Convolutional Neural Network-based medical checkup system for Pigmented Skin Lesions Classification.

School of Computing, Engineering and Mathematics
Coventry University

Bsc Computer Science
Project Supervisor: Dr.David Croft
Submission Date: 21/04/2020

Author Name: Vinayak Sareen
Student Id Number: 7651331

Statement of originality

Declaration of originality

I Declare that This project is all my own work and has not been copied in part or in whole from any other source except where duly acknowledged. As such, all use of previously published work (from books, journals, magazines, internet etc.) has been acknowledged by citation within the main report to an item in the References or Bibliography lists. I also agree that an electronic copy of this project may be stored and used for the purposes of plagiarism prevention and detection.

Statement of copyright

I acknowledge that the copyright of this project report, and any product developed as part of the project, belong to Coventry University. Support, including funding, is available to commercialise products and services developed by staff and students. Any revenue that is generated is split with the inventor/s of the product or service. For further information please see www.coventry.ac.uk/ipr or contact ipr@coventry.ac.uk.

Statement of ethical engagement

I declare that a proposal for this project has been submitted to the Coventry University ethics monitoring website (https://ethics.coventry.ac.uk/) and that the application number is listed below

Sign: **Vinayak Sareen**. Date: **07/03/20**

Project Details

FirstName:	Vinayak
LastName:	Sareen
Student ID Number	7651331
Ethics Application Number	P101878
1st Supervisor Name	Dr.David Croft
2nd Supervisor Name	Unknown

Abstract

Abstract should be a succinct and self-standing summary of the basis, context and achievements of the project. Minimally an abstract does three things: (1) It states the problem that you set out to solve, (2) It describes your solution and method, (3) It states a conclusion about the success of the solution. Be straightforward and factual and avoid vague statements, confusing details and "hype". Do not be tempted to use acronyms or jargon to keep within the half-page limit. Consider that search engines, librarians and non-computer scientists wishing to classify your Report rely on the abstract. You may if you wish provide a short list of keywords (2-6 is reasonable) at the end of the abstract.

Contents

Acknowledgements

I want to express my gratitude to my supervisor Dr David Croft for guiding me thought the investigations. The support of all medical professionals who contribute by providing data to the research is sincerely acknowledged. At last, I would thank my parents, who supported me finically and psychologically throught this journey of computer science.

Chapter 1

Introduction

1.1 Introduction to Problem

Skin cancer is categorized into two types: melanoma and non-melanoma skin cancer. Melanoma, is the most dangerous kind of skin cancer accounted for an estimated 16,000 deaths each year from 2014 to 2016 in the United Kingdom (Cancer Research UK, 2020). The melanoma tumour caused by melanocytes can result in uncontrolled and abnormal growth which can spread in the human body (Korotkov & Garcia 2012). Detection and classification of unknown pigmented skin lesions can result in early diagnosis of the medical problem. The research data provided by cancer research organisation in 2017 has shown that melanoma was the 20th most common disorder with new incidents of 81,00 and 83,00 in males and females respectively in the United Kingdom (Korotkov & Garcia 2012). Dermoscopy is a non-invasive method of examining the pigmented skin, which includes microscopic imaging of the surface structure of pigmented skin lesions (Korotkov & Garcia 2012). Early diagnosis of pigmented skin lesions is crucial to classify skin disorders to decrease mortality concerning particular skin disorders. Dermoscopy improves the detection of melanoma compared to detection of disease with naked eyes by analysing the pigmented skin lesion. Previous studies have shown that such tumours can result in higher chances of better treatment and cure of disease by removing the tumour (Celebi, Kingravi, Uddin, Iyatomi, Aslandogan, Stoecker & Moss 2007). The current diagnosis method of detection involves using ABCD rule which considers the Asymmetry, Border irregularity, Colour irregularities, Darmascopic structures respectively of common pigmented skin lesions (Loescher, Janda, Soyer, Shea & Curiel-Lewandrowski 2013). People working in busy work environments or less mobility can be victims of belated and slow diagnosis of such dangerous skin cancers. The automated analysis of pigmented skin lesions using artificial neural networks can be beneficial in optical analysis of microscopic images of pigmented skin lesions. The primary targeted audience who benefits from the outcome is the people who are working in busy work environments or people with less mobility

are best to use cases which can use such an automated system. Booking a prior appointment with medical professionals based on the urgency of detected medical problems can result in the immediate treatment of patients with more critical conditions. The people with less mobility such as older audiences or people with special needs can detect pigmented skin lesions through online systems in an inconvenient manner. Medical institutions can use such technologies to automate the process of pre-health checkups and overcome the problem of shortage of staff members in case of emergency. Such automated systems can also result in faster diagnosis of medical problems compared to a manual analysis by a clinician. Furthermore, manufacturing companies which supply the microscopic medical instruments can also use such intelligent models with their products to provide value to customers and medical institutions.

1.2 Objective of Research

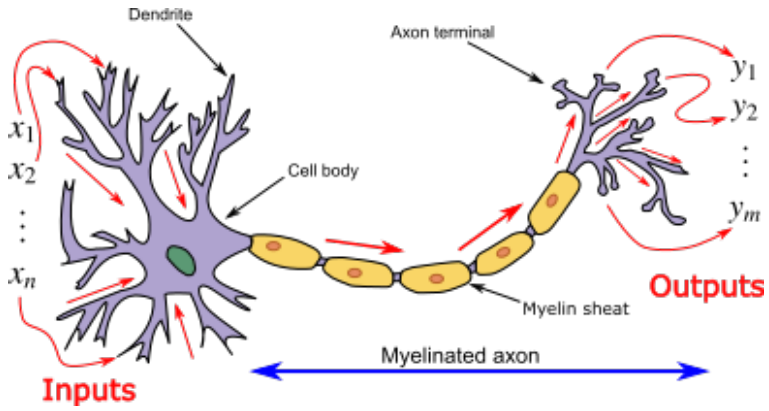
The research concentrates on developing a type of artificial neural network called a convolutional network to perform automated optimal analysis to identify the class of pigmented skin lesion. The research focuses on providing the quantitative analysis on comparing the results predicted by the automated intelligent machine are compared with medical professionals to identify the classes of pigmented skin lesions. The study employs different experiments by applying different model architectures and analysing accurate hyper-parameters for optimal performance. The research concentrate on analysing limited skin tumours such as melanoma, benign keratosis, melanocytic nevi and basal cell carcinoma. The investigation employs publicly available HAM10,000 dataset (Tschandl 2018). The extensive collection of 10,000 images of labelled data units were collected from a diverse population of subjects over twenty years. The outcome of the research project will to be analyse the effectiveness of the automated pigmented lesion system. Furthermore, the intelligent model will be deployed on web based system to provide interface to use the system to analyse by general audience.

1.3 Perceptron Model

1.3.1 Inspiration from biological Structure of Neurons

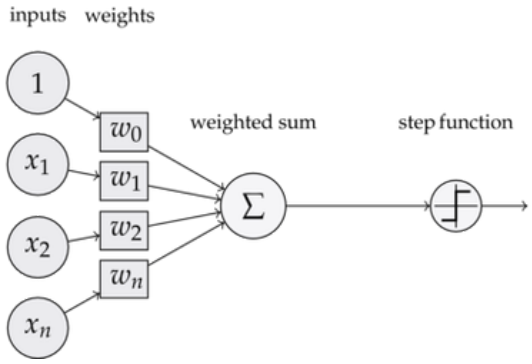
The human brain are componsed of millions specialised cell "neurons" which are interconnected to each other which carry electrical and chemical signals from neuron to another to function There are an estimated 500 trillion connections between neurons in the human nervous system which helps communicate signals (Patterson & Gibson 2017). The inspiration of artifical neural networks are taken from functioning of human brain and learn to recoganise patterns and relationship in the data (Agatonovic-Kustrin & Beresford 2000).

1.3.2 Neural Structure



The figure above represents the biological structure of neurons that helps in the communication of electric signals in the nervous system to learn and process information. Biological structure of neurons includes three major parts dendrites which are responsible for accepting the electrical and chemical signals to the neuron. Furthermore, the neuron contains the nucleus which is accountable for the processing input information with the neurons. At last, the processed information is passed to another neuron which is interconnected to neuron in the human nervous system through axon terminals (Agatonovic-Kustrin & Beresford 2000).

1.3.3 Perceptron Model



Preceptron model was proposed by Frank Rosenblatt in 1943 to design the model to mimic the human brain (Kussul, Baidyk, Kasatkina & Lukovich 2001). Preceptron model or single layered feed forwards network networks takes the vectors of inputs and multiply with a randomly initialised weights and add random bias to network and process the information by providing data to the activation function to process the information (Agatonovic-Kustrin & Beresford 2000). Figure above represents the simple preceptron model which consists inputs x and weights w for which weighted sum of muliplication result of inputs and weights will be passed to step activation function.

1.3.4 Mathematical Representation

The equation below represents the preceptron model mentioned above in Mathematical notations.

$$y = \Big[\sigma(\sum_{k=0}^{n} x_k.w_k + b_k)\Big] \tag{1.1}$$

In, euation above $\sigma = Activation function$, x and w represents the inputs provided and weights vectors in the network respectively. Furthermore symbol b represents the bias in the equation, during the optimisation of the network weights and bias are adjusted accordingly to improve the predection. There are various activation functions σ such as step function sigmoid, relu, leaky relu and others that can be applied based on the requirments of model prediction.

1.4 Multi Layered Feed Forward Neural Network

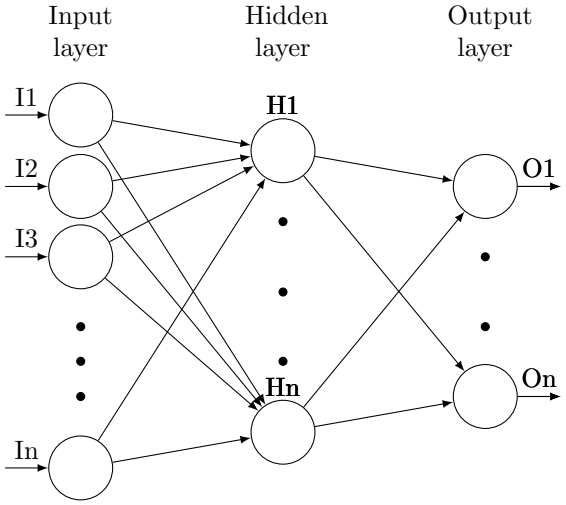


Figure 1.1: Multi Layered Neural Network diagram.

The preceptron model acts the base for all the functioning of modern multi-layered neural networks. The output of the single neuron described in the the above preceptron model is taken as the input in next layer of network. The above figure(1.1) shows an example of multi-layered neural where the first layer is known as input layer and intermediate layers are called hidden layer and last layer is known as output layer. The single neuron can performe limited computation but the computation power increases with inter-connected neurons in the network (Agatonovic-Kustrin & Beresford 2000) where at each layer information is processed through the appropriate activation function. Artifical neural networks are designed to learn the patterns and relationships in the data and requires enough amount of data to train and predict accurate outputs (Agatonovic-Kustrin & Beresford 2000). The training data is passed to the neural network to recoganise the patterns in data and in each iteration or epoch the model predection gets improved with the optimisation algorithm which propagtes through the network to update the weights and bias to increase the accuracy (Agatonovic-Kustrin & Beresford 2000). The accuracy of the neural networks are determined through various hyper-parameters such as learning rate, number of hidden layers in the model and number of epochs for which model is trained. The fundamental rule for learning in artifical model

1.4.1 Cost Function and Backpropagation Algorithm

Chapter 2

Literature Review

2.0.1 Convential Diagnosis Methods

The most common conventional diagnosis method of detection involves using ABCD rule which considers the Asymmetry, Border irregularity, Colour irregularities, Darmascopic structures respectively of common pigmented skin Lesions (Loescher, Janda, Soyer, Shea & Curiel-Lewandrowski 2013). The above method of analysis is performed on dermoscopic images of pigmented skin lesions. Dermatoscopy is non-invasive microscopic imaging of pigmented skin lesions which provides clear imaging to perform proper analysis on pigmented skin lesions (Loescher, Janda, Soyer, Shea & Curiel-Lewandrowski 2013). Furthermore, the result of dermatoscopic images is examined by dermatologists to classify the pigmented skin lesion.

2.0.2 Support Vector Based Machine

Thompson Felsia and Jeyakumar proposed research in 2017 on support vector machine based classifier to detect multi-lesions skin cancer by analysing pigmented skin lesions with an accuracy of 86.37 percent. The proposed investigation with SVM based classifier has performed image segmentation using SRM (support region merging) algorithm. Furthermore, it employs SURF (speed up robust features) to find the region of interest for feature extraction to get optimal classification performance based on vector-based technique (Thompson & Jeyakumar 2017). However, the research does not include image augmentation which generalises the predictions accurate to test in real-world environment. The research papers mention that support vector machine for automated classification of pigmented skin lesions is sensitive to the artefacts and can potentially increase the false positives which mean that predicted result for analysis was wrong positive prediction instead of an actual negative result. The investigation will perform image augmentation to generate random samples of images with different rotation angle and flipped images will be used to train and test the model to generalise the overall performance.

2.0.3 Border Detection Based System

Rahil Garnavi and his other co-researchers purposed research based on a state of the art border detection method combined with the colour space analysis and clustering-based histogram hybrid thresholding to classify pigmented skin lesions. The research was primarily focused on the research was to develop the hair removal mechanism to perform colour channels transformation. Furthermore, for all the image channels the noise reduction and clustering-based histogram thresholding were performed for optimal border detection. The predicted outcomes of novel broder detection system were compared with the borders detected by the actual dermatologists on a sample of dermoscopic pigmented skin lesions to understand the reliability of the system (Garnavi, Aldeen, Celebi, Varigos & Finch 2011).However, the system was only tested on a data sample set of 30 dermoscopic images and four sample sets of dermatologist hand-drawn images were used as ground truth to compare the results. The system was tested on overall 85 dermoscopic images. Border detection can be used to analyse the pigmented skin lesions but convolutional networks have the potential to find more data patterns in the images to minimise the cost function using the backpropagation algorithm. The current research will employe basic image segmentation based on the binary threshold algorithm as an experiment to help network detecting more accurate borders of pigmented skin lesions.

Deep Feature to classify Pigmented Skin lesions

In 2016, a research paper from Simon Fraser University's computer science and medical image analysis lab had researched using deep residual network architecture with ten labelled classes of pigmented skin lesions. The research was based on very deep convolutional network architecture with the accuracy of 85.8 percent in classifying five distinct classes and 81 percent in classifying 10 classes of pigmented skin lesions (Kawahara, BenTaieb & Hamarneh 2016). Although the performance of the overall convolutional network was accurate, the training and testing data were limited to 13,00 overall images of 10 distinct classes. However, In the current research project, the classes of labelled images will be five and around 9,000 overall images will be used during the investigation. Estimated 80 percent of data will be consumed for training the model, and the rest of the label images will be used as validation and testing datasets to evaluate the performance of the model. Research is also consuming such artificial neural network-based technologies to various areas of investigations.

Chapter 3

Methodology

This is a sample methodology review content. BLAH BLAH BLAH .... This chapter should describe what you did to answer your research question (or to support your thesis, if you think of it that way), and how you went about it (essentially your research design). You should describe your research design in sufficient detail that another researcher could recreate your work to check your results.

Chapter 4

Evaluations and Results

In this chapter, you should evaluate what you have done, and say what answer (to your research question) you have arrived at. It may be that in your method you describe some experiments, and this section records your results and analysis of those results. This is an important section – most students gain or lose marks in either their literature review or evaluation. The key to producing a convincing evaluation is to plan very early in the project what information or results you will need to write this section.

Chapter 5

Project Managment

This is a discussion based chapter BLAH BLAH BLAH .... Your first supervisor may have a very good idea of how well you tackled your project - however second supervisors may not have any idea. For this reason you need to include an account of the conduct of the project. What problems you encountered, how you overcame them, how diligently you worked, how you sought advice, how you responded to feedback. This chapter will be evidence driven – which is why you need to keep a log or diary of your project, maybe a project management timeline with milestones, keep evidence of each supervision meeting (signed off by your supervisor), Keep notes of supervisor feedback to your presentation and reflect on them in this chapter.

Chapter 6

Reflection

This is dummy text for reflection

Chapter 7

Conclusion

conclusion can be drawn from these examples to be continued .....

Bibliography

Agatonovic-Kustrin, S. & Beresford, R. (2000), 'Basic concepts of artificial neural network (ann) modeling and its application in pharmaceutical research', *Journal of Pharmaceutical and Biomedical Analysis* **22**(5), 717 – 727.
URL: *http://www.sciencedirect.com/science/article/pii/S0731708599002721*

Celebi, M. E., Kingravi, H. A., Uddin, B., Iyatomi, H., Aslandogan, Y. A., Stoecker, W. V. & Moss, R. H. (2007), 'A methodological approach to the classification of dermoscopy images', *Computerized Medical Imaging and Graphics* **31**(6), 362 – 373.
URL: *http://www.sciencedirect.com/science/article/pii/S0895611107000146*

Garnavi, R., Aldeen, M., Celebi, M. E., Varigos, G. & Finch, S. (2011), 'Border detection in dermoscopy images using hybrid thresholding on optimized color channels', *Computerized Medical Imaging and Graphics* **35**(2), 105 – 115. Advances in Skin Cancer Image Analysis.
URL: *http://www.sciencedirect.com/science/article/pii/S0895611110000819*

Kawahara, J., BenTaieb, A. & Hamarneh, G. (2016), Deep features to classify skin lesions, *in* '2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)', pp. 1397–1400.

Korotkov, K. & Garcia, R. (2012), 'Computerized analysis of pigmented skin lesions: A review', *Artificial Intelligence in Medicine* **56**(2), 69 – 90.
URL: *http://www.sciencedirect.com/science/article/pii/S0933365712001108*

Kussul, E., Baidyk, T., Kasatkina, L. & Lukovich, V. (2001), Rosenblatt perceptrons for handwritten digit recognition, *in* 'IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No.01CH37222)', Vol. 2, pp. 1516–1520 vol.2.

Loescher, L. J., Janda, M., Soyer, H. P., Shea, K. & Curiel-Lewandrowski, C. (2013), 'Advances in skin cancer early detection and diagnosis', *Seminars in Oncology Nursing* **29**(3), 170 – 181. Skin Cancer.
URL: *http://www.sciencedirect.com/science/article/pii/S0749208113000338*

Patterson, J. & Gibson, A. (2017), *Deep learning: A practitioner's approach*, " O'Reilly Media, Inc.".

Thompson, F. & Jeyakumar, M. (2017), 'Vector based classification of dermoscopic images using surf', *International Journal of Applied Engineering Research* **12**(8), 1758–1764.

Tschandl, P. (2018), 'The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions'.
URL: *https://doi.org/10.7910/DVN/DBW86T*

Appendices

Participation Consent Form

Participant No.

INFORMED CONSENT FORM:

Convolutional Neural Network-based medical check-up system for Pigmented Skin Lesions Classification.

You are invited to take part in this research study for the purpose of collecting data on evaluating the reliability of automated skin check-up system for classification of common types of pigmented skin lesions.

Before you decide to take part, you must **read the accompanying Participant Information Sheet.**

Please do not hesitate to ask questions if anything is unclear or if you would like more information about any aspect of this research. It is important that you feel able to take the necessary time to decide whether or not you wish to take part.

If you are happy to participate, please confirm your consent by circling YES against each of the below statements and then signing and dating the form as participant.

1	**I confirm that I have read and understood the Participant Information Sheet for the above study and have had the opportunity to ask questions**	**YES**	**NO**
2	**I understand my participation is voluntary and that I am free to withdraw my data, without giving a reason, by contacting the lead researcher and the Research Support Office at any time until the date specified in the Participant Information Sheet**	**YES**	**NO**
3	**I have noted down my participant number (top left of this Consent Form) which may be required by the lead researcher if I wish to withdraw from the study**	**YES**	**NO**
4	**I understand that all the information I provide will be held securely and treated confidentially**	**YES**	**NO**
5	**I am happy for the information I provide to be used (anonymously) in academic papers and other formal research outputs**	**YES**	**NO**
6	**I am happy to answer questions asked in this questionnaire**	**YES**	**NO**
7	**I agree to take part in the above study**	**YES**	**NO**

Thank you for your participation in this study. Your help is very much appreciated.

Participant's Name	**Date**	**Signature**
Researcher	**Date**	**Signature**

Supervisor Meeting Logs