

SUMMARY

Conclusions

The field of computer vision has been found effective in multiple sectors: manufacturing, autonomous systems and the medical. One of the branches of computer vision is image segmentation. Today popular segmentation methods are tested on datasets containing structured scenes. These datasets typically contain scenes that have good contrast, sharp features and good lighting. However, datasets that contain unstructured scenes is equally important. These can contain features that have poor contrast, fuzzy features and low lighting conditions. Segmentation methods that target unstructured scenes have been proposed, however these methods typically are supervised. Some unsupervised methods for unstructured segmentation have been proposed, but these require temporal data which is not always feasible for static images.

In this thesis, various semantic segmentation methods that can segment a dataset of unstructured scenes in an unsupervised manner were explored and compared. To aid with comparison an evaluator tool built in the Python is created. Upon conducting the Literature review it was found that semantic segmentation algorithms can be broadly classified into those based on clustering and neural networks. In the case of segmentation algorithms based on clustering, three were chosen. These include K-means, mean shift and graph cut. The neural network segmentation methods use concepts of deep learning. In the case of semantic segmentation methods employing CNNs became popular. The covered neural network methods that are capable of unsupervised segmentation where Segment Anything and Differentiable Feature Clustering. For the evaluation of segmentation algorithms an evaluation metric is used. This metric compares the segmentation result with a ground mask and gives a score. From the literature popular evaluation metrics include Intersection over Union, F1 measure, Panoptic Quality and the Region Over and Under segmentation measure.

The covered segmentation methods were implemented and evaluated by a purpose built tool in the Python programming language. This tool exposes the Evaluator class to the user which can be instantiated with a path to a COCO JSON dataset and the segmentation method to be evaluated. The dataset of unstructured scenes was provided by the university and then was hand segmented.

Based purely on the segmentation performance as reported by the evaluation tool, K-means and Segment Anything performed the best in their categories. However, there have been cases where the performance of these methods may not be properly accounted. For example Segment Anything tends to over segment the road region while Differentiable Feature Clustering performs better in this area. In the case of K-means, it tends to produce more segmentations that are noisy in comparison to graph cut or mean shift.

Future works

As previously stated, our evaluation tool may fail to properly account in cases where over segmentation or many segmentations are present. Future works could explore how to better account for cases where over segmentation or under segmentation are penalized.

The evaluation tool currently operates under a single thread. Thus, comparisons of large datasets can take a significant amount of time. Further, works can seek to implement multithreading or explore other performance optimization methods.

When saving an evaluation result the private data is saved to disk. During the tests, some of these results approached a gigabyte. Thus, future works could explore how to better compress this data or optimize the saving of only necessary data.

In this thesis only some unsupervised segmentation methods were covered. The field of computer vision is always evolving thus this topic may be explored again with the newest available methods. Future work may also cover how supervised or other segmentation methods can be applied for the segmentation of unstructured scenes.

KOKKUVÕTE

Arvutinägemise valdkond on osutunud tõhusaks mitmes sektoris: Tootmises, autonoomsetes süsteemides ja meditsiinis. Üks arvutinägemise harudest on pildi segmenteerimine. Tänapäeval katsetatakse populaarseid segmenteerimismeetodeid struktureeritud stseene sisaldavate andmekogumite peal. Need andmekogumid sisaldavad tavaliselt stseene, millel on hea kontrastsus, teravad tunnused ja hea valgustus. Kuid sama oluline on ka struktureerimata stseene sisaldavate andmekogumite kasutamine. Need võivad sisaldada halva kontrasti, hägusaid tunnuseid ja halba valgustust. tingimused. Välja on pakutud segmenteerimismeetodid, mis on suunatud struktureerimata stseenidele, kuid need meetodid on tavaliselt juhitud. Struktureerimata stseenide segmenteerimiseks on välja pakutud mõned järelevalveta meetodid, kuid need nõuavad ajalisi andmeid, mis ei ole staatiliste piltide puhul alati teostatav.

Käesolevas töös käsitletakse erinevaid semantilisi segmenteerimismeetodeid, mida saab segmenteerida struktureerimata stseenide andmestikku järelevalveta, uuriti ja võrreldi. Võrdluse abistamiseks on loodud Pythonis ehitatud hindamisvahend. Kirjanduse läbivaatamisel leiti, et semantilisi segmenteerimisalgoritme võib laias laastus liigitada järgmiselt. klastrimisel ja närvivõrkudel põhinevateks. Klasterdamisel põhinevate segmenteerimisalgoritmide puhul valiti välja kolm. Nende hulka kuuluvad K-means, keskmine nihe ja graafi lõikamine. Neuronivõrkude segmenteerimismeetodid kasutavad süvaõppe kontseptsioone. Semantilise segmenteerimise puhul muutusid populaarseks CNN-i kasutavad meetodid. Kaetud närvivõrgumeetodid, mis on võimelised järelevalveta segmenteerimiseks kus Segment Anything ja Differentiable Feature Clustering.

Segmenteerimisalgoritmide hindamiseks kasutatakse hindamismeetrit. Selle mõõdiku abil võrreldakse segmenteerimistulemust põhimaskiga ja antakse hinne. Kirjanduses on populaarsed hindamismõõdikud Intersection over Union, F1-mõõdikud, Panoptic Quality ja Region Over and Under segmentatsioonimõõdikud. Kaetud segmenteerimismeetodid, mille puhul rakendati ja hinnati Pythoni programmeerimiskeeles spetsiaalselt loodud tööriistaga. See tööriist eksponeerib kasutajale Evaluator klassi, mida saab instantsseerida käsuga COCO JSON-andmekogumile ja hinnatavale segmenteerimismeetodile. Struktureerimata stseenide andmekogumi esitas ülikool ja seejärel kasutati segmenteeriti käsitsi.

Põhineb üksnes hindamisvahendi poolt teatatud segmenteerimise tulemuslikkusel, K-means ja Segment Anything saavutasid oma kategooriates parimad tulemused. Siiski on esinenud juhtumeid, kus nende tulemuslikkus on meetodid ei pruugi olla nõuetekohaselt arvestatud. Näiteks Segment Anything kipub maanteepiirkonda liigselt segmenteerima, samas kui Differentiable Feature Clustering töötab selles piirkonnas paremini. K-means'i puhul kipub see tekitama rohkem segmenteerimisi, mis on mürarikkad. võrreldes graafiku lõikamise või keskmise nihkega.

Nagu eelnevalt öeldud, ei pruugi meie hindamisvahend nõuetekohaselt arvesse võtta juhtudel, kui esineb liigset segmenteerimist või palju segmenteerimisi. Tulevaste töödega võiks uurida, kuidas paremini arvesse võtta juhtumeid, kus üle- või alamsegmentatsiooni karistatakse. Hindamisvahend töötab praegu ühe niidi all.

Seega võib suurte andmekogumite võrdlemine võtta märkimisväärselt palju aega. Edasistes töödes võib püüda rakendada mitmikkeelset töötlust või uurida muid jõudluse optimeerimise meetodeid. Hindamistulemuse salvestamisel salvestatakse privaatsed andmed kettale. Katsete käigus lähenesid mõned neist tulemustest gigabaidile. Seega võiks tulevastes töödes uurida, kuidas neid andmeid paremini kokku suruda. või optimeerida ainult vajalike andmete salvestamist. Käesolevas väitekirjas käsitleti ainult mõningaid mittekontrollitud segmenteerimismeetodeid. Arvutinägemise valdkond areneb pidevalt, seega võib see teema olla uuesti uurida uusimate olemasolevate meetoditega. Tulevases töös võib käsitleda ka seda, kuidas jälgitavad või muud segmenteerimismeetodid saab rakendada struktureerimata stseenide segmenteerimiseks.