

## **Abstract**

Lung cancer disease has turned into the first and second reasons (for men and women respectively) for death in the world. Many research efforts and investigations are being developed to understand and treat this illness. Medical crew and radiologists are working for to detect this on this illness all around the world. However, methods based on traditional chest X-rays do not provide sufficient reliability to detect lung tumours at an early stage.

In this context, numerous computer aided detection systems (CAD) and image databases are being created to improve the examination of lung recognition systems in the early phases of lung cancer. CAD systems use computer tomography images which are stored in image databases. CT scans are obtained from a patient's lung area via special X-ray scanners and stored in a database. CAD systems examine these databases and try to identify the nodule spots on the images. However, these systems are still not trustful enough and generate (to various extents) so-called false positive and false negatives, which impairs the work of the medical personnel.

In relation to an ongoing research project at T.J. Seebeck Department of Electronics, this MSc thesis concentrates on methods used for increasing lung cancer detection accuracy for CAD systems. The main objective is to figure out how to reduce false positives and false negatives during the nodule detection phase. To meet this objective, we investigate methods for feature extraction (2D shape and size features) and nodule classification (3D features such as intensity, shape, size and texture).

These methods are implemented in a framework developed in the MATLAB environment and applied to the LIDC/IDRI CT scans dataset. The experimental results show how features are extracted from the CT scans and the corresponding values are plotted as graphs. The results also show that applying threshold values can increase the accuracy of the CAD systems. All in all, this work may contribute to the early detection of lung cancer and can help reducing the workload of radiologists.

Keywords: Computer Aided Detection (CAD), Computer Tomography (CT), Lung cancer, Nodule, False Positives, False Negatives.

The thesis is in English and contains 43 pages of text, 5 chapters, 25 figures and 3 tables.

## **Annotatsioon**

### **KOPSU NOODULITE AVASTAMINE JA KLASSIFITSEERIMINE KOMPUUTERTOMOGRAAFIA KUJUTISTELT**

Tänapäeval on kopsuvähk kujunenud meeste surmapõhjuseks number üks ja naiste surmapõhjuseks number kaks. Palju jõupingutusi ja uurimistöid on tehtud, et paremini mõista ja ravida nimetatud haigust. Meditsiinipersonal ja radioloogid tegelevad selle haiguse avastamisega ülemaailmselt. Kuid traditsionaalsed röntgenmeetodid ei ole veel piisavalt usaldusväärsed kopsuvähi varaseks tuvastamiseks.

Sellega seoses on loodud mitmeid arvutipõhiseid tuvastamis-süsteeme (CAD) ja kujutiste andmebaase, et parandada kopsuvähi avastamist varases staadiumis. CAD-süsteemid kasutavad andmebaasidesse salvestatud tomograafiakujutisi. Kompuutertomograafia käigus tehakse patsientide kopsupiirkonnast röntgenskaneringud ja kujutised salvestatakse andmebaasidesse. CAD-süsteemid analüüsivad neid andmekogumeid, et leida kujutistelt võimalikke kasvajate olemasolu. Kuid senised süsteemid pole piisavalt usaldusväärsed ja annavad valepositiivseid ja valenegatiivseid tulemusi, mis raskendavad personali tööd.

T.J. Seebecki elektroonikainstituudis läbiviidava uurimistööga seotud antud magistr töö keskendub meetoditele, mis aitavad parandada CAD-süsteemide kopsuvähi tuvastamise täpsust. Peamise eesmärk on vähendada kasvaja tuvastamisel valepositiivsete ja valenegatiivsete tulemuste hulka. Selle saavutamiseks uuritakse tunnuste eraldamise (2D kuju- ja suurusomadused) ja kasvaja klassifitseerimise (3D tunnused; näit. intensiivsus, kuju, suurus ja tekstuur) meetodeid.

Need meetodid on teostatud MATLAB keskkonnas ja rakendatakse kompuutertomograafia andmekogude LIDC/IDRI analüüsimiseks. Eksperimentaalsed tulemused näitavad, kuidas kompuutertomograafia kujutistest on eraldatud nimetatud tunnused, saadud tulemused on esitatud graafikutel. Tulemused näitavad samuti, et otsustusläävede kasutuselevõtt suurendab CAD-süsteemide täpsust. Kokkuvõttes, antud töö tulemused parandavad kopsuvähi varajast tuvastamist ja vähendavad radioloogide töökoormust. Võtmesõnad: arvutipõhine tuvastamine (CAD), kompuutertomograafia (CT), kopsuvähk, kasvaja, valepositiivne, valenegatiivne. See magistr töö on inglise keeles ja sisaldab 43 lehekülge teksti, 5 peatükki, 25 joonist ja 3 tabelit.