, S., & Moosavi-Nasab, M. (2019a). Non-destructive Determination of Microbial, Biochemical, and Chemical Changes in Rainbow Trout (*Oncorhynchus mykiss*) During Refrigerated Storage Using Hyperspectral Imaging Technique. *Food Analytical Methods*, 12(7), 1635–1647. <https://doi.org/10.1007/s12161-019-01494-8>
- Khoshnoudi-Nia, S., & Moosavi-Nasab, M. (2019b). Prediction of various freshness indicators in fish fillets by one multispectral imaging system. *Scientific Reports*, 9(1), 1–11. <https://doi.org/10.1038/s41598-019-51264-z>
- Lalonde, M., Chapdelaine, C., Foucher, S., Team, I., Ave, O., & Qc, M. (2015). *Multispectral imaging: An application to density measurement in photographic paper manufacturing process control*. 9405(February 2015), 1–7. <https://doi.org/10.1117/12.000000>
- Li, J., Meng, X., Wang, W., & Xin, B. (2019). A novel hyperspectral imaging and modeling method for the component identification of woven fabrics. *Textile Research Journal*, 89(18), 3752–3767. <https://doi.org/10.1177/0040517518821907>
- Li, Q., He, X., Wang, Y., Liu, H., Xu, D., & Guo, F. (2013). Review of spectral imaging technology in biomedical engineering: achievements and challenges. *Journal of Biomedical Optics*, 18(10), 100901. <https://doi.org/10.1117/1.jbo.18.10.100901>
- Liang, F., Liu, H., Wang, X., & Liu, Y. (2018). Hyperspectral image recognition based on artificial neural network. *NeuroQuantology*, 16(5), 699–705. <https://doi.org/10.14704/nq.2018.16.5.1244>
- Lim, J., Kim, G., Mo, C., Kim, M. S., Chao, K., Qin, J., ... Cho, B. K. (2016). Detection of melamine in milk powders using near-infrared hyperspectral imaging combined with regression coefficient of partial least square regression model. *Talanta*, 151, 183–191. <https://doi.org/10.1016/j.talanta.2016.01.035>
- Liu, C., Hao, G., Su, M., Chen, Y., & Zheng, L. (2017). Potential of multispectral imaging combined with chemometric methods for rapid detection of sucrose adulteration in tomato paste. *Journal of Food Engineering*, 215, 78–83. <https://doi.org/10.1016/j.jfoodeng.2017.07.026>
- Ma, F., Yao, J., Xie, T., Liu, C., Chen, W., Chen, C., & Zheng, L. (2014). Multispectral imaging for rapid and non-destructive determination of aerobic plate count (APC) in cooked pork sausages. *Food Research International*, 62, 902–908. <https://doi.org/10.1016/j.foodres.2014.05.010>
- Mair, P. (2018). *Principal Component Analysis and Extensions*. (Icces), 179–210. https://doi.org/10.1007/978-3-319-93177-7_6
- Mäkelä, M., Rissanen, M., & Sixta, H. (2020). Machine vision estimates the polyester content in recyclable waste textiles. *Resources, Conservation and Recycling*, 161(June), 105007. <https://doi.org/10.1016/j.resconrec.2020.105007>
- Menesatti, P., Costa, C., & Aguzzi, J. (2010). Quality Evaluation of Fish by Hyperspectral Imaging. *Hyperspectral Imaging for Food Quality Analysis and Control*, 273–294. <https://doi.org/10.1016/B978-0-12-374753-2.10008-5>

- Michael, M., Phebus, R. K., & Amamcharla, J. (2019). Hyperspectral imaging of common foodborne pathogens for rapid identification and differentiation. *Food Science and Nutrition*, 7(8), 2716–2725. <https://doi.org/10.1002/fsn3.1131>
- Mirschel, G., Daikos, O., & Scherzer, T. (2019). In-line monitoring of the thickness distribution of adhesive layers in black textile laminates by hyperspectral imaging. *Computers and Chemical Engineering*, 124, 317–325. <https://doi.org/10.1016/j.compchemeng.2019.01.015>
- Mirschel, G., Daikos, O., Scherzer, T., & Steckert, C. (2017). Near-infrared hyperspectral imaging of lamination and finishing processes in textile technology. *NIR News*, 28(1), 20–25. <https://doi.org/10.1177/0960336016687949>
- Mirschel, G., Daikos, O., Scherzer, T., & Steckert, C. (2018). Near-infrared chemical imaging used for in-line analysis of functional finishes on textiles. *Talanta*, 188(February), 91–98. <https://doi.org/10.1016/j.talanta.2018.05.050>
- Mishra, P., Asaari, M. S. M., Herrero-Langreo, A., Lohumi, S., Diezma, B., & Scheunders, P. (2017). Close range hyperspectral imaging of plants: A review. *Biosystems Engineering*, 164, 49–67. <https://doi.org/10.1016/j.biosystemseng.2017.09.009>
- Mishra, P., & Nordon, A. (2020). Classifying green teas with near infrared hyperspectral imaging. *NIR News*, 31(1–2), 20–23. <https://doi.org/10.1177/0960336019889321>
- Mustafic, A., Jiang, Y., & Li, C. (2016). Cotton contamination detection and classification using hyperspectral fluorescence imaging. *Textile Research Journal*, 86(15), 1574–1584. <https://doi.org/10.1177/0040517515590416>
- Naganathan, G. K., Grimes, L. M., Subbiah, J., Calkins, C. R., Samal, A., & Meyer, G. E. (2008). Visible/near-infrared hyperspectral imaging for beef tenderness prediction. *Computers and Electronics in Agriculture*, 64(2), 225–233. <https://doi.org/10.1016/j.compag.2008.05.020>
- Neath, R. C., & Johnson, M. S. (2010). Discrimination and classification. *International Encyclopedia of Education*, 135–141. <https://doi.org/10.1016/B978-0-08-044894-7.01312-9>
- Onoyama, H., Ryu, C., Suguri, M., & Iida, M. (2013). Potential of hyperspectral imaging for constructing a year-invariant model to estimate the nitrogen content of rice plants at the panicle initiation stage. In *IFAC Proceedings Volumes (IFAC-PapersOnline)* (Vol. 4). <https://doi.org/10.3182/20130828-2-SF-3019.00054>
- Pirouz, D. M. (2012). An Overview of Partial Least Squares. *SSRN Electronic Journal*, (October 2006). <https://doi.org/10.2139/ssrn.1631359>
- Qiao, J., Ngadi, M. O., Wang, N., Gariépy, C., & Prasher, S. O. (2007). Pork quality and marbling level assessment using a hyperspectral imaging system. *Journal of Food Engineering*, 83(1), 10–16. <https://doi.org/10.1016/j.jfoodeng.2007.02.038>
- Qin, J., Chao, K., Kim, M. S., Lu, R., & Burks, T. F. (2013). Hyperspectral and multispectral imaging for evaluating food safety and quality. *Journal of Food Engineering*, Vol. 118, pp. 157–171. <https://doi.org/10.1016/j.jfoodeng.2013.04.001>
- Romaniello, R., & Baiano, A. (n.d.). Discrimination of flavoured olive oil based on hyperspectral imaging. *Journal of Food Science and Technology*, 55. <https://doi.org/10.1007/s13197-018-3160-8>
- Rostom, N. G., Shalaby, A. A., Issa, Y. M., & Afifi, A. A. (2017). Evaluation of Mariut Lake water quality using Hyperspectral Remote Sensing and laboratory works. *Egyptian Journal of Remote Sensing and Space Science*. <https://doi.org/10.1016/j.ejrs.2016.11.002>

- Rungpichayapichet, P., Nagle, M., Yuwanbun, P., Khuwijitjaru, P., Mahayothee, B., & Müller, J. (2017). Prediction mapping of physicochemical properties in mango by hyperspectral imaging. *Biosystems Engineering*, *159*(2011), 109–120. <https://doi.org/10.1016/j.biosystemseng.2017.04.006>
- Seidl, T. (2009). Nearest Neighbor Classification. *Encyclopedia of Database Systems*, *1*, 1885–1890. https://doi.org/10.1007/978-0-387-39940-9_561
- Serranti, S., Palmieri, R., & Bonifazi, G. (2015). Hyperspectral imaging applied to demolition waste recycling: innovative approach for product quality control. *Journal of Electronic Imaging*, *24*(4), 043003. <https://doi.org/10.1117/1.jei.24.4.043003>
- Setser, A. L., & Waddell Smith, R. (2018). Comparison of variable selection methods prior to linear discriminant analysis classification of synthetic phenethylamines and tryptamines. *Forensic Chemistry*, *11*(October), 77–86. <https://doi.org/10.1016/j.forc.2018.10.002>
- Sivertsen, A. H., Kimiya, T., & Heia, K. (2011). Automatic freshness assessment of cod (*Gadus morhua*) fillets by Vis/Nir spectroscopy. *Journal of Food Engineering*, *103*(3), 317–323. <https://doi.org/10.1016/j.jfoodeng.2010.10.030>
- Spyrelli, E. D., Doulgeraki, A. I., Argyri, A. A., Tassou, C. C., Panagou, E. Z., & George-John, E. N. (2020). Implementation of multispectral imaging (MSI) for microbiological quality assessment of poultry products. *Microorganisms*, *8*(4), 1–14. <https://doi.org/10.3390/microorganisms8040552>
- System Overview — Benchtop Assembly for Spectronon Version 3.1.1. (n.d.). Retrieved November 25, 2020, from http://docs.resonon.com/spectronon/BenchtopAssemblyGuide/html/SystemOverview_include.html
- Vidal, M., & Amigo, J. M. (2012). Pre-processing of hyperspectral images. Essential steps before image analysis. *Chemometrics and Intelligent Laboratory Systems*, *117*, 138–148. <https://doi.org/10.1016/j.chemolab.2012.05.009>
- Wallace, M. B., Wax, A., Roberts, D. N., & Graf, R. N. (2009). Reflectance Spectroscopy. *Gastrointestinal Endoscopy Clinics of North America*, *19*(2), 233–242. <https://doi.org/10.1016/j.giec.2009.02.008>
- Wang, K., Pu, H., & Sun, D. W. (2018). Emerging Spectroscopic and Spectral Imaging Techniques for the Rapid Detection of Microorganisms: An Overview. *Comprehensive Reviews in Food Science and Food Safety*, *17*(2), 256–273. <https://doi.org/10.1111/1541-4337.12323>
- Wu, D., Shi, H., Wang, S., He, Y., Bao, Y., & Liu, K. (2012). Rapid prediction of moisture content of dehydrated prawns using on-line hyperspectral imaging system. *Analytica Chimica Acta*, *726*, 57–66. <https://doi.org/10.1016/j.aca.2012.03.038>
- Wulder, M. A., Loveland, T. R., Roy, D. P., Crawford, C. J., Masek, J. G., Woodcock, C. E., ... Zhu, Z. (2019). Current status of Landsat program, science, and applications. *Remote Sensing of Environment*, *225*, 127–147. <https://doi.org/10.1016/j.rse.2019.02.015>
- Xing, J., Bravo, C., Jancsó, P. T., Ramon, H., & De Baerdemaeker, J. (2005). Detecting bruises on “Golden Delicious” apples using hyperspectral imaging with multiple wavebands. *Biosystems Engineering*, *90*(1), 27–36. <https://doi.org/10.1016/j.biosystemseng.2004.08.002>
- Yang, B., Gao, Y., Li, H., Ye, S., He, H., & Xie, S. (2019). Rapid prediction of yellow tea free amino acids with hyperspectral images. *PLoS ONE*, *14*(2), 1–17. <https://doi.org/10.1371/journal.pone.0210084>
- Yu, P., Huang, M., Zhang, M., Zhu, Q., & Qin, J. (2020). Rapid detection of moisture content and shrinkage ratio of dried carrot slices by using a multispectral imaging system. *Infrared Physics and Technology*, *108*(May), 103361. <https://doi.org/10.1016/j.infrared.2020.103361>

Yuan, L., Yan, P., Han, W., Huang, Y., Wang, B., Zhang, J., ... Bao, Z. (2019). Detection of anthracnose in tea plants based on hyperspectral imaging. *Computers and Electronics in Agriculture*, 167(June), 105039. <https://doi.org/10.1016/j.compag.2019.105039>

Zahavi, A., Palshin, A., Liyanage, D. C., & Tamre, M. (2019). Influence of Illumination Sources on Hyperspectral Imaging. *2019 20th International Conference on Research and Education in Mechatronics (REM)*, 5, 1–5. <https://doi.org/10.1109/REM.2019.8744086>

ADDITIONAL READING

Chang, Chein-I. (Ed.). (2003). *Hyperspectral Imaging*. Springer, Boston, MA.

Liyanage, D. C. (2021). *Smart terrain perception using hyperspectral imaging* [Unpublished doctoral dissertation]. Tallinn University of Technology, Tallinn, Estonia.

Martens, H. & Martens, M. (2001). *Multivariate Analysis of Quality: An Introduction*. Wiley & Sons, Inc.

Vinzi, V. E., Chin, W. W, Henseler, J, & Wang, H. (Ed). (2010). *Handbook of Partial Least Squares*. Springer, Boston, MA.

KEY TERMS AND DEFINITIONS

At-line – The measurement task is carried out near the production process. However, they are physically separate processes.

In-situ – The necessary measurements or analysis of the process are conducted on-site and directly integrated into the main production flow.

In-line – The measurement process is directly integrated into the main production process.

On-line – Same as in-situ. The measurement or analysis is done in the exact location where the process occurs.

Off-line – The measurement samples are taken from the process and further analyses separately.

Publication II

D. C. Liyanage, R. Hudjakov, and M. Tamre, "Hyperspectral Image Band Selection Using Pooling," in 2020 International Conference Mechatronic Systems and Materials (MSM), 2020, pp. 1–6.

Hyperspectral Image Band Selection Using Pooling

Dhanushka C. Liyanage*, Robert Hudjakov† and Mart Tamre‡

Dept. of Electrical Power Engineering and Mechatronics

Tallinn University of Technology

Tallinn, Estonia

E-mail: *dhanushka.liyanage@taltech.ee, †robert.hudjakov@taltech.ee, ‡mart.tamre@taltech.ee

Abstract—Hyperspectral images contain hundreds of spectral bands. These bands contain abundant information and more often redundant information. This article presents an unsupervised band selection method to choose most significant spectral image bands from hyperspectral datacube which maximize the relevance and minimize redundancy. The outcome of the research can be used for two purposes. One of them is to fabricate multispectral sensor which is more effective in identifying the subjects of interest. The other purpose is to enhance computational efficiency in classification of the hyperspectral image by processing only the selected spectral bands. Proposed method is based on spectral pooling of hyperspectral data cube. Since spectral pooling methods cause information loss which reduces the discrimination of material classes. This article suggests a novel approach which can minimize the loss while improving material classes discrimination. The proposed technique uses min-max pooling together with reflectance intensity gradient of neighboring pixels of hyperspectral image in spectral data axis. Experiment results of classifying terrain hyperspectral datasets are presented to validate the consistency of image classification which contains several material classes. A comparative analysis has been conducted between max pooling, principle component analysis and the proposed min-max pooling.

Index Terms—hyperspectral imaging, pooling, feature selection, band selection, dimensionality reduction, spectral image classification

I. INTRODUCTION

Hyperspectral images contain hundreds of spectral bands within a single image data cube. These images contain contiguous spectral distributions of the pixels related to the objects in the scene. This gives abundant information about the subjects under inspection. The higher the number of spectral bands, the higher the computational cost for process and classify the hyperspectral image. Most of these spectral bands contain redundant data and also highly correlated [1]. Even though having higher number of spectral bands increases the classification accuracy, it fails and drastically decrease accuracy when it reaches certain band count. This phenomenon is called Hughes phenomenon which demands careful selection of appropriate spectral bands [1]. This is one of the problems in hyperspectral imaging (HSI) tackled as dimensionality reduction [1] [2]. In certain industrial applications such as product quality inspection in production lines, it is required to process the spectral image in real time. Classification of objects based on large number of spectral bands which contains redundant and correlated data with smaller number of training samples causes decrease in classification accuracy. On the otherhand, extracting the most informative bands from

the data cube which are unique to the subject of interest and post process the chosen bands will yield better classification accuracy and computational efficiency. Therefore, the objective of this study is to develop a band selection method which minimizes the redundant and correlated data while taking unique spectral characteristics into consideration.

There are various band selection methods have been proposed by various authors during past decades [3]–[8]. Principle component analysis (PCA) has been widely used for dimensionality reduction and for analysis of HSI images. There have been previous studies on band selection (feature selection) methods based on different techniques. Some of them PCA based variations such as selective principle component analysis based on genetic algorithms [3], Fisher linear discriminants [4], Independent Component Analysis (ICA) [5], statistical methods with skewness – kurtosis parameters [6], convolution neural networks some additional functions such as attention mechanism [7] distance density [8], etc.

In state-of-the-art image recognition algorithms, spatial pooling is one of the typical steps where it uses convolution neural networks (CNN). The pooling operation is used to group a subset of features in a selected region of the image which intends to preserve important feature while discarding irrelevant features. Pooling operation summarizes the outputs of neighboring neuron groups in the same kernel [9]. Various pooling operations are in use for that purpose. These methods can be maximum, average, minimum, stochastic or mixed pooling. In this study, novel gradient based min-max pooling method will be introduced in selecting hyperspectral image bands. The max pooling method with reference to band selection also presented as it is the basis of novel method. Both band selection methods are described, experiment results presented and compared based on selected data set for terrain classification and Indian Pines data set which is widely used in remote sensing research. Even though the proposed band selection (feature selection) method tested in visible and near infrared spectrum for hyperspectral imaging, the same method can be used for the other light wavelength ranges as well.

II. RELATED WORK

Spectral band selection based on mean pooling has been carried out for remote sensing image analysis. This particular research was used support vector machines for classification [10]. Maximin distance algorithm based band selection method proposed by Ganesan and Vasuki. In this method, they estimate

virtual dimensionality of the spectral data cube which defines how many bands needed for class discrimination. According to the maximum length of the band which is maximum normalized reflectance, it selects the characteristic bands [14]. Shape similarity based approach presented by Shijin Li and others in which they propose spectral clustering followed extraction of key points in time series data. Then conditional filtering of mutual information to obtain fewer number of bands [15]. In contrast to these methods, proposed minmax pooling method works on clustering the datacube according to kernels and then searching for extremes of pooling operation.

III. PREPARATION OF DATA SETS

Hyperspectral images were acquired using Specim IQ hyperspectral camera (Specim Spectral Imaging Ltd, Finland) which has spectral range of 400-1000 nm [11]. The camera has spectral resolution of 7 nm and has 204 spectral bands. The hyperspectral data cubes were captured under sunlight. The data cubes were processed, analyzed using MATLAB.

The below Fig. 1 scene was captured to classify terrains with gravel, asphalt and grass. There are different materials such as dry tree trunks, dry leaves, mud and sand present in the image too. Due to complexity of labeling, dry matter, mud and sand considered as one class. In total, the image contains five classes.



Fig. 1. Dataset with asphalt, gravel, grass in a terrain image.

The dataset was calibrated using white and dark references taken while capturing the image. The data cube contains 12-bit resolution reflection data for each band. Therefore, calibrated image kept with the same resolution. In (1) I_i denotes the calibrated reflection intensity at i^{th} band while I_{oi} is the reflectance of original image at that wavelength band. Similarly, W_i and D_i denote the white reference reflectance and dark reference reflectance at respective band.

$$I_i = 4096 \times \left(\frac{I_{oi} - D_i}{W_i - D_i} \right) \quad (1)$$

IV. BAND SELECTION METHOD

Since, the objective of band selection is to extract certain number of spectral bands which characterize the objects in the scene to define spectral bands for a multispectral sensor, initially it is required to define the number of spectral bands desired. Depending on this, the number of pooling iterations calculated together with the size of pooling window for each iteration. To extract the max pooling values from the spectral bands, it is required to re-organize the image where primary axis spatial pixels while secondary axis gives spectral distribution for the given pixel. To do that, the spectral image needs to be converted to 2-D image from 3-D data cube. In the original data cube columns converted into rows of modified image as in the Fig. 2. Purpose of the above transform is to perform pooling operation only on spectral data instead of spatial data.

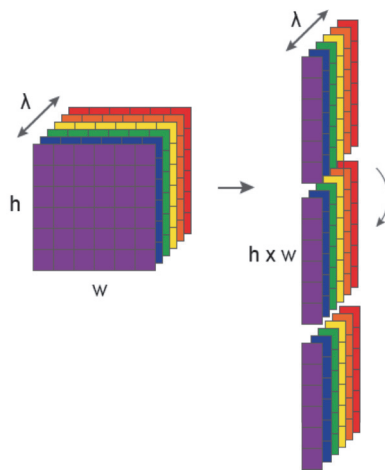


Fig. 2. Unfolding spectral data cube.

A. Max Pooling

This method sequentially finds the largest feature in the spectral bands kernel. The spectral kernels are non-overlapping. Max reflectance of the filter is calculated using (2).

$$y_{\max} = \max_{i=1, j=1}^{h, w} x_{ij} \quad (2)$$

B. Min - Max Pooling (Proposed method)

In contrast to max pooling method, min-max pooling method involves searching for the local minima and maxima (local extremum) in the pooling kernel. In the pooling kernel, there can be significant local minima which doesn't taken into consideration when it is only choose with max pooling. To take such a situation into account, this method searches for minima and maxima in the current n^{th} kernel. The minima calculated using (3).

$$y_{\min} = \min_{i=1, j=1}^{h, w} x_{ij} \quad (3)$$

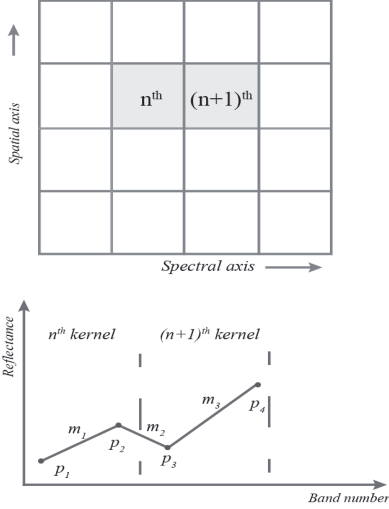


Fig. 3. Min-max pooling.

Similarly, maxima and minima for $(n + 1)^{\text{th}}$ kernel will be calculated using the same equations (2) and (3). Find the maxima and minima coordinates in the spectral axis which are the band numbers. The adjacent gradients calculated between three consecutive extreme locations. Depending on the calculated gradient differences between extreme points, the band of either maxima or minima will be selected from the filter.

In this algorithm, let the consecutive points to be p_1, p_2 and p_3 and the gradients of p_1p_2 as m_1, p_2p_3 as m_2 . The Fig. 3 shows how these points located in spectral axis. If the gradient difference is more than 0.15 which was obtained empirically, the point p_2 will be more significant in the kernel and the else p_1 .

The band number B taken from (4). Below 0.15 taken empirically and depending on the variation of spectral signature, this value can be changed.

$$\mathbf{B} = \begin{cases} p_2, & \text{if } m_2 - m_1 > 0.15 \\ p_1, & \text{otherwise} \end{cases} \quad (4)$$

In this geometric technique, if two consecutive points lie on consecutive bands, then the band with low reflectance intensity will be ignored. Eventhough it can catch subtle points in the spectral signature, it has downside over max pooling. This method takes considerable computing time as it requires to calculate gradient in each iteration.

V. CLASSIFICATION METHOD

For HSI classification, deep feed forward neural network (DNN) has been used. The network comprised on input layer, output layer and four hidden layers. In the input layer, there were 20 neurons. This was according to chosen number of spectral bands. First hidden layer has 20 neurons, second layer with 50 neurons, third layer with 25 neurons and the fourth layer was 5 neurons. Last layer of the network used 'softmax' as activation function where all the other layers were 'purelin'. The output layer has 5 neurons which is based on number of classes in the image.

Training data classes extracted from the same image which corresponds to different materials in the HSI data cube. Therefore, training data labelled according to material in the data cube.

VI. RESULTS

HSI classification performance for the proposed methods are presented in below I. Results include two pooling based methods and widely used PCA method for the chosen data set. Number of bands denoted in N , while detection accuracy for each class presented as a percentage based on its ground truth. Classes named as follows, AP - Asphalt (green), GV - Gravel (red), GR - Grass (blue), DM - Dry matter (black) and SK - Sky (white). Below Fig.4 show the ground truth of the terrain dataset and Fig. 5 shows the classification result for 15 bands using min-max pooling method.



Fig. 4. Ground truth of the dataset.

A. Max Pooling

With max pooling method, the classification accuracy decreases with large number of image bands. However, in overall it gives 86 – 87 % classification accuracy.

B. Min-Max Pooling (Proposed method)

This method gives slightly better accuracy over other methods in overall. Most importantly, it gives higher classification average results for all across the classes even with lower number of spectral bands. The objective of the method is to distinguish spectral signatures of closely correlated classes along with band selection. This objective has been achieved with above results.

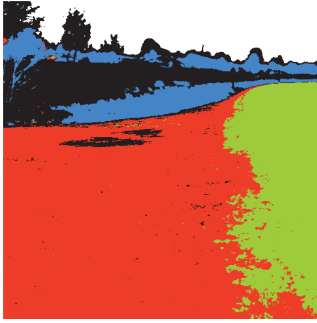


Fig. 5. Classification result for 15 bands using Min-Max Pooling.

TABLE I
TERRAIN HSI DATASET CLASSIFICATION RESULTS

Method	N	Classification accuracy %					
		AP	GV	GR	DM	SK	Overall
Max Pooling	5	81.5	94.2	0	96.9	97.5	86.2
	10	83.0	95.4	83.6	74.3	98.7	89.6
	15	78.5	91.2	84.2	69.7	98.3	86.1
	20	83.1	90.5	84.7	70.3	98.5	86.7
	25	83.9	89.5	85.6	69.2	98.0	86.0
	Avg	82.1	92.2	67.6	76.1	98.2	
Min-Max Pooling	5	73.9	86.8	78.8	77.2	98.1	83.8
	10	82.1	96.1	77.9	77.8	97.7	89.7
	15	86.0	96.0	83.6	76.0	98.9	90.6
	20	85.6	86.1	86.1	80.0	97.8	86.5
	25	83.0	90.4	79.3	76.6	97.9	87.1
	Avg	82.1	91.1	81.1	77.5	98.1	
PCA	5	70.9	79.2	66.5	0	95.3	67.7
	10	78.7	81.5	70.9	0	98.5	70.9
	15	79.7	81.7	68.0	0	97.9	71.0
	20	78.6	78.9	70.6	0	95.9	69.4
	25	82.5	89.2	81.8	0	97.3	75.9
	Avg	82.1	91.1	81.1	0	98.1	

C. Comparison with other methods

Principle component analysis (PCA) method also used to extract spectral bands from the hyperspectral dataset. PCA methods gave lower overall classification accuracy compared to the pooling methods. It couldn't classify dry matter against grass with PCA based spectral bands. Below Fig. 6 shows the overall accuracy of each method.

VII. REMOTE SENSING IMAGE CLASSIFICATION

There are popular remote sensing datasets such as Indian Pines dataset used in most of the researches to compare performance of proposed methods. In this study, the proposed pooling methods also tried out using Indian Pines data set [12]. The corrected dataset contains 145 x 145 pixels with 200 spectral bands which excludes spectral bands containing water absorption spectral signature [13]. The ground truth of the dataset shown in Fig. 7.

The classifier was modified by changing the number of neurons in the last layer to 9. Due to low resolution of the dataset, there weren't sufficient amount of training data for all the 17

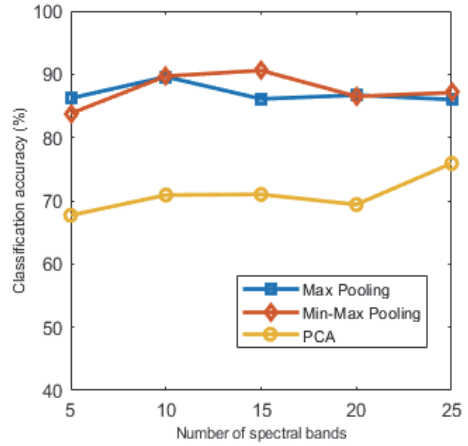


Fig. 6. Comparison of overall accuracy of band selection methods.

classes in the original image. In order to minimize the impact of limited training data on the classification, 9 classes were selected based on number of spatial pixels available in each class sorted in descending order. The classification outcome is shown in Fig. 8 for the chosen 9 classes. Those classes weren't included in the training, were detected according to highest probable class. Even though it has been used as a benchmark, neural network methods require large amount of data for accurate classification.

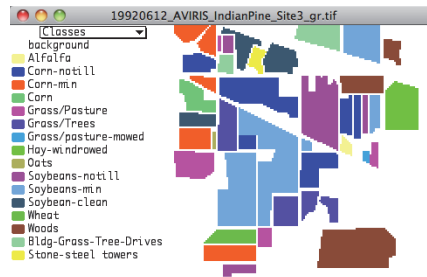


Fig. 7. Ground truth of AVIRIS Indian Pines dataset from Purdue University [13].

TABLE II
INDIAN PINES DATASET CLASSIFICATION RESULTS

Method	Overall accuracy %
Max Pooling	62.39 %
Min-Max Pooling	64.92 %
PCA	56.18 %

Above II, shows the classification results for the Indian Pines dataset. Based on classification results, the pooling methods are performing better than PCA method. Among the



Fig. 8. The HSI classification result for the Indian Pines dataset. Number of classes has reduced.

pooling methods, proposed min –max pooling method shows slightly better classification accuracy. Since number of samples were low, the accuracy of each class has not been investigated.

VIII. EARTH MINERAL MATERIAL CLASSIFICATION

This method was used to classify two earth mineral materials which are Apatite and Tourmaline. The materials have similar visual appearance Fig. 9. The spectral characteristics for each material plotted after obtaining from min-max pooling operation which are shown in Fig. 10 and Fig. 11. The material classification result is shown in Fig. 12. According to the classification result, the materials have been detected with reasonable accuracy. However, earth minerals are mostly contaminated with several other minerals, it might need to experiment with relevant light wavelength ranges which provide information on other contaminants.

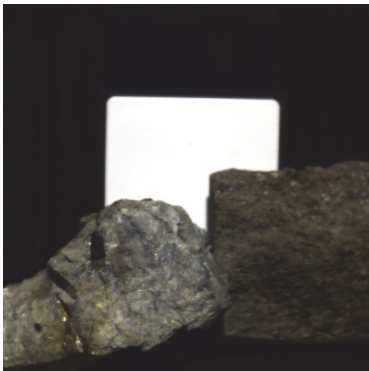


Fig. 9. Earth mineral samples. Left: Tourmaline and Right: Apatite. White reference in the center.

IX. CONCLUSION

Hyperspectral band selection method based on max pooling and min-max pooling has presented in this article. Both pooling methods delivered better classification accuracy compared to PCA method. Experiments has been conducted on terrain image classification and remote sensing public dataset classification with all three methods. Among pooling methods,

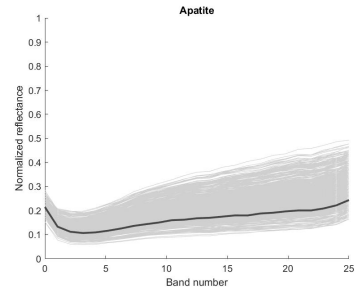


Fig. 10. Spectral signature of Apatite with 25 bands taken from pooling operation .

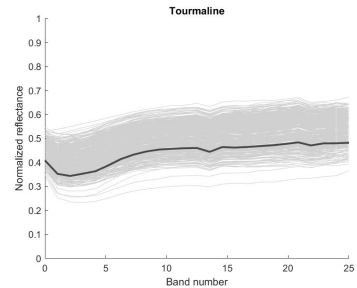


Fig. 11. Spectral signature of Tourmaline with 25 bands taken from pooling operation.



Fig. 12. HSI image classification result showing Tourmaline in cyan and Apatite in red.

the proposed min-max pooling technique shows better classification accuracy among all the classes. The proposed min-max pooling method is capable of capturing distinct characteristics of different classes quite well.

In general, the min-max pooling takes more computing time over max pooling method. However, band selection performed for identifying suitable spectral bands to classify the scene which doesn't affect much for overall classification process.

The only pre-processing step was spectral calibration with white and dark references.

REFERENCES

- [1] W. Ma, C. Gong, Y. Hu, P. Meng, and F. Xu, "The Hughes phenomenon in hyperspectral classification based on the ground spectrum of grasslands in the region around Qinghai Lake," *Int. Symp. Photoelectron. Detect. Imaging 2013 Imaging Spectrom. Technol. Appl.*, vol. 8910, no. August 2013, p. 89101G, 2013.
- [2] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, 2014.
- [3] Liu Ying, Gu Yanfeng, and Zhang Ye, "Hyperspectral Feature Extraction using Selective PCA based on Genetic Algorithm with Subgroups," *First International Conference on Innovative Computing, Information and Control - Volume I (ICICIC'06)*, pp. 652–656, Beijing, 2006.
- [4] X. Cheng, Y. Tao, Y. R. Chen, and X. Chen, "Integrated PCA-FLD method for hyperspectral imagery feature extraction and band selection," *2006 3rd IEEE Int. Symp. Biomed. Imaging From Nano to Macro - Proc.*, vol. 2006, no. 1, pp. 1384–1387, 2006.
- [5] H. Du, H. Qi, X. Wang, R. Ramanath, and W. E. Snyder, "Band selection using independent component analysis for hyperspectral image processing," *Proc. - Appl. Imag. Pattern Recognit. Work.*, vol. 2003-Janua, pp. 93–98, 2004.
- [6] X. Du, H. Chen, Z. Liu, and C. Yang, "A novel unsupervised bands selection algorithm for hyperspectral image," *Optik (Stuttg.)*, 2018.
- [7] P. R. Lorenzo, L. Tulczyjew, M. Marcinkiewicz, and J. Nalepa, "Band Selection from Hyperspectral Images Using Attention-based Convolutional Neural Networks," in *IEEE Access*, vol. 8, pp. 42384–42403, 2020.
- [8] Y. Zhan, D. Hu, H. Xing, and X. Yu, "Hyperspectral Band Selection Based on Deep Convolutional Neural Network and Distance Density," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2365–2369, 2017.
- [9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
- [10] A. Paul and N. Chaki, "Dimensionality Reduction of Hyperspectral Images Using Pooling," *Pattern Recognit. Image Anal.*, vol. 29, no. 1, pp. 72–78, 2019.
- [11] S. S. I. Ltd, "Specim IQ Tech Specs," 2019. [Online]. Available: <https://www.specim.fi/iq/tech-specs/>. [Accessed: 14-Nov-2019].
- [12] M. F. Baumgardner, L. L. Biehl, and D. A. Landgrebe, "220 Band AVIRIS Hyperspectral Image Data Set: June 12, 1992 Indian Pine Test Site 3." *Purdue University Research Repository*, doi:10.4231/R7RX991C, 2015.
- [13] "Hyperspectral Remote Sensing Scenes - Grupo de Inteligencia Computacional (GIC)," [Online]. Available: http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes. [Accessed: 16-Jan-2020].
- [14] V.S.K.Ganesan, and S. Vasuki, "Maximin distance based band selection for endmember extraction in hyperspectral images using simplex growing algorithm," *Multimedia Tools Applications*, vol 77, pp. 7221–7237, Springer, 2018.
- [15] S. Li, J. Qiu, X. Yang, H. Liu, D. Wan, and Y. Zhu, "A novel approach to hyperspectral band selection based on spectral shape similarity analysis and fast branch and bound search," *Engineering Applications of Artificial Intelligence*, vol. 27, pp. 241–250, 2014.

Publication III

D. C. Liyanage, R. Hudjakov, and M. Tamre, "Hyperspectral Imaging Methods Improve RGB Image Semantic Segmentation of Unstructured Terrains," in 2020 International Conference Mechatronic Systems and Materials (MSM), 2020, pp. 1–5.

Hyperspectral Imaging Methods Improve RGB Image Semantic Segmentation of Unstructured Terrains

Dhanushka C. Liyanage*, Robert Hudjakov[†] and Mart Tamre[‡]

Dept. of Electrical Power Engineering and Mechatronics

Tallinn University of Technology

Tallinn, Estonia

E-mail: *dhanushka.liyanage@taltech.ee, [†]robert.hudjakov@taltech.ee, [‡]mart.tamre@taltech.ee

Abstract—Terrain recognition for off-road unmanned ground vehicles driven on unstructured terrains are far complex in contrast to terrain recognition for road vehicles which are driven on structured terrains. Large number of terrain classes mixed-up in unstructured terrains make it difficult to classify using convolution neural networks based on RGB images. Partly it is attributed to lack of sufficiently annotated training data for neural network, and partly it is difficult to label such a large number of object classes which shows visual similarities. Introducing additional details about the scene with hyperspectral or multispectral cameras, the scene classification can be greatly improved for annotation of training data for neural network training. Using spectral signatures of different materials, hyperspectral imaging can detect different materials in the scene. This article discusses a method to annotate RGB images and semantic segmentation for autonomous driving on unstructured terrain applications by using hyperspectral imaging. The RGB images will be generated using same hyperspectral data cube by extracting certain spectral bands in the visible light spectrum. Using semantic segmentation network *ResNet18*, manually annotated training data will be compared with hyperspectral method assisted annotated data by classifying terrain scenarios.

Index Terms—Image annotation, hyperspectral, unstructured terrain recognition, UGV

I. INTRODUCTION

Machine vision for autonomous ground vehicle navigation is one of the rapidly evolving research area in the recent years. In those applications, recognition of different terrains, road surfaces and objects are done by using Deep Convolutional Neural Networks (DCNN) based on RGB images. Object detection accuracy highly depends on the labelled training image data given to the network. These labelled training data coming from assigning class label to pixels in the image. Semantic labeling can be done by manually or by using annotation tools. Manual labeling is a time consuming and expensive task done by human labellers. On the otherhand, there are various commercial tools and research work to automate image labeling process to reduce labour cost for manual labeling [1] [2] [3] [4]. The annotation can be either strong-annotation which involved labelling every pixel in the image or weak-annotation which is defining bounding box around the object.

However, researchers claim that state-of-the-art segmentation accuracy can be achieved in image semantic segmentation using fully-supervised methods. Modern weakly supervised networks produce inferior results while combining weakly-annotated images with small number of strongly-annotated images can produce comparative results [5]. Therefore it is necessary to have fully supervised strongly-annotated images for state-of-the-art segmentation. There are various CNNs developed over the time to improve semantic segmentation [6], [7]. Several other researches indicate use of multi-modal approaches to enhance semantic segmentation [8], [9], [10]. Using RGB-D images and four-stage CNN for classification, Nico and others have used depth maps to improve segmentation accuracy [8]. Transforming 3D LiDAR point clouds to spherical image, PointSeg network [10] has been developed for semantic segmentation even though it suffers small object detection performance.

Off-road autonomous navigation is far more complex than urban road navigation due to complexity of the terrain. Different terrain classes such as swamps, water puddles, wetlands, mud, etc. have close visual appearance which reduces the detection using RGB images [11]. For off-road terrain semantic segmentation, it has been proved that additional 3D Lidar fusion with images can help segmenting roads and muddy water puddles which is rather complex terrain scenario [11]. Fusing additional sensors such as LiDARs, depth-cameras, the terrain detection can be enhanced even under low light conditions [12].

Using Hyperspectral Imaging (HSI) method, this article investigates an alternative approach for image labelling. Hyperspectral image comprises of hundreds of contiguous spectral bands in a single data cube. The reflectance characteristics at different wavelength bands for different materials produce unique signature. Based on this so called spectral signature, image pixels of similar materials can be identified. Similar approach has been used in remote sensing using Multispectral images (MSI) in visible Near infrared (VNIR) wavelength range [13]. Since HSI contains more spectral bands than MSI, the aforementioned HSI method can provide higher accuracy in image classification with the help of suitable band selection.

Therefore, scene labeling can be improved by combining HSI classification with RGB annotation which reduces the manual segmentation operation.

Objective of this study is to investigate HSI image classification possibilities to assist annotation of RGB images as training data for neural networks. This study involves shallow neural network for HSI classification. For the segmentation results comparison, in addition to HSI images, RGB images were used with semantic segmentation network called ‘resnet18’.

II. RELATED WORK

Since the focus of this work is to investigate RGB scene labeling with the help of HSI, object labeling with spectral-spatial data were explored. Use of hyperspectral images for terrain classification in UGV applications has been researched using snapshot hyperspectral cameras by C. Winkens and others [14]. Their study was based on random forest classifier where the authors claim that VNIR camera based method capable of distinguishing drivable areas from non-drivable areas [14]. Instead of vehicle mounted or ground based sensors, using aerial sensor platform which comprises of HSI, RGB camera and Lidar, urban scene labeling has been investigated and presented in reference [15]. Similar research has been done by Eslami and others with hyperspectral thermal infrared imagery and visible imagery for urban object detection such as roofs, trees, vegetation and soil [16]]. Instead of single multispectral camera, Valada and others have used two RGB cameras and removing Near InfraRed (NIR) filter from one of the cameras to capture RGB + NIR images for semantic segmentation. The authors claimed that their NIR wavelengths fusion with RGB provide more accurate segmentation in forested environments [17].

With the findings of all these studies, it is a valid hypotheses to investigate the semantic segmentation capabilities of HSI in order to reduce labour involved in pixel-wise labeling of RGB images.

III. EXPERIMENTS

The experiments were conducted including HSI classification using shallow neural networks and comparing the results with best performing semantic segmentation deep neural network results for RGB images.

A. Image dataset

Hyperspectral images were captured by using Specim IQ mobile hyperspectral camera in VNIR range of 400 – 1000 nm. The camera captures hyperspectral data cube with 204 spectral bands, 7 nm spectral resolution. The spatial resolution of the hyperspectral sensor is 512 px x 512 px and it provide RGB image with 1200 px x 720 px resolution from viewfinder camera [18]. The images were acquired under sunlight. Specim IQ camera uses push-broom image acquisition technique as it comprises of line scan imaging sensor [19].

The data sets were captured in Estonia covering different terrain classes which includes muddy areas, water puddles,

grass, gravel roads, stone roads, paved roads and various natural objects. The captured hyperspectral images were weakly annotated for each terrain class in order to perform spectral image classification. Below Fig 1 shows the spectral data samples taken for grass, mud, trees, sky and calibration object classes.

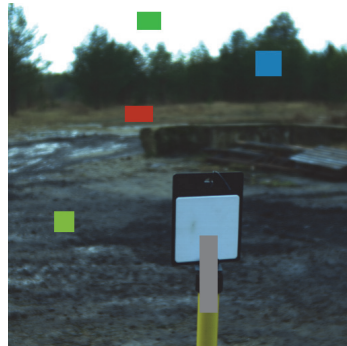


Fig. 1. Training data taken for spectral classifier training. Terrain classes are, Blue: trees, Green: sky, Light green: mud, Red: grass and Light brown: calibration object.

Several HSI data cubes have been used to extract training data samples in a similar fashion which covers all the above mentioned unstructured terrain surface classes.

B. Band reduction

The terrain surface classes in spectral images need to be able to use as general signature for the same object type available in anywhere else. Since HSI datacubes contain hundreds of bands, band selection has been carried out in order to eliminate redundant data from the spectral image while preserving characteristic information pertaining to a particular terrain surface class. The Min-Max pooling method has been used for band selection which was proposed by us [20]. These experiments have been conducted by using 25 spectral bands taken from the HSI data cube.

C. HSI image classification

For HSI classification, 4 layered neural network has been used. The network comprised on input layer, output layer and two hidden layers. In the input layer, there were 25 neurons. This was according to chosen number of spectral bands. First hidden layer has 50 neurons, second hidden layer with 25 neurons, and output layer was comprised of 7 neurons. Last layer of the network used ‘softmax’ as activation function where all the other layers were ‘purelin’. The output layer has 7 neurons which is based on number of classes considered for the classification. The neural network training was done by using graphics processing unit (GPU).

D. Comparison with RGB image segmentation

For RGB image semantic segmentation, the training RGB images were taken from the specim IQ mobile hyperspectral

camera which provides false-*RGB* image corresponds to HSI data cube. Considerable number of those false-*RGB* images were pixel-wise annotated and used as training data. And also false-*RGB* image brightness was manually altered as a pre-processing step before used in segmentation. Since the number of training images were insufficient, camvid open source dataset has been added by cropping the images into 512 px x 512 px resolution.

The 'resnet18' deep neural network has been used for the image semantic segmentation for comparative analysis. This network contains fewer number of classes compared to state-of-the-art ResNet101. Since the network with less number of parameters can be trained with fewer number of images in comparison to the large networks which have large number of parameters.

However, it is required to generate false-*RGB* images from the algorithm by adjusting the image brightness dynamically. This will be covered as part of future work.

IV. RESULTS

The test scenes were selected with muddy unstructured terrain which has grass, trees and water puddles Fig 2 which is scene 1 and more complex terrain with water stream which filled with dry leaves fallen from nearby trees as scene 2 shown in Fig 3. The ground truth shown in Fig 4 which is the actual semantic segmentation needs to obtain from the classification. The obtained classification results are shown in Fig 5 which resulted in 81.9% overall pixel level accuracy. The classification accuracy for each class presented in Table I. Grass area has been detected with 98% accuracy while water puddles were not detected with considerable accuracy which stands at 34%. However, by considering overall classification result, the resultant classification images can be used as labelled data for semantic segmentation network training.

The corresponding false-*RGB* image was tested using 'resnt18' network. The segmentation result is poor compared to hyperspectral image classification. One of the reason is small number of annotated training images were used for train the network. This was due to unavailability of annotated unstructured terrain *RGB* image datasets. Large amount of training data require to achieve better segmentation result. For the scene 1, corresponding 'resnet18' *RGB* image segmentation result shown in Fig 6.

However, for the more complex scene (scene 2) which has multiple classes mixed together, the overall classification accuracy stands at 77.3%. The accuracy for each class shown in Table II. Ground truth for the scene and HSI result shown in Fig 7 and Fig 8 respectively. Similar to the scene 1, the scene 2 also suffers from poor classification accuracy for water bodies. Grass patch wasn't differentiated from wet fallen leaves on the stream in HSI classification. Muddy and water stream areas of the scenes also contains considerable amount of dry leaves which affected the result. In the case of unstructured terrains, it will never be a clean homogeneous terrain with one terrain surface class. Therefore additional processing methods need to be incorporated to improve the results. On the otherhand,

different wavelength range can help improve the classification such as short wave infrared (SWIR). The *RGB* image classification comparison using 'resnet18' gave poor result for the scene 2 as well. In contrast, HSI classification performs better than *RGB* semantic segmentation.



Fig. 2. Scene 1 - false *RGB* image of HSI datacube with muddy terrain.



Fig. 3. Scene 2 - false *RGB* image of HSI datacube which contains water stream full of dry leaves.

TABLE I
SCENE 1 - HYPERSPECTRAL IMAGE CLASSIFICATION RESULTS

Object type	Label color	Accuracy %
Muddy terrain	Blue	85.4%
Grass	Red	98.1%
Trees	Yellow	82.1%
Water	Light green	34.3%
Calibration target	Grey	54.4%
Sky	Dark green	71.0%

All the neural network classification experiments have been performed using MATLAB r2019b.



Fig. 4. Scene 1 - Ground truth of the muddy terrain. Refer Table I for label colors.



Fig. 7. Scene 2 - Ground truth of the complex terrain containing water stream.

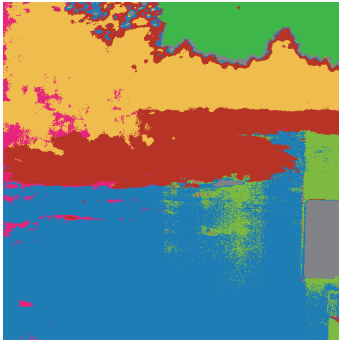


Fig. 5. Scene 1 - HSI classification for the muddy terrain. Refer Table I for label colors.

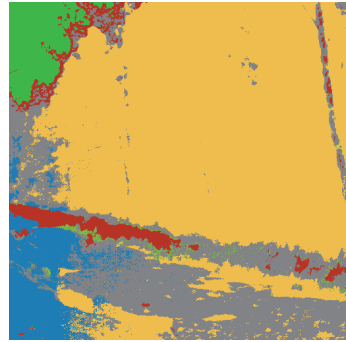


Fig. 8. Scene 2 - HSI classification result. Refer Table II for label colors.

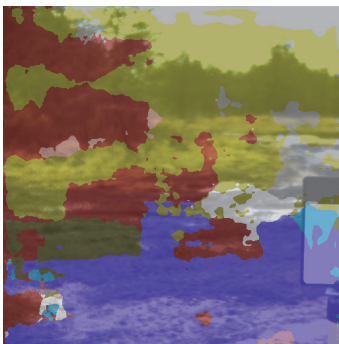


Fig. 6. RGB image semantic segmentation using resnet18. Purple color shows the muddy ground, yellow color shows the trees and grey indicates sky. Red color stands for objects like buildings which resulted in wrong segmentation.

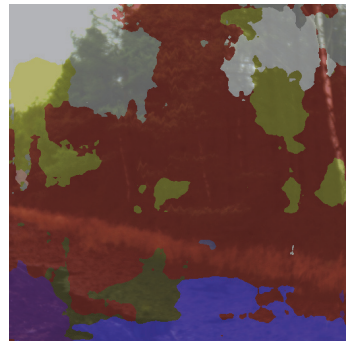


Fig. 9. RGB image semantic segmentation using resnet18. Purple color shows the muddy ground, yellow color shows the trees and grey indicates sky. Red color stands for objects like buildings which resulted in wrong segmentation.

TABLE II
SCENE 2 - HYPERSPECTRAL IMAGE CLASSIFICATION RESULTS

Object type	Label color	Accuracy %
Muddy terrain	Blue	65.7%
Grass	Red	33.7%
Trees	Yellow	89.7%
Water	Light green	2.4%
Wet leaves on stream	Grey	74.4%
Sky	Dark green	77.1%

V. CONCLUSION

In this article we proposed HSI based approach for pixel-wise image annotation to reduce workload of human annotation. With shallow neural network based hyperspectral image classification gives considerably accurate pixel-wise classification in certain scenarios. The demonstrated classification results were achieved by using few hundreds of training pixels from few HSI data cubes. Considering manual pixelwise labeling of a RGB image, such HSI training pixel extraction takes very little time. There were some classes mixed up with others such as grass detected on the top edges of tree line and wet-fallen leaves mixed up with dry grass which degraded classification of those terrain classes. These artefacts can be removed by post processing. However, for more complex terrains, the classification method need to be improved. The obtained HSI classification result can be used as pixel-wise annotated label images for neural networks. The training images for the neural network can be obtained from HSI data cube as false-RGB images. In conclusion, with the help of hyperspectral imaging, pixel-wise labeling can be used to reduce manual labelling process.

VI. FUTURE WORK

This research is currently on-going in order to develop efficient pixel-wise labeling method for training deep neural networks for semantic segmentation. One of the future task is to use pixel-wise classification images taken from HSI classification as the ground truth along with false-RGB images to train semantic segmentation neural network and measure its segmentation performance.

The false-RGB images are required to automatically generate from HSI datacubes according to CIE color matching functions. Due to ambient light variation during HSI acquisition, it is essential to compensate illumination with dynamic image brightness adjustment.

The comparison neural network which was 'resnet18' for the RGB image segmentation was not performing upto its highest accuracy which needs to be trained with large number of strongly-annotated data.

REFERENCES

[1] A. Dutta and A. Zisserman, "The VIA Annotation Software for Images, Audio and Video," in Proceedings of the 27th ACM International Conference on Multimedia - MM '19, 2019, pp. 2276–2279.
[2] "Supervise.ly - Web platform for computer vision. Annotation, training and deploy." [Online]. Available: <https://supervise.ly/>. [Accessed: 29-Jan-2020].

[3] "Image Annotation - Image tagging services for Computer Vision." [Online]. Available: <https://www.imageannotation.ai/>. [Accessed: 30-Jan-2020].
[4] "Labelbox: The leading training data solution." [Online]. Available: <https://labelbox.com/>. [Accessed: 29-Jan-2020].
[5] G. Papandreou, L. C. Chen, K. P. Murphy, and A. L. Yuille, "Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation," Proc. IEEE Int. Conf. Comput. Vis., vol. 2015 Inter, pp. 1742–1750, 2015.
[6] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 12, pp. 2481–2495, 2017.
[7] L.-C. Chen, G. Papandreou, S. Member, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 834–848, 2018.
[8] N. Höft, H. Schulz, and S. Behnke, "Fast Semantic Segmentation of RGB-D Scenes with GPU-Accelerated Deep Neural Networks," in KI 2014: Advances in Artificial Intelligence - 37th Annual German Conference on AI, Stuttgart, Germany, September 22–26, 2014. Proceedings, 2014, vol. 8736, pp. 80–85.
[9] D.Feng, C. Haase-Schuetz, L.Rosenbaum, H.Hertlein, C.Glaeser, F.Timm, W.Wiesbeck and K.Dietmayer, "Deep multi-modal object-detection and semantic segmentation for autonomous driving: datasets, methods and challenges", arXiv:1902.07830, 2019.
[10] Y. Wang, T. Shi, P. Yun, L. Tai, and M. Liu, "PointSeg: Real-Time Semantic Segmentation Based on 3D LiDAR Point Cloud", arXiv:1807.06288, 2018.
[11] S. Laible, Y. N. Khan, K. Bohlmann, and A. Zell, "3D LIDAR- and camera-based terrain classification under different lighting conditions." Proc. Autonomous Mobile Systems 2012, Inform. aktuell, pp. 21–29, Springer, 2012.
[12] Y. Wang, W.-L. Chao, D. Garg, B. Hariharan, M. Campbell, and K. Q. Weinberger, "Pseudo-LiDAR from Visual Depth Estimation: Bridging the Gap in 3D Object Detection for Autonomous Driving", arXiv:1812.07179, 2018.
[13] R. Kemker, C. Salvaggio, and C. Kanan, "Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning," ISPRS J. Photogramm. Remote Sens., vol. 145, pp. 60–77, 2018.
[14] C. Winkens, F. Sattler, and D. Paulus, "Hyperspectral terrain classification for ground vehicles," VISIGRAPP 2017 - Proc. 12th Int. Jt. Conf. Comput. Vision, Imaging Comput. Graph. Theory Appl., vol. 5, no. Visigrapp, pp. 417–424, 2017.
[15] M. Weinmann and M. Weinmann, "Urban scene labeling based on multi-modal data acquired from aerial sensor platforms," 2019 Jt. Urban Remote Sens. Event, JURSE 2019, pp. 1–4, 2019.
[16] M. Eslami and A. Mohammadzadeh, "Developing a Spectral-Based Strategy for Urban Object Detection From Airborne Hyperspectral TIR and Visible Data," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 9, no. 5, pp. 1808–1816, 2016.
[17] A. Valada, G. L. Oliveira, T. Brox, and W. Burgard, "Deep Multispectral Semantic Scene Understanding of Forested Environments Using Multi-modal Fusion," Proc. 2016 International Symposium on Experimental Robotics, pp. 465–477, Springer 2017.
[18] S. S. I. Ltd, "Specim IQ Tech Specs," 2019. [Online]. Available: <https://www.specim.fi/iq/tech-specs/>. [Accessed: 14-Nov-2019].
[19] J. Behmann et al., "Specim IQ: Evaluation of a new, miniaturized handheld hyperspectral camera and its application for plant phenotyping and disease detection," Sensors (Switzerland), vol. 18, no. 2, 2018.
[20] D.C. Liyanage, M.Tamre, R.Hudjakov, "Hyperspectral image band selection using pooling," in Proceedings of MSM2020, Poland., [Unpublished].

Publication IV

A. Zahavi, A. Palshin, D. C. Liyanage, and M. Tamre, "Influence of illumination sources on hyperspectral imaging," Proc. 2019 20th Int. Conf. Res. Educ. Mechatronics, REM 2019, vol. 5, pp. 1–5, 2019.

Influence of Illumination Sources on Hyperspectral Imaging

Ali Zahavi, Andrei Palshin, Dhanushka Chamara Liyanage, Mart Tamre

Mechatronics and Autonomous Systems Centre
School of Engineering, Tallinn University of Technology
Tallinn, Estonia

ali.zahavi@taltech.ee, andreypalshin@gmail.com, dhanushka.liyanage@taltech.ee, mart.tamre@taltech.ee

Abstract—This paper discusses the influence of different illumination sources on hyperspectral imaging. In order to find the best illumination source for hyperspectral image acquisition, sun, fluorescent, light-emitting diodes (LED) and incandescent light sources have been used. The hyperspectral images were acquired using two hyperspectral cameras, Resonon Pika II in the range of 400-1000 nm and Resonon Near Infrared camera in the range of 1000-1700 nm. Then, the results have been compared with the literature. This chapter of the paper suggests the incandescent lights as the best illumination source in HSI. The second part of the paper presents light absorption spectra of different materials including Aluminium Oxide (Al_2O_3), Aluminium Oxide plus Cubic Boron Nitride (cBN), Wood and unknown Metal Disk by using two hyperspectral cameras, Resonon Pika II and Resonon Near Infrared by the presence of incandescent lights as the illumination source. All experiments have been done in Mechatronics and Autonomous Systems Centre of Tallinn University of Technology.

Keywords— *Hyperspectral imaging; Illumination sources; Incandescent light;*

I. INTRODUCTION

Hyperspectral Imaging (HSI), similar to other spectral imaging, processes the information from the electromagnetic spectrum in order to find or identify different materials or objects by obtaining the spectrum of each pixel in the image of a scene [1][2].

Like other Machine Vision's applications, illumination source plays a vital role in the validity of extracted data. As a practice, we experimented HSI by the presence of different illumination sources including sun, fluorescent, LED and incandescent to find the best light source for hyperspectral image acquisition.

After finding the best illumination source for HSI, we did some experiments to obtain light absorption spectra graphs of different materials in the visible range and near-infrared (NIR).

II. INVESTIGATING THE INFLUENCE OF ILLUMINATION SOURCE

A. Sun

In the literature, it has been investigated (Fig.1) the Solar Spectrum Distribution entering the lower part of the atmosphere [3]. As one can see, the intensity is highest in the light blue region, around 460 nm.

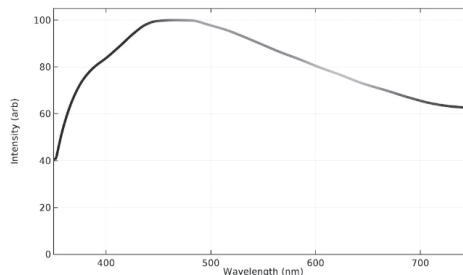


Fig. 1. Solar spectral distribution entering the lower parts of the atmosphere

In a HSI intensity distribution graph of an object (having some has extra peaks or depths), we will be able to state that probably these peaks or depths belong to a material or object absorption. The certain statement needs more experiments with different lighting sources. If the peaks or depths remain in all the graphs for different illumination sources, we certainly can say they belong to the material or object.

The simple spectral distribution of the illumination source will help to distinguish peaks or depths related to the material absorption from the illumination source. From this aspect, the sun seems good for being as a test example the illumination source. However, many factors may affect the above graph (Fig.1) such as geographical location, air pollution and etc.

B. Fluorescent

A typical fluorescent lamp consists of a low-pressure mixture of mercury and a rare gas like Argon. Next figure (Fig. 2) shows the Fluorescent lamp spectral distribution. As one can see in the graph, the fluorescent light spectrum is full of peaks and depths. It means this illumination source will emit different amount of energy in different wavelengths. It is recommended not to use the fluorescent light for hyperspectral image acquisition because it will be hard to distinguish if peaks or depths belong to this source from the peaks for material absorption. Also, there are two major peaks for this light source around 540 nm and 620 nm which mean in these two wavelengths, fluorescent light source emits a large amount of energy to the environment.

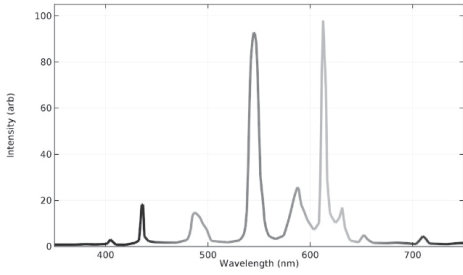


Fig. 2. Fluorescent Spectral Distribution

C. Light-Emitting Diodes (LED)

These days LEDs are known as an efficient illumination source in the industry because they are often more efficient in term of luminous and more durable than traditional incandescent lights [4]. In the following figure, one can find LED light spectral distribution (Fig.3). As can be seen in the graph, there is only one peak around 450 nm (Blue) which means this source will emit a large amount of energy in this wavelength. Because unlike incandescent bulbs, LEDs emit light over a narrow range of wavelengths, we can state that LED light seems good enough for the visible range of spectrum and we can distinguish the peaks for the light source from peaks or depths belong to material (object) absorption. However, from the experiment that we have done with hyperspectral cameras in the range of 400 nm to 1000 nm, it seems that beyond 750 nm, LED spectrum will have many peaks and depths (recognized as noise). Therefore, it is suggested to use LED light only in the range of 380 nm to 700 nm.

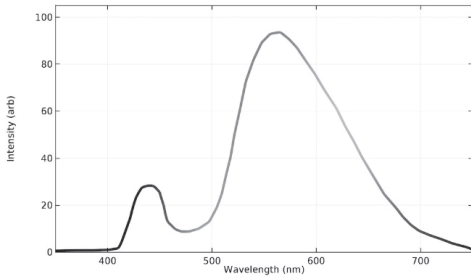


Fig. 3. LED Spectral Distribution

D. Incandescent

An incandescent light bulb has a tungsten filament which is heated when the current is conducted through it. In the following figure (Fig. 4), one can find the incandescent lamp spectral distribution. Although it is for an ideal incandescent light, we can understand that if we see some depths during HSI in the visible range, it can be interpreted as material (object) absorption.

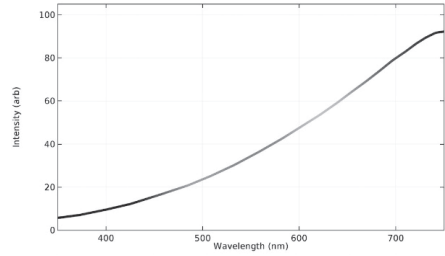


Fig. 4. Incandescent Bulb Spectral Distribution

III. BEST ILLUMINATION SOURCE

As it is mentioned before in this paper, simple light source spectral distribution will help to distinguish peaks or depths related to material absorption from illumination source. From this aspect, all typical light sources except fluorescent seem good for different machine vision applications such as HSI in the visible range. Also, LED emits light in a narrow range of wavelength which causes to be good for visible range, not NIR. Furthermore, as it is discussed, many factors will have an effect on the amount of sun's energy which is entering the lower part of the atmosphere. Therefore, incandescent light bulb has chosen as the best light source for the different machine vision application such as HSI. In the following table (Table.1), one can find the best operational range of each illumination source.

TABLE I. COMPARISON OF LIGHT SOURCES

Light sources	380-700nm	400-1000nm
Sun	✓	✓
Incandescent	✓	✓
Fluorescent	✗	✗
LED	✓	✗

IV. EXPERIMENTS

All the experiments have been performed in a dark room with only the presence of incandescent light as the illumination source without any external light interference. In the following figure (Fig. 5) presented the measurement setup.

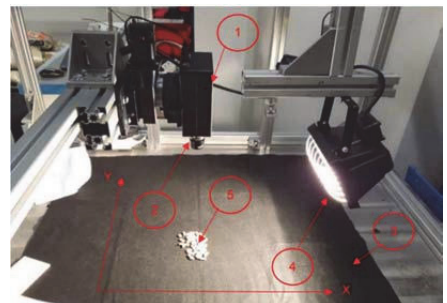


Fig. 5. Measurement setup, 1. Camera 2. Optic lens 3. light-absorbing material 4. Incandescent light 5. Material or Object

The camera was located directly above the investigated material, and the light sources were located to the left and right of the object. The material is located on the surface covered with light-absorbing material.

For each material, a series of images with two hyperspectral cameras were acquired and the best results have been selected for analyzing. The graphs represent the average value of the spectral distribution on the part of a surface of the material.

Abbreviation NIR is short for the camera that operating in the near infrared range (Resonon Near Infrared), and VR for the camera operating in the visible range (Resonon Pika II).

Three vertical lines on VR graph is related to RGB colour spectra. The first one is for blue, the second is for green and the last is for red.

Below presented experiment results for aluminium oxide, aluminium oxide with cubic boron nitride, wood and metal.

A. Aluminium oxide (Al₂O₃)

In the next figure (Fig.6), taken by NIR, the spectrum of aluminium is almost flat over the entire range. We do not see the characteristic peaks of aluminium because they are outside [5] of the NIR camera range [6].

However, we can see from this graph that when using an incandescent lamp, we minimize the effect of the light source on the absorption of the material due to the fact that the light source emission spectrum is almost uniform over the entire considered range.

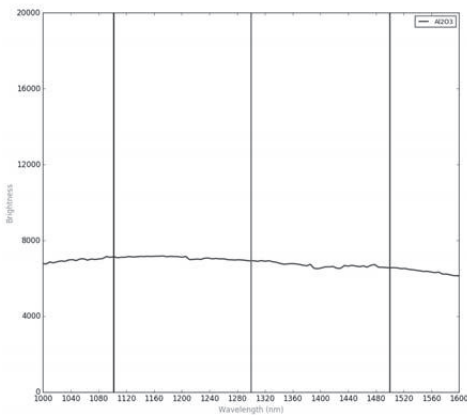


Fig. 6. NIR image of AL₂O₃

The next graph (Fig.7) shows an image taken by VR camera in the range of 400 – 1000 nm and it is very close to the reference incandescent light bulb graph (Fig.4). Its brightness increases from 300 nm to 900 nm on the whole interval.

In the experiment, incandescent light was used, and the illumination of the material was uneven, for this reason, we can observe small fluctuations that slightly affect the intensity of light absorption by the material.

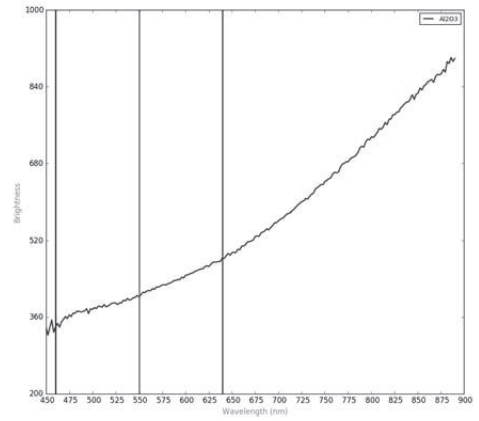


Fig. 7. VR image of AL₂O₃

B. Aluminium oxide (Al₂O₃) plus cubic boron nitride (cBN)

To further explore the possibility of working with the material, we took aluminium oxide interspersed with a cubic form of boron nitride (cBN) [7].

The below figure (Fig.8) is different from what was seen when we examined the hyperspectral image of pure aluminium oxide. The reason for the change in the amount of light absorption is that the position of the material relative to the light source has been changed. We can observe similar oscillations of about 1380 nm and the graph is similar to the previous one. The difference is only in the amount of light on the surface.

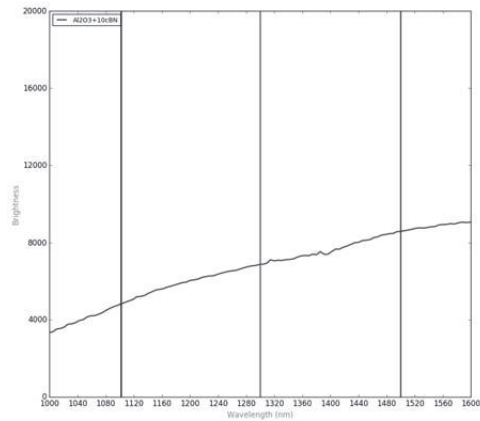


Fig. 8. NIR image of Al₂O₃ + 10% cBN

Similar results have been seen when considering a hyperspectral image taken in the visible range (Fig.9). Light absorption changes in the same way as the incandescent light bulb spectrum. The change in the amount of cBN in aluminium had no significant effect on the spectrum.

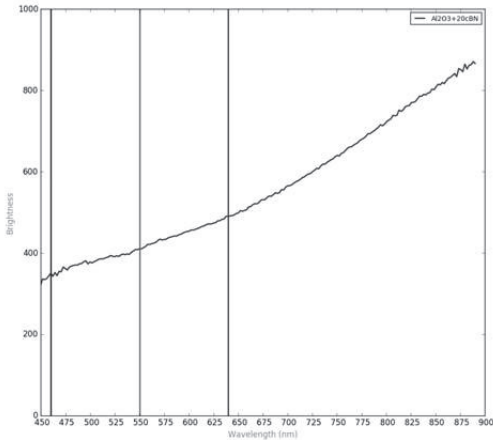


Fig. 9. VR image of Al₂O₃ + 20% cBN

An increase in the amount of cBN in aluminium oxide did not affect the infrared range either (Fig.10). The spectrum is almost the same as that of pure aluminium oxide.

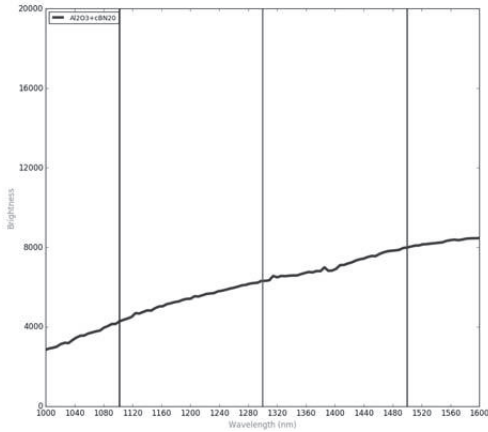


Fig. 10. NIR image of Al₂O₃ + 20% cBN

The absorption spectrum peaks for aluminium oxide and cBN are outside the operating range of our cameras. Changes in the composition of the material also had no effect on the spectra captured by cameras. However, we can clearly see that incandescent light source do not add any significant peaks to the result and we can easily suppress its influence in any calculations.

C. Metal Disc

In the hyperspectral image of the metal disk (Fig.11) made in the infrared spectrum range, we see that the reflection level is much higher and in near-infrared range, the brightness starts from 9000 and goes up to 14000. The graph shows small peaks at 1060 nm and 1330 nm, but on their basis, it is impossible to determine the exact composition of this material.

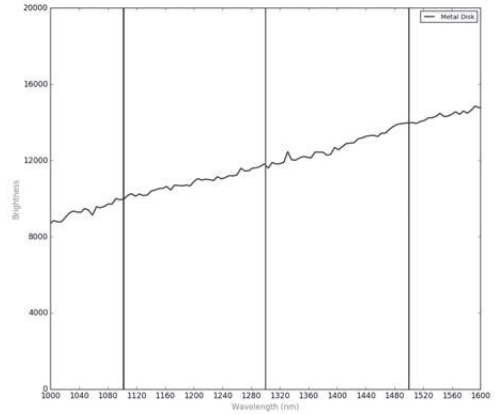


Fig. 11. NIR image of the metal disc

The graph in the next figure (Fig.12) is almost flat and there are no peaks on it. In this case, the light source also does not distort the data.

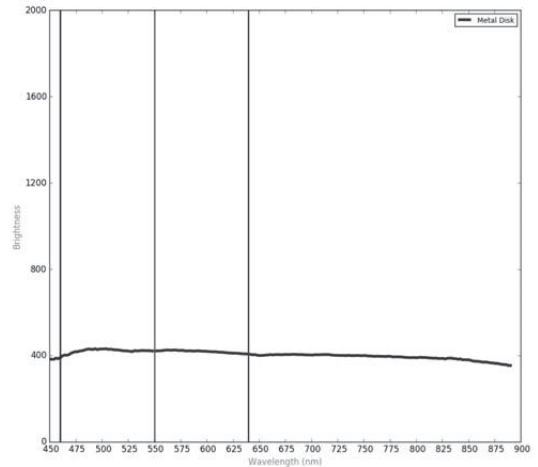


Fig. 12. VR image of the metal disc

D. Wood

The last material that we took for experiments is wooden chips consisting of different wood. Unlike previous materials, part of the absorption spectrum of wood chips is in the field of view of the infrared camera (Fig.13). Peaks on the chart clearly indicate the presence of water in the tested material [8].

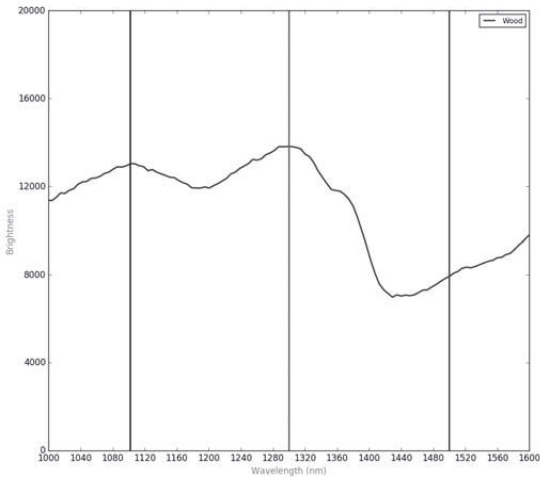


Fig. 13. NIR image of wood

The last spectrum (Fig.14) contains data about the wood in the visible range. It is similar to incandescent bulb spectra. There is no data about the material in this range.

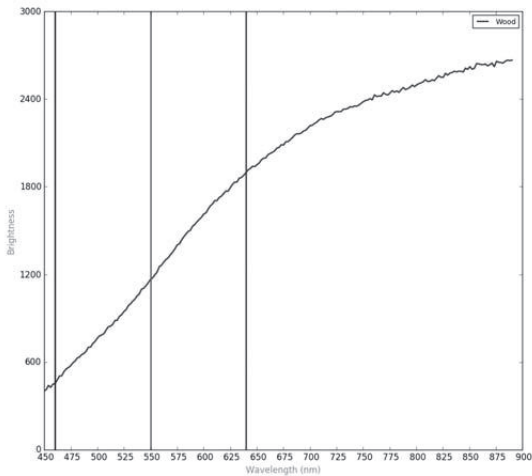


Fig. 14. VR image of wood

V. CONCLUSION

Graphs of light absorption of all these materials show us that the use of incandescent lamps for hyperspectral imaging is optimal from the point of view of covering a wide spectral range of light. This light source adds no significant peaks throughout the operating spectrum. This makes it easy to neutralize its influence in the analysis of the results of hyperspectral imaging and further calculations.

Furthermore, significant peaks belong to water have been observed in the NIR range, although other tested material did not show any peaks neither in visible range nor near- infrared. In other words, identifying pure aluminium oxide, aluminium oxide with cBN and the tested metal disk in the range of 400-1700 by using hyperspectral imaging was impossible, because no peaks or depths belong to these materials have been observed.

ACKNOWLEDGMENT

This research was supported by "Mobilitas Plus ERA-NET", ETAG 18012, Durable ceramic composites with superhard particles for wear-resistant cutting tools, Tallinn University of Technology.

REFERENCES

- [1] C.-I. Chang, "Hyperspectral imaging: techniques for spectral detection and classification", Springer Science & Business Media LLC, 2003.
- [2] H. Grahn, P. Geladi, "Techniques and applications of hyperspectral image analysis", John Wiley & Sons, 2007.
- [3] "Environmental decision making", Science and Technology, 2019. [Online]. Available: <http://environ.andrew.cmu.edu/m3/s2/02sun.shtml> .
- [4] D. Smith, "Calculating the Emission Spectra from Common Light Sources", COMSOL Multiphysics, 2016. [Online]. Available: <https://www.comsol.com/blogs/calculating-the-emission-spectra-from-common-light-sources/>
- [5] "ATR-FT-IR spectra of Aluminium oxide (Al₂O₃)", Database of ATR-FT-IR spectra of various materials, 2019. [Online]. Available: http://lisa.chem.ut.ee/IR_spectra/paint/fillers/aluminium-oxide/
- [6] "RESONON hyperspectral imaging cameras", Laser 2000, 2018. [Online]. Available: https://www.laser2000.fr/en/hyperspectral-imaging/48082-pika-nir-near-infrared-hyperspectral-imaging-camera.html#download_anker_link
- [7] M.-M. Attallah, "Boron in Materials Technology", biotsavart.tripod.com, 2019. [Online]. Available: <http://biotsavart.tripod.com/bci.htm>
- [8] M. Chaplin, "Water absorption spectrum", 2019. [Online]. Available: http://www1.lsbu.ac.uk/water/water_vibrational_spectrum.html

Curriculum Vitae

Personal data

Name: Dhanushka Chamara Liyanage
Date of birth: 21.01.1983
Place of birth: Colombo (Sri Lanka)
Citizenship: Sri Lanka

Contact data

E-mail: dhanushka.liyanage@taltech.ee / liyanagedc@gmail.com

Education

2017–2021 Tallinn University of Technology, Estonia – PhD
2014–2016 Tallinn University of Technology, Estonia – MSc (Mechatronics)
(*Cum Laude*)
2003–2008 University of Moratuwa, Sri Lanka – BSc Hons (Mechanical
Engineering)
1993–2002 Mahanama College, Colombo, Sri Lanka, Secondary education

Language competence

English Fluent
Sinhalese Native

Professional employment

2014–Present Engineer (Projects Eng., Supervisor of master theses and
teacher for MSc and BSc degree courses), Tallinn University of
Technology, Estonia
2010–2014 Mechanical Design Engineer, Toshiba Tec Singapore Pte Ltd,
Singapore
2008–2010 Mechanical Engineer, Imperial Tea Exports (Pvt) Ltd, Sri Lanka

Elulookirjeldus

Isikuandmed

Nimi: Dhanushka Chamara Liyanage
Sünniaeg: 21.01.1983
Sünnikoht: Colombo (Sri Lanka)
Kodakondsus: Sri Lanka

Kontaktandmed

E-post: dhanushka.liyanage@taltech.ee / liyanagedc@gmail.com

Hariduskäik

2017–2021 Tallinna Tehnikaülikool, Eesti – doktorantuur, PhD
2014–2016 Tallinna Tehnikaülikool, Eesti – magistrantuur, MSc
(mehhatroonika) (*Cum Laude*)
2003–2008 Moratuwa Ülikool, Sri Lanka – bakalaureuseõpe, BSc
(Mehaanika)
1993–2002 Mahanama kolleegium, Colombo, Sri Lanka, keskharidus

Keelteoskus

Inglise keel Kõrgtase
Sinhala keel Emakeel

Teenistuskäik

2014– Insener (projekti insener, magistritööde juhendaja ja BSc ning MSc kursuste õppejõud), Tallinna Tehnikaülikool, Eesti
2010–2014 Mehaanika tootearenduse insener, Toshiba Tec Singapore Pte Ltd, Singapur
2008–2010 Mehaanika insener, Imperial Tea Exports (Pvt) Ltd, Sri Lanka

ISSN 2585-6901 (PDF)
ISBN 978-9949-83-766-3 (PDF)