Evaluation of VGG & ResNet Very Deep Convolutional Neural Networks for Detecting Lung Cancer in CT Scans

W.P.H.D. Perera

168253V

This dissertation submitted in partial fulfillment of the requirements for the Degree of MSc in Computer Science specializing in Data Science

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

May 2018

DECLARATION

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another per-son except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature:…………… Date:……………….

Name: W.P.H.D.Perera

The supervisor/s should certify the thesis/dissertation with the following declaration.

I certify that the declaration above by the candidate is true to the best of my knowledge and that this report is acceptable for evaluation for the CS5999 PG Diplo-ma Project.

Signature of the supervisor: …………………………. Date: ………………..

Name: Dr. Charith Chitraranjan

ABSTRACT

Lung cancer is the second most common destructive cancer in the world. It is important to detect Lung cancer at its earliest possible time because this dreadful illness spreads in a rapid pace weakening and killing the entire body of a human. The lung cancer identification process is not easy as its symptoms are visible to outside mostly at its final stage. Lung cancer nodules are detected by radiologists through CT (Computed tomography) scans, but there is a high probability to fail to spot where the actual lung cancer nodule is because, the lung lesions are low in contrast. Therefore, there should be a Computer Aided Diagnosis (CAD) system to assist radiologists in identifying lung nodules efficiently, accurately so the results given by the CAD systems can be taken as a second opinion to detect lung nodules for radiologists. Accurate CAD systems can improve the quality and productivity of radiologists' image interpretation. There are many research subjects ongoing in medical imaging and diagnostic radiology. But it is needed to continuously improve the accuracy and the consistency of radiological diagnoses because still there are high false positive rates associated with CAD system results.

Current CAD systems have been developed using two main different approaches. First is the conventional old framework which detects lung cancer nodules using manual feature extraction and conventional image preprocessing approach. New approach is the Deep Neural Network architecture which automatically and directly uncovers features from the training data. In this approach the three steps, feature extraction, selection and supervised classification have been realized within the optimization of the same deep architecture.

This research evaluates two existing very deep learning architectures, Resnet-50 and VGG19 for learning high-level image representation to achieve high classification accuracy with low variance in medical image binary classification tasks. The classification accuracy was performed with two different datasets, NDSB and LUNA16+LIDC. For NDSB dataset the Restnet-50 model outperformed VGG19 model by giving the accuracy, sensitivity and specificity as 68%, 73%, 65% respectively. And for LUNA16+LIDC dataset 80.3%, 79.8%, 80.6% results were obtained for accuracy sensitivity and specificity out performing again the VGG19 network architecture.

ACKNOWLEDGEMENT

I would like to express profound gratitude to my advisor, Dr. Charith Chitraranjan, for his invaluable support by providing relevant knowledge, materials, advice, supervision and useful suggestions throughout this research work. His expertise and continuous guidance enabled me to complete my work successfully.

Last but not least, I present my appreciation to my family and my friends who were behind me, encouraging and directing me towards the success of my project.

Table of Content

LIST OF FIGURES

LIST OF TABLES