Appendix A – Front Cover Exemplar (for using University Card Covers with cutout)

MSc Applied AI and Data Science
2022
Itoro-abasi Iton

Artificial Intelligence System to improve melanoma detection

SOLENT UNIVERSITY

FACULTY OF BUSINESS LAW AND DIGITAL TECHNOLOGIES

MSc Applied AI and Data Science
Academic Year 2021-2022

Itoro-abasi Iton

Artificial Intelligence System to improve melanoma detection

Supervisor: Kenton Wheeler
September 2022

This report is submitted in partial fulfilment of the requirements of Solent University for the degree of MSc Artificial Intelligence and Data Science

Acknowledgements

I thank the Almighty God for the strength and guidance throughout the commencement of this dissertation. I am grateful to my family and friends for the constant encouragement and support, and my supervisor Kenton Wheeler for consistently showing his support. I appreciate you all.

Abstract

Skin cancer has become increasingly prevalent in several countries across the world. The steady rise of the disease in the UK has become a healthcare burden on both medical practitioners and patients. In order to reduce mortality rates, it is necessary to identify and diagnose the disease early and accurately before it becomes malignant. The main types of skin cancer are squamous cell carcinoma, basal cell carcinoma and melanoma. This dissertation is focused on the most severe type of skin cancer which is melanoma; this cancer begins in the melanin making cells melanocytes and is likely to spread to other body parts and tissues if not identified immediately (Skin Cancer Screening (PDQ®)–Patient Version, 2022). There are current diagnostic practices used to identify if a skin lesion is malignant or benign, these procedures are mostly based on visual perception from a specialist. However, this thesis would highlight the flaws in this procedure, it is unreliable and there are several risks that could be incurred from it. This project would also consider the use of smartphone applications in appropriately identifying skin lesions. A number of deep learning algorithms such as; CNN, SVM, DERM and ANN, would alleviate the problems associated with current melanoma detection practices and facilitate effective and early diagnosis of the disease thereby decreasing health and financial burdens on the patients.

Table of Contents

1. Introduction

The exponential advancement of Artificial Intelligence (AI) AI is progressively transforming societies and industries across the world. (Wang and Siau, 2019) This is due to recent progress in digitized data acquisition, machine learning and computing infrastructure. As a result, the application of AI is expanding into areas that were previously thought to be solely occupied by human experts. (Yu, Beam and Kohane, 2018). The advancement of AI in medical practice is demonstrated through techniques like medical imaging; where the interior or exterior of the human body are represented in visual form by using radiations, radio frequencies, sound waves etc for diagnosis. Imaging modalities such as; Computed Tomography (CT) scan, X-ray and Magnetic Resonance Imaging (MRI) are used to study interiors of the body. (Adhikari 2019) This is vital in ensuring that a patient is accurately diagnosed and provided the appropriate treatment plan for their diagnosis. As opposed to sole reliance on examinations conducted with the naked eye from a human expert.

This does not necessarily imply that human medical practitioners are going to be replaced by machines, however, it reinforces the healthcare sector by scaling its processes and establishing quality control systems. There are areas where the medical profession cannot be automated as medical experts are essential to show care, empathy, mutual understanding and support for the patients. (Artificial Intelligence in Medical Imaging 2019). However, AI aims to revolutionize this industry through 'development of personalized and automated diagnostics, new diagnostic data-based methods, systematic monitoring of potential diagnostic errors and supporting correct medical decisions'. (Artificial intelligence in medical imaging 2019)

The AI system utilizes an ensemble of techniques where algorithms learn iteratively from data. This learning process is defined as Machine Learning (ML). When an input data is assigned its label, ML algorithm tries to learn features associated with those labels and this process is known as supervised learning. There is another process known as unsupervised learning where an algorithm tries to learn from data with no labels. Most medical diagnostic problems use supervised learning as data in medical

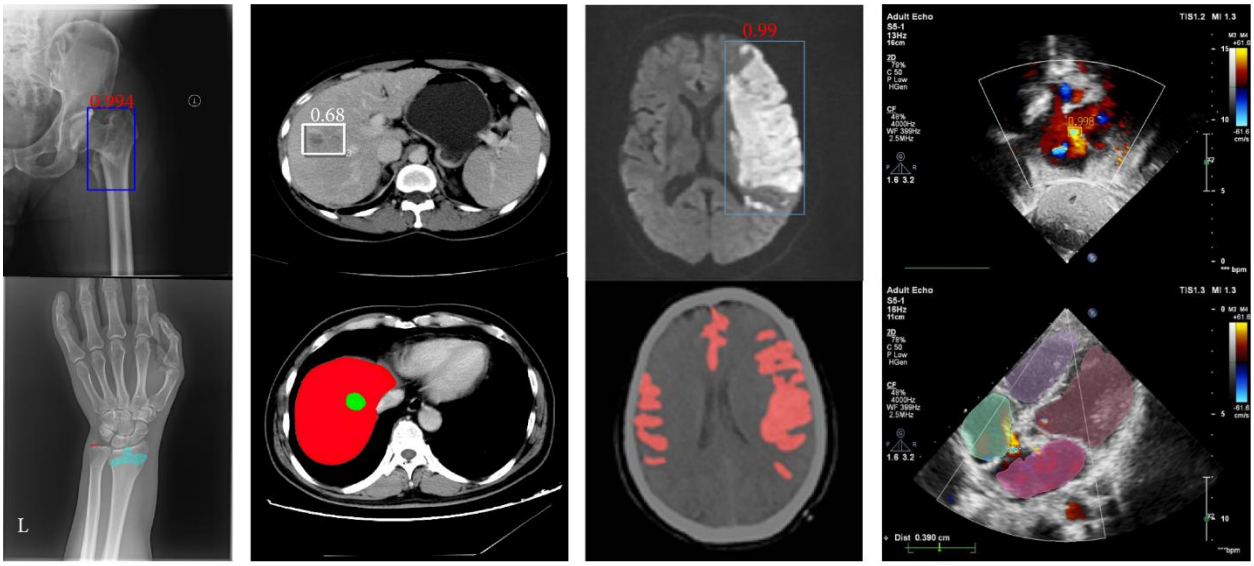
imaging are often labeled or annotated by human experts. After the labeled data has been trained, to verify the efficiency of a ML algorithm, the algorithm would be tested with unseen test data and would appropriately be regarded as 'efficient' if it can show similar performance and correctly identify test data. (Adhikari 2019)

1.1 Background

Melanoma is a serious form of skin cancer that occurs when melanin pigment forming cells known as melanocytes grow uncontrollably and this growth results in formation of malignant tumors. Melanoma is usually characterized by structure, features, size and colors, knowledge of these characteristics is vital to aid in early diagnosis of the disease. (The skin cancer foundation 2022) An experienced dermatologist would usually make diagnosis using an image observation technique called 'ABCD rule'. It is an abbreviation which states the melanoma characteristics to detect; A-Asymmetry, B- Border irregularity, C- Color distribution, D- Diameter length. (B. Sreedhar, M. Swamy B.E and M. S. Kumar 2020).

Over the years, skin cancer has become much more prevalent in the UK. The NHS accounts around 16,000 new cases of melanoma skin cancer are diagnosed each year making it the 5th most common cancer in the UK, with a record of more than 2,300 deaths each year. The age demographic being diagnosed are people under 50 years, which is unusually early in comparison with other cancer types. (NHS 2022) A factor thought to be a major contributor to the prevalence of melanoma is; increased exposure to ultraviolent rays from the sun or through tanning devices. other factors include; genetics and family history, having pale skin and having a large number of moles. (World Cancer Research Fund 2022)

melanoma diagnostics are usually carried out through visual inspection by a medical expert. The emergence of medical imaging has improved diagnostic accuracy through imaging modalities.



Bone X-ray Liver CT Brain MRI Cardiac ultrasound

Figure 1: the image above is displaying different examples of medical image modalities and their corresponding applications.

The imaging modality used to capture skin images through the use of polarized light is dermatoscopy and this is achieved by using an instrument named digital epiluminescence dearmatoscope. (Adhikari 2019). There is a risk of a medical practitioner making a misdiagnosis and collecting skin biopsy for lesions which turn out to not be skin cancer. Henceforth, there is a need for a reliable system which would be able to accurately detect the disease and decrease the mortality rate. (B. Sreedhar, M. Swamy B.E and M. S. Kumar 2020)

In this thesis a suitable machine learning model which would provide the best prediction accuracy for melanoma detection has been proposed. Additionally, the possibility of designing a layout to develop a prototype application for melanoma detection has been discussed. This application would be able to educate its users on how to carry out skin examinations, making it possible for them to detect skin changes and would be able to allocate them to their nearest dermatologists or General Practitioner for examination or surgery where necessary.

1.2 Research Question

The research question which would be considered in this report has been phrased according to the PICOT framework. PICOT is a mnemonic derived from the elements

of a clinical research question; (Formulating a PICOT Question | Duquesne University, 2022) P stands for Patient, I stands for Intervention, C stands for Comparison, O stands for Outcome and T stands for Time.

The research question is as follows; Is there an artificial intelligence system that can successfully detect images of melanoma as malignant or benign in individuals, with suitable prediction accuracy in comparison with other existing detection processes like clinical examinations.?

Using the PICOT framework;

- P –Individuals
- I – An AI system that can successfully detect images of melanoma as malignant or benign.
- C – In comparison with regular clinical examinations
- O – Suitable prediction accuracy
- T – Before metastasis occurs

1.3 Aims and Objectives

The aim of this project is to carry out research on existing relevant case studies, literature and data related to melanoma skin cancer and to display technical skills by identifying a suitable machine learning model which would aid in detecting melanoma. Objectives that would be achieved are;

I. To improve the quality of existing diagnostic systems by proposing a suitable machine learning algorithm that would aid in early detection of melanoma, thereby facilitating early treatment and improving survival rate.
II. To compare the proposed machine learning model with the results of other models with justifications.
III. To train the chosen model with a suitable dataset containing a sufficient number of case-study images of moles and lesions, to appropriately classify them as malignant and benign.
IV. To propose the possible development of a smartphone application that can be used by non-experts, which would be able to screen moles on all parts of the

user's body to determine whether they are malignant or benign and subsequently refer them to a nearby specialist for further examination or surgery where necessary.

2. Literature Review

In this chapter, there will be a critical review of journal articles, books and other relevant sources related to skin cancer, the risks associated with skin cancer screening and an overview of artificial intelligence systems that have been utilized to provide early detection of the disease would also be discussed.

2.2 Skin Cancer and clinical examinations

National Cancer Institute defines cancer as a term for diseases in which abnormal cells divide uncontrollably and invade nearby tissues. These cancer cells can also spread to other areas of the body through the blood and lymph systems and this process is known as metastasis.

Cell division occurs in every human body. This process allows the cells to grow and multiply to form new cells which replace old and damaged cells. However, when this orderly process breaks down and abnormal or damaged cells begin to grow and multiply when they should not, they form tumors which can either be malignant(cancerous) or benign(non-cancerous) tumors. (Skin Cancer Screening (PDQ®)–Patient Version, 2022)

There are several types of skin cancer, such as; basal cell carcinoma (BCC), squamous cell carcinoma (SCC) and melanoma. This thesis is going to focus on melanoma which is the skin cancer with the worst prognosis, unlike BCC and SCC which develop in the upper layer of the skin (the epidermis), melanoma spreads to other tissues and body parts. For an increased chance of survival, early diagnosis is necessary in order for appropriate and efficient treatment to be given. This would usually be conducted through screening tests.

Screening tests for early detection of skin cancer are normally carried out by the patient and the health care provider. A visual self-exam would be performed by the

patient while the health care provider would perform clinical examinations using dermoscopy. During a skin exam, the medical practitioner checks the skin for lesions, moles, birthmarks and other pigmented areas that look abnormal in color, size, shape or texture. If there is an area on the skin that looks abnormal, a biopsy is usually done. The medical practitioner would use a surgical excision to remove as much of the suspicious tissue as possible, then a pathologist inspects the tissue under a microscope to check for cancer cells (Skin Cancer Screening (PDQ®)–Patient Version, 2022).

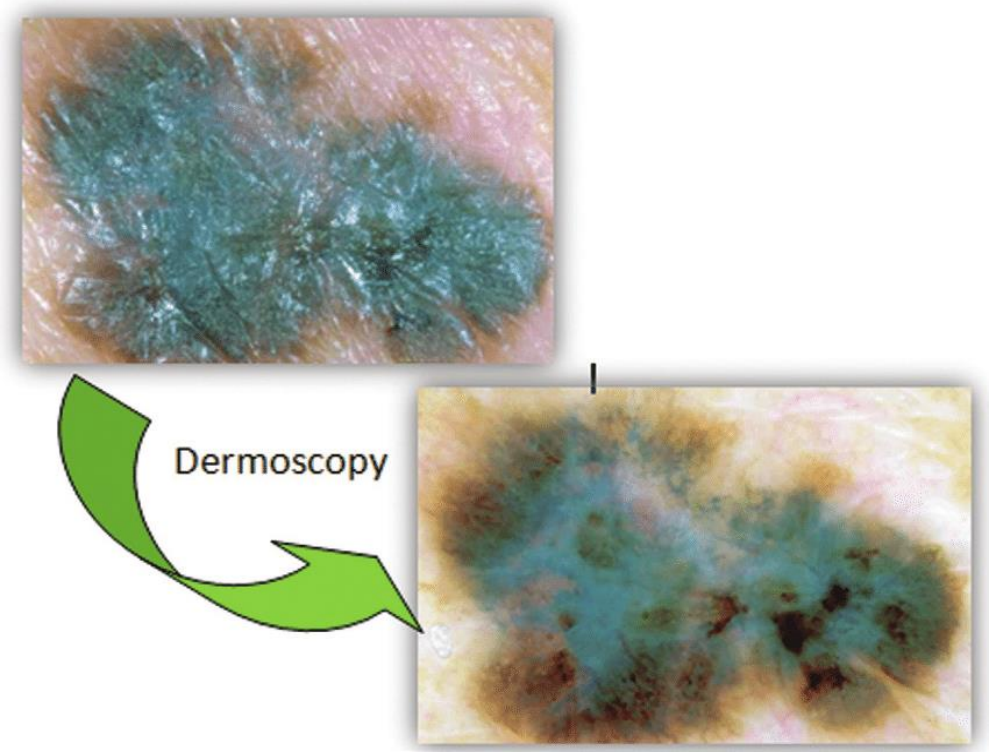


Figure 2: lesion observed with the naked eye in comparison to the dermoscopy examination

From the diagram above it can be seen that the features of the lesion are more visible using dermoscopy and the specialist can effectively visualize the irregularities of the lesion.

There are some risks associated with skin cancer screening;

- False-negative and false-positive test results are very likely to occur. In a false-negative test result, screening test results may appear normal even though

cancer is present, thereby, delaying the patient seeking urgent medical attention. For false-positive test results; screening test results may appear to be abnormal even though no cancer is present. This could affect a patient psychologically, causing them to have anxiety and have more tests like a biopsy conducted on them. (Skin Cancer Screening (PDQ®)–Patient Version, 2022)

- There is controversy regarding the screening of skin cancer that it promotes overdiagnosis and there is insufficient evidence showing the benefits of screening in asymptomatic adults. A study from Germany showed no reduction in mortality after 10 years of follow up after initiation of a skin cancer screening program. (Nicholas Brownstone, 2022)
- There is a risk of scarring and infection from a biopsy. As necessary as a skin biopsy is, there are several factors such as the cost of the procedure, the invasive nature of the procedure, the associated pain and risk of scarring, and the need for repeated samples in suspected lesions to be provided, that make it limited in clinical diagnosis (Das et al., 2021).

From the risks and limitations detailed above, of clinical examinations. There is a need to consider automated systems which can aid in early diagnosis of skin cancer.

2.3 Deep Learning

Deep learning is an intuitive process whose complexity of learning increases with the increase in the number of layers. Recently, deep learning has contributed greatly for skin lesion classification problems with high performance. However, potential groundbreaking research in medical diagnostics using deep learning is limited, this is due to the algorithm's dependency on large training datasets, as it requires millions of parameters and large amounts of labeled data to learn. It should be noted that when limited data is used to train a deep learning model, it uses a large amount of its resources for training which leads to overfitting issues. (Aakriti Adhikari).

In deep learning, each level learns to transform its input data into a slightly more abstract and composite representation.' The underlying architecture for deep learning was inspired by the structure of the human brain. Just as neurons form the

fundamental building blocks of the brain, deep learning architecture contains a computational unit called perceptron, that allows modeling of nonlinear functions. The perceptron receives a list of input signals and transforms them into output signals. It aims to understand data representation by stacking together a number of layers, each layer is responsible for understanding some part of the input (IBM Developer, 2022). Using the dataset, 'the first representational layer may abstract the pixels and encode edges; the second layer may compose and encode arrangements of edges, the third layer may encode anatomical patterns and the fourth layer may recognize that the image contains a disease.'(2019 Maryellen, phD)

There is a research study that was carried out to compare the accuracy of human expert readers versus machine learning algorithms in classifying pigmented skin lesions (Tschandl et al., 2019) the study was an open, web-based one were human readers were asked to diagnose dermatoscopic images selected randomly in 30-image batches from a test set of 1511 images. A comparison of the diagnosis from the human readers were made with those of 139 algorithms created by 77 machine learning labs. "The two main outcomes were the differences in the number of correct specific diagnoses per batch between all human readers and the top three algorithms, and between human experts and the top three algorithms." (Tschandl et al., 2019). The research was conducted in 2018 with 511 human readers from 63 countries who were all certified and professional medical practitioners. When comparing the results of all the human readers and all the machine learning algorithms, findings showed that the algorithms achieved a more correct diagnosis. This study proves that advanced machine learning classifiers should have a more prominent role in clinical practice. (Tschandl et al., 2019)

An Artificial Neuron Network (ANN) is the underlying architecture behind deep learning. The model is based on the structure and operations of biological neural networks. (Das et al., 2021) CNN is a variant of deep, feed-forward ANN used typically for analyzing visual imagery. It is made up of convolutional as well as pooling layers that allow the network to encode picture characteristics. (Das et al., 2021)

Further studies identify a model called DERM that was trained and tested using 7,102 dermoscopic images of both histologically confirmed melanoma and benign pigmented lesions. (Phillips, Greenhalgh, Marsden and Palamaras, 2019) and its accuracy compared to that of current clinical practices. The conclusion reached in this study showed that DERM has the potential to be used in primary care by providing dermatologist-grade recommendation on the likelihood of malignant melanoma.(Phillips, Greenhalgh, Marsden and Palamaras, 2019)

There are other technologies available for skin cancer detection, a prevalent one is smartphone applications(apps). There is a large number of smartphone apps that have been developed to aid in melanoma detection. Kassianos et al reviewed 39 applications that addressed skin cancer issues; 19 involved smartphone photography and 4 provided an estimate of the probability of malignancy. None of these applications had been assessed for diagnostic accuracy. There is concern about the possible harm to patients that a poorly designed or inaccurate diagnostic from these applications may cause. However, with suitable evaluation and appropriate development a smartphone app that can improve diagnostic accuracy of skin cancer would be essential in the medical field. (Phillips, Greenhalgh, Marsden and Palamaras, 2019) there are some current skin care apps that provide a diverse variety of functions. These apps can provide an information resource to educate and advise its users also, modern smartphones with high quality cameras assist greatly in monitoring skin conditions and conducting self-examinations (Freeman et al., 2022). Further studies on the risks of reliance on smartphone apps for skin cancer detection would be discussed in the next chapter.

3. Proposed Artefact and Societal Impact

The reason melanoma skin cancer is the chosen topic area for this project is because there is evidence from clinical studies indicating that the cancer cells can stay dormant for years and then reappear on the patient’s body even after surgery has been conducted, this is because the cancer has already spread to other tissues and

organs in the body evidencing, that there is no absolute cure for melanoma when the cancer cells are widespread or the disease has reached an advanced state. 'Clinical research has prompted the development of specific drugs such as vemurafenib and dabrafenib that are effective in decreasing tumor burden but only for a limited period of time.' (G. Merlino et al, 2016), however this is not sufficient enough in decreasing the mortality rate. Medical checkups of a suspicious mole are necessary however, it would be expensive for a person to regularly seek examinations from a specialist with a slight suspicion for every mole that appears on their body.

In this project the possible development of a smartphone application that would be able to detect if a mole or lesion is malignant or benign would be considered. Most of the features on the proposed app would be influenced from features contained in an already existing premium service skin cancer app called 'Miiskin' which uses high-res photography to take photos of large parts of the users body and compares individual moles over time to detect changes (Williamette Valley Cancer Institute 2022). The proposed app would be able to detect melanoma at a timely manner as early detection of the disease is especially paramount to aid in prevention and its accuracy results would be based on the effectiveness and reliability of the chosen machine learning model. During my research I discovered that most available mobile apps for detecting skin cancer, often offer premium paid service. The proposed app if implemented would offer a free service to promote awareness of the disease and cater to a wider rural demographic. It will also cut costs and resources that are usually needed in traditional methods of medical checkups.

There are undeniable benefits of developing a smartphone app that can effectively detect if a skin lesion is malignant or benign. However, it should be noted that previous research has been conducted that examined the accuracy of smartphone apps in detecting the risk of skin cancer in suspicious skin lesions. The criteria to determine eligibility of these studies was; assessing sources that evaluated algorithm-based smartphone apps, with the reference standards being; histological diagnosis or follow-up, and expert recommendation for further intervention or investigation. Two authors independently extracted data and the validity was assessed using QUADAS-2

(Quality Assessment of Diagnostic Accuracy Studies 2 tool) and reported estimates of sensitivity and specificity for each app. (Freeman et al., 2022)

Results from nine studies that evaluated six different smartphone apps provided differing results. Six verified results by using histology or follow-up and three verified results by using expert recommendations. The studies were cited as having poor methodological quality, with selective recruitment, high rates of unevaluable images and differential verification. Lesion selection and image acquisition were also performed by clinicians rather than smartphone users (Freeman et al., 2022).

The app SkinVision was evaluated in three different studies for the detection of malignant or premalignant lesions and its accuracy compared to expert recommendations was poor.

In order to successfully develop an algorithm-based smartphone app to assist people in detecting skin cancer suspicions, to determine whether they should seek further medical attention. It has to be safe for public use especially in populations where the disease is rampant, which is why clinical validation is a necessary process that should be taken.

3.1 Models:

This project would consider suitable supervised machine learning and deep learning algorithms that would aid in the early detection of melanoma. Most medical diagnostic problems use supervised learning because data obtained from medical imaging are often labeled by human experts. After the labeled data is trained, testing to verify the efficiency of the chosen algorithm is carried out by using unseen test data. (Adhikari 2019). The problem to be identified in this project is a classification one. For this report, the dataset contains several images of skin lesions which would be appropriately classified into malignant or benign using a suitable algorithm.

Deep learning is an evolving area of AI that is widely used in medical image analysis, achieving computer vision related tasks such as classification, detection and

segmentation. Deep learning as opposed to traditional machine learning algorithms allows for automatic feature extraction and classification. (X. Liu et al. Vol 2021)

Some proposed classification methods which would be explored in order to provide the best outcomes for accuracy prediction and detection of melanoma skin cancer are;

Support Vector Machines (SVM): SVM is a supervised classification algorithm that is based on the concept of decision planes that define decision boundaries This decision plane also known as the hyperplane separates the objects to distinguish classes. (H. Alquran *et al*. 2017) Research has shown that SVM has been successful in providing sufficient results for skin cancer detection. "Doukas, et al.(2012) developed a smart phone based system that was able to store the captured images of skin areas, extract a region of interest and then perform a self-assessment of the images. The system uses a mobile application to acquire and identify the moles in skin images and classify them as melanoma, nevus and benign lesions based on their brutality. The system implemented using 11 classifiers and the experimental result shows that the Support Vector Machine (SVM) has the highest accuracy of 77.06%". (H. Alquran *et al*. 2017)

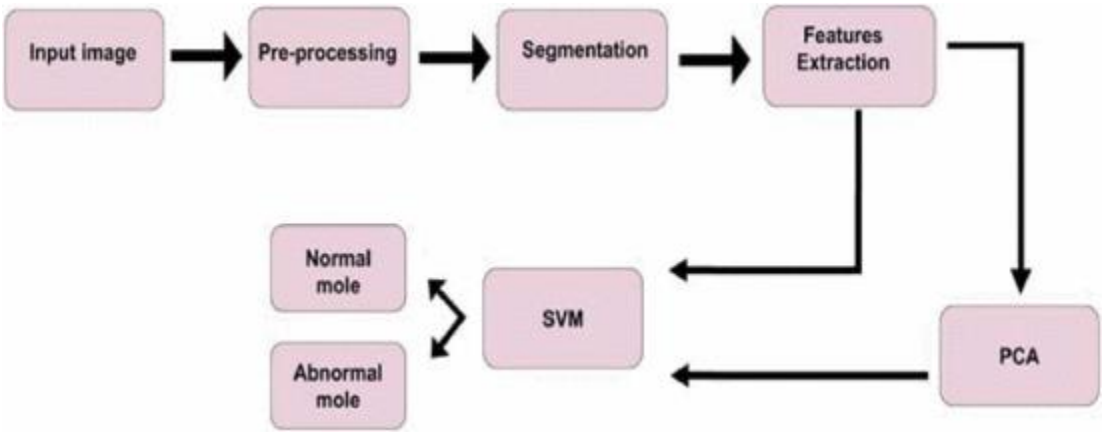


Figure 3: a proposed system block diagram

Convolution Neural Networks (CNN): CNN is a deep learning algorithm and a dominant classification framework for image analysis. A breakthrough test that demonstrated this was conducted by a research team from Germany, the United

States and France where an AI system was used to detect and classify malignant and benign skin lesions from a dataset of more than 100,000 images that were fed into a CNN. The results showed that in comparison to human experts who were able to perform with 86.6% accuracy, the AI system was able to classify with 95% accuracy. This established that CNN is able to accurately make image classification to identify skin cancer. **(**Adhikari 2019**)**. CNN is commonly used with architecture comprising of multiple layers that are trained. The two stages to train the network are; the forward stage and the backward stage. The goal of the forward stage is to represent the input image with the current parameters in each layer. The first layers find low-level features such as; edges and corners, and the other layers find mid-level and high-level feature such as; objects and shapes. (Traore, Kamsu-Foguem and Tangara 2018)

4. Project Implementation

There are several necessary steps which should be carried out to ensure a successful project implementation. They would be discussed in this chapter.

Data Collection: for this project the dataset used was obtained from a publicly available online data source called Harvard Dataverse Repository. The dataset is the (Humans Against Machines 10000) HAM10000 dataset. More information on this dataset would be discussed on chapter 5.1 of this project.

Pre-Processing: in this stage the dataset is going to be cleaned and prepared for modelling. In this step modifications to the images would also be carried out to make it befitting for classification. Such modifications could include; histogram modification; to enhance the contrast of the image and noise filtering: this reduces the impact of hair covered on the skin. (H. Alquran *et al*. 2017) during this stage, the intensity of the images is improved to decrease inconsistencies among photos. The image is additionally scaled and standardized to fit the training model.

Image Segmentation: the next stage is to detect and segment the region of interest (ROI) which is the lesion region. “The segmentation stage includes steps: Image thresholding, image filling, image opening, converting extracted region to gray level,

and then performing histogram equalization to the extracted gray level image". (H. Alquran *et al*. 2017)

Classification Algorithms: the chosen classification models that would be used for this project are SVM and CNN. Their end results would also be compared to the outcomes of other classification models.

Regularization: a central problem in machine learning is how to make an algorithm that will perform well not just on the training data but on the new inputs as well. There are multiple strategies in machine learning that are explicitly designed to reduce test errors, possibly at the expense of increased training error. These strategies are known collectively as regularization. (deep Learning: Methods and Applications pg. 201). Chapter 6 of this thesis would demonstrate a regularization technique that was utilized on the project dataset.

5. Methodology and Resources

This chapter would focus on the research processes undertaken for the project. There will be a discussion on how the data and other relevant sources used for this project were collected, analyzed and interpreted to communicate relevant findings and define the given objective.

5.1 The Dataset

There is sufficient availability of datasets online which contain various different kinds of skin diseases. However, most of the datasets I discovered while conducting my research were insufficient to showcase satisfactory academic skill, most of the datasets either contained very limited training data which is risky as the model needs to learn from a large amount of training data and would have likely caused overfitting. Another dataset was already split into training and test data which would have not allowed me to undergo pre-processing steps on the dataset.

The dataset used for this report was obtained from the Harvard Dataverse Repository which is a free data repository open to all researchers. The dataset is titled HAM10000 (Humans Against Machines 10000). This particular dataset has been used to conduct

milestone findings in the past, it was the training dataset used in the ISIC 2018 challenge. 'Test-set evaluations of the ISIC 2018 challenge were compared to physicians on an international scale, where the majority of challenge participants outperformed expert readers' (Tschandl P. et al., Lancet Oncol 2019).

When downloaded, the dataset contains a csv file with the headers; 'lesion_id', 'image_id', 'dx', 'dx_type', 'age', 'sex' and 'localization'. the important headers based on feature importance that were used for data preparations, analysis, visualization and modelling were 'dx' and 'image id'

Additionally, the dataset contains 10015 dermatoscopic images which are stored in two different zip files due to size limitations. (The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions, 2022). Each image falls into one of seven predefined disease classes;

- actinic keratoses akiec),
- basal cell carcinoma(bcc),
- benign keratinocytic lesions including solar lentigo(bkl),
- vascular lesions(vas),
- dermatofibroma(df),
- melanoma(mel)
- melanocytic nevi(nv).

The work environment where this dataset has been analyzed is Pycharm with Keras as the backend to provide high level building blocks for developing the deep learning models.

The research approach used to gather relevant sources and case studies for this dissertation was the 'quantitative research approach' this is because a bulk of the observations which have been discussed in this project are built upon pre-existing theories and data. Studies were experts examined and critically analyzed the accuracy rates of AI systems in detecting skin cancer were considered by accessing sources such as library books, Google scholar, ScienceDirect, Solent Online Library, health websites and journal articles.

6. Design and Implementation

This chapter would analyze the code snippets and their functions in providing the desired output from the dataset.

```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import os
from glob import glob
import seaborn as sns
from PIL import Image
import shutil

np.random.seed(33)
from sklearn.preprocessing import StandardScaler
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split, KFold, cross_val_score, GridSearchCV
from sklearn.metrics import accuracy_score, confusion_matrix
import itertools
```

Figure 4: importing the necessary libraries

The first step in data preparations is to import the necessary libraries.

Matplotlib.pyplot is used to plot graphs and plots using python scripts.

Numpy to act as a random number generator

Pandas for manipulating numerical tables

Seaborn is used to plot the confusion matrix

Glob to work through the folder structure and return a list of files or folders that match the specified path.

From PIL image function is used to load the images making it easy to resize the images and convert them to numpy array.

Random seed has been set to 42 to achieve a more accurate result as opposed to a random one.

Confusion matrix from sklearn has been used to plot and calculate the confusion matrix between the predicted and true output (y values).

Train_test_split to split the data into training and testing datasets

LabelEncoder to convert the skin disease labels into numbers 0-6.

```
import keras
from keras.utils.np_utils import to_categorical # used for converting labels to one-hot-encoding
from keras.models import Sequential, Model
from keras.layers import Dense, Dropout, Flatten, Conv2D, MaxPooling2D
from keras import backend as K
```

Figure 4.1: importing libraries

Keras is an open source library that provides a python interface for deep learning.

Import to_categorical is necessary because it is a multi-class classification and having labels of 0, 1, 2, 3, 4 would not be sufficient, it should be in a categorical format

```
df = pd.read_csv('HAM10000_metadata.csv')
print(df)
```

```
2       vidir_modern
3       vidir_modern
4       vidir_modern
...              ...
10010   vidir_modern
10011   vidir_modern
10012   vidir_modern
10013   vidir_modern
10014   vidir_modern

[10015 rows x 8 columns]
```

Figure 5: load the dataset

The next step is to read in the csv file. It can be seen that there are 10015 rows and 8 columns in the dataset.

These columns are; lesion_id, image_id, dx, dx_type, age, sex and localization

The dx column is the classification column. Therefore, it is the most important feature in the dataset.

```
df['image_id']=df['image_id'].apply(lambda x: x+ '.jpg')
print (df.head())
```

```
     lesion_id          image_id   dx dx_type   age   sex localization  \
0  HAM_0000118  ISIC_0027419.jpg  bkl   histo  80.0  male        scalp
1  HAM_0000118  ISIC_0025030.jpg  bkl   histo  80.0  male        scalp
2  HAM_0002730  ISIC_0026769.jpg  bkl   histo  80.0  male        scalp
3  HAM_0002730  ISIC_0025661.jpg  bkl   histo  80.0  male        scalp
4  HAM_0001466  ISIC_0031633.jpg  bkl   histo  75.0  male          ear
```

Figure 6: to edit the files in the image_id header to jpg

```
im = Image.open('images/ISIC_0024306.jpg')
plt.figure()
plt. imshow(im)
```

```
<matplotlib.image.AxesImage at 0x1dbcef45d00>
```

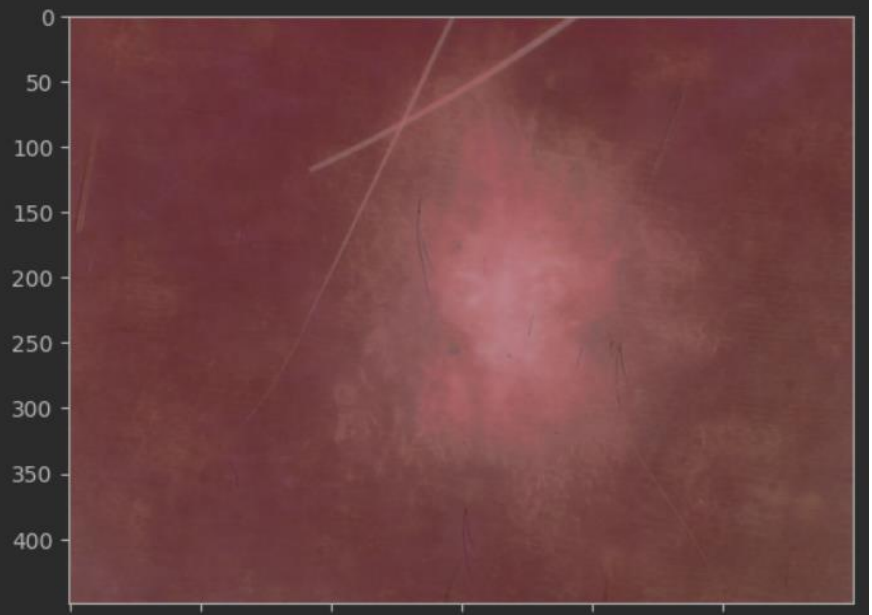


Figure 7: visualization of an image from the image folder.

```
labelencoder = LabelEncoder()
labelencoder.fit(df['dx'])
LabelEncoder()
print(list(labelencoder.classes_))

df['label'] = labelencoder.transform(df["dx"])
print(df.sample(5))
```

```
['akiec', 'bcc', 'bkl', 'df', 'mel', 'nv', 'vasc']
        lesion_id      image_id   dx    dx_type   age     sex  \
9126  HAM_0002936  ISIC_0024380   nv      histo  60.0    male
285   HAM_0002691  ISIC_0031345  bkl      histo  80.0  female
8022  HAM_0006295  ISIC_0032720   nv      histo  25.0  female
4689  HAM_0004969  ISIC_0025521   nv  follow_up  55.0  female
2785  HAM_0006727  ISIC_0024582  bcc      histo  75.0    male

         localization        dataset  label
9126             back      rosendahl      5
285              hand  vidir_molemax      2
```

Figure 8: LabelEncoding

The above figure displays the code to convert the labels in the column ‘dx’ into numbers from 0 to 6. The output created a new column titled label that appropriately assigned a number between o and 6 to a specific skin cancer class in the column ‘dx’.

```
fig = plt.figure(figsize=(13, 8))

ax1 = fig.add_subplot(221)
df['dx'].value_counts().plot(kind='bar', ax=ax1)
ax1.set_ylabel('count', size=12)
ax1.set_title('Cell Type')

ax2 = fig.add_subplot(222)
df['sex'].value_counts().plot(kind='bar', ax=ax2)
ax2.set_ylabel('Count', size=12)
ax2.set_title('Sex')

plt.tight_layout()
plt.show()
```

Figure 9: code for data distribution visualization

Data visualization is crucial as it is used to gain insights on the data and represent relationships and/or correlations between data. The code above is used to create bar plots to visualize the distribution of data in the columns ‘dx’ and ‘sex’

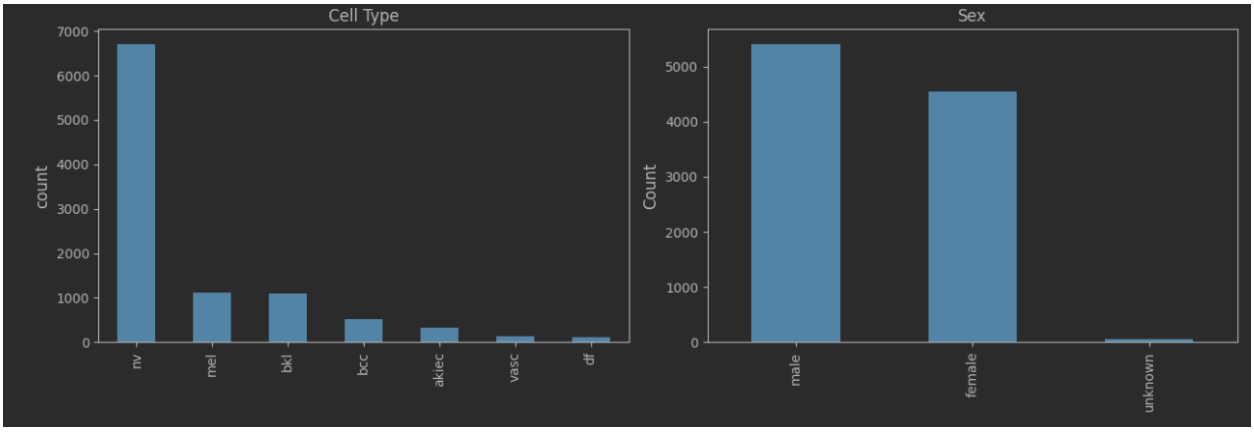


Figure 10: bar plots

From the image above it can be seen that the data in ‘dx’ is heavily unbalanced

```
from sklearn.utils import resample
print(df['label'].value_counts())

 5    6705
 4    1113
 2    1099
 1     514
 0     327
 6     142
 3     115
 Name: label, dtype: int64
```

Figure 11: the spread of the data

The above figure shows how unbalanced the data is. In order to balance the data, a code to take 500 images from each class to enable all the labels to have an even amount of images so the data does not overfit when modelling is necessary the function of this code is to upscale the classes that have low amounts to 500 by randomly copying images from a class with a high amount of image data.

```
df_0 = df[df['label'] == 0]
df_1 = df[df['label'] == 1]
df_2 = df[df['label'] == 2]
df_3 = df[df['label'] == 3]
df_4 = df[df['label'] == 4]
df_5 = df[df['label'] == 5]
df_6 = df[df['label'] == 6]
```

Figure 12: code to copy the images and duplicate them

Capture all the rows where the value equals 0, capture that into a new dataframe called df_0. Likewise for the other 5 labels

Separate each class so you can augment them individually

```
n_samples=500
df_0_balanced = resample(df_0, replace=True, n_samples=n_samples, random_state=42)
df_1_balanced = resample(df_1, replace=True, n_samples=n_samples, random_state=42)
df_2_balanced = resample(df_2, replace=True, n_samples=n_samples, random_state=42)
df_3_balanced = resample(df_3, replace=True, n_samples=n_samples, random_state=42)
df_4_balanced = resample(df_4, replace=True, n_samples=n_samples, random_state=42)
df_5_balanced = resample(df_5, replace=True, n_samples=n_samples, random_state=42)
df_6_balanced = resample(df_6, replace=True, n_samples=n_samples, random_state=42)
```

Figure 13: resample the data into the amount stated on n_samples which is 500

If its more than 500 it downscales to 500, if its more than 500 it upscales to 500

```
df_balanced = pd.concat([df_0_balanced, df_1_balanced,
                         df_2_balanced, df_3_balanced,
                         df_4_balanced, df_5_balanced, df_6_balanced])
```

Figure 14: code snippet to balance the data

Used to concatenate; put them all together in one dataframe.

```
print(df_balanced['label'].value_counts())
```

```
0    500
1    500
2    500
3    500
4    500
5    500
6    500
Name: label, dtype: int64
```

Figure 15: the output of the balanced data

```
folder_dir = os.getcwd() + r"C:\Users\Samuel\PycharmProjects\pythonProject19\images"#C:\Users\Samuel\Downloads\HAM10000 MAIN
 FOLDER\images"

dest_dir = os.getcwd() + r"C:\Users\Samuel\PycharmProjects\pythonProject19\reorganized" #r"C:\Users\Samuel\Downloads\HAM10000
 MAIN FOLDER\reorganized"
```

Figure 16: create a new folder containing all the image data and create a new folder called destination directory

The first step is to sort these images into sub folders labeled according to their classes. This would allow for easy extraction of the images.

The code snippet above would would take images from the folder_dir and put them in the dest_dir folder labeledd according to their classes.

```
df2 = pd.read_csv(r"C:\Users\Samuel\
print(df['dx'].value_counts())

#label=df2['dx'].unique().tolist()
#label_images = []
```

```
nv       6705
mel      1113
bkl      1099
bcc       514
akiec     327
vasc      142
df        115
Name: dx, dtype: int64
```

Figure 17: assigning a unique label

7. Conclusion

The fear of Early detection of melanoma skin cancer is necessary to ensure prevention and facilitate early treatment thereby reducing the mortality rate of the disease. The current practice for detecting skin cancer is through clinical examinations by a health care professional by using dermoscopy to assess suspected skin lesions. Diagnosis is later confirmed with biopsy, histological examination and a specialist pathological

interpretation by a specialist. To reduce the burden which the increased demand for these services has on health care specialists and to alleviate the financial and mental burden of these procedures on patients. AI systems which would aid in early detection of melanoma have been considered in this dissertation project.

At the beginning of this project, there was a goal to create a prototype layout of a smartphone application which would aid in early detection of melanoma. However, further research and studies brought a realization of the limitations of developing a smartphone app for clinical purposes. There are ethical and legal issues to be considered, as there is concern on the harmfulness an inaccurate and poorly designed app would have on patients. Consequently, further testing needs to be undertaken in a real-world clinical setting before inputting an ml algorithm to mobile devices. Nevertheless, if a smartphone app is developed accurately and undergoes all the necessary clinical validations, this technology would adequately improve diagnostic accuracy.

Several deep learning algorithms which are capable of accurately detecting skin cancer were considered in this project. However, there is a notable limitation with using deep learning, which is its heavy reliance on large training datasets. Nevertheless, with more availability of huge amounts of data, deep learning would continue to create milestones in the health care sector.

8. References

ADHIKARI, A., *A Thesis entitled Skin Cancer Detection using Generative Adversarial Network and an Ensemble of deep Convolutional Neural Networks 2019.*

Doukas Charalampos et al., "Automated skin lesion assessment using mobile technologies and cloud platforms", *Engineering in Medicine and Biology Society (EMBC) 2012 Annual International Conference of the IEEE. IEEE*, 2012.

TRAORE, B.B., B. KAMSU-FOGUEM and F. TANGARA, 2018. Deep convolution neural network for image recognition. *Ecological Informatics,* 48, 257-268

H. ALQURAN *et al.,* 2017. The melanoma skin cancer detection and classification using support vector machine. *- 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT).* pp.1-5

N. SHARMA *et al.,* 2021. A Stochastic Neighbor Embedding Approach for Cancer Prediction. *- 2021 International Conference on Emerging Smart Computing and Informatics (ESCI).* pp.599-603

Oregoncancer.com. 2022. *5 Apps That Help Detect Skin Changes That May Lead to Skin Cancer*. [online] Available at: <https://www.oregoncancer.com/blog/apps-to-help-detect-skin-cancer> [Accessed 3 July 2022].

Liu, X., Gao, K., Liu, B., Pan, C., Liang, K., Yan, L., Ma, J., He, F., Zhang, S., Pan, S. and Yu, Y., 2021. Advances in deep learning-based medical image analysis. *Health Data Science*, *2021*.

Merlino G, Herlyn M, Fisher DE, Bastian BC, Flaherty KT, Davies MA, Wargo JA, Curiel-Lewandrowski C, Weber MJ, Leachman SA, Soengas MS, McMahon M, Harbour JW, Swetter SM, Aplin AE, Atkins MB, Bosenberg MW, Dummer R, Gershenwald JE, Halpern AC, Herlyn D, Karakousis GC, Kirkwood JM, Krauthammer M, Lo RS, Long GV, McArthur G, Ribas A, Schuchter L, Sosman JA, Smalley KS, Steeg P, Thomas NE, Tsao H, Tueting T, Weeraratna A, Xu G, Lomax R, Martin A, Silverstein S, Turnham T, Ronai ZA. The state of melanoma: challenges and opportunities. Pigment Cell Melanoma Res. 2016 Jul;29(4):404-16. doi: 10.1111/pcmr.12475. Epub 2016 Apr 17. PMID: 27087480; PMCID: PMC5228487.

B. SREEDHAR, M. SWAMY B.E and M. S. KUMAR, 2020. A Comparative Study of Melanoma Skin Cancer Detection in Traditional and Current Image Processing Techniques. - *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*. pp.654-658

team, T., 2022. *Skin cancer | Cancer types | World Cancer Research Fund UK*. [online] World Cancer Research Fund. Available at: <https://www.wcrf-uk.org/cancer-types/skin-cancer/> [Accessed 4 July 2022].

nhs.uk. 2022. *Melanoma skin cancer*. [online] Available at: <https://www.nhs.uk/conditions/melanoma-skin-cancer/> [Accessed 5 July 2022].

The Skin Cancer Foundation. 2022. *Melanoma*. [online] Available at: <https://www.skincancer.org/skin-cancer-information/melanoma> [Accessed 6 July 2022].

Yu, K., Beam, A. and Kohane, I., 2018. Artificial intelligence in healthcare. *Nature Biomedical Engineering*, 2(10), pp.719-731.

Ranschaert, E.R., Morozov, S. and Algra, P.R. eds., 2019. *Artificial intelligence in medical imaging: opportunities, applications and risks*. Springer.pg. 4-5

Wang, W. and Siau, K., 2019. Artificial Intelligence, Machine Learning, Automation, Robotics, Future of Work and Future of Humanity. *Journal of Database Management*, 30(1), pp.61-79.

Tschandl, P., Codella, N., Akay, B., Argenziano, G., Braun, R., Cabo, H., Gutman, D., Halpern, A., Helba, B., Hofmann-Wellenhof, R., Lallas, A., Lapins, J., Longo, C., Malvehy, J., Marchetti, M., Marghoob, A., Menzies, S., Oakley, A., Paoli, J., Puig, S., Rinner, C., Rosendahl, C., Scope, A., Sinz, C., Soyer, H., Thomas, L., Zalaudek, I. and Kittler, H., 2019. Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. *The Lancet Oncology*, 20(7), pp.938-947.

National Cancer Institute. 2022. *Skin Cancer Screening (PDQ®)–Patient Version*. [online] Available at: <https://www.cancer.gov/types/skin/patient/skin-screening-pdq> [Accessed 25 August 2022].

Nicholas Brownstone, N., 2022. *Controversies in Skin Cancer Screening: The Debate Rages On*. [online] Dermatology Times. Available at: <https://www.dermatologytimes.com/view/controversies-in-skin-cancer-screening> [Accessed 26 August 2022].

Harvard Dataverse. 2022. *The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions*. [online] Available at: <https://doi.org/10.7910/DVN/DBW86T> [Accessed 29 August 2022].

Phillips, M., Greenhalgh, J., Marsden, H. and Palamaras, I., 2019. Detection of Malignant Melanoma Using Artificial Intelligence: An Observational Study of Diagnostic Accuracy. *Dermatology Practical & Conceptual*, p.e2020011.

Malvehy, J., Marchetti, M., Marghoob, A., Menzies, S., Oakley, A., Paoli, J., Puig, S., Rinner, C., Rosendahl, C., Scope, A., Sinz, C., Soyer, H., Thomas, L., Zalaudek, I. and Kittler, H., 2019. Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. *The Lancet Oncology*, 20(7), pp.938-947.

Duquesne University School of Nursing. 2022. *Formulating a PICOT Question | Duquesne University*. [online] Available at: <https://onlinenursing.duq.edu/blog/formulating-a-picot-question/#:~:text=The%20word%20PICOT%20is%20a,phrased%20to%20elicit%20an%20answer.> [Accessed 10 August 2022].

Das, K., Cockerell, C., Patil, A., Pietkiewicz, P., Giulini, M., Grabbe, S. and Goldust, M., 2021. Machine Learning and Its Application in Skin Cancer. *International Journal of Environmental Research and Public Health*, 18(24), p.13409.

Developer.ibm.com. 2022. *IBM Developer*. [online] Available at: <https://developer.ibm.com/articles/an-introduction-to-deep-learning/> [Accessed 12 August 2022].

Icpme.us. 2022. [online] Available at: <https://www.icpme.us/courses/CIBC_PDF/4.1%20CAD_Radiomics_and%20AI%20in%20Breast%20Imaging_Giger.pdf> [Accessed 25 August 2022].

Freeman, K., Dinnes, J., Chuchu, N., Takwoingi, Y., Bayliss, S., Matin, R., Jain, A., Walter, F., Williams, H. and Deeks, J., 2022. *Algorithm based smartphone apps to assess risk of skin cancer in adults: systematic review of diagnostic accuracy studies*.

Phillips, M., Greenhalgh, J., Marsden, H. and Palamaras, I., 2019. Detection of Malignant Melanoma Using Artificial Intelligence: An Observational Study of Diagnostic Accuracy. *Dermatology Practical & Conceptual*, p.e2020011.

Tschandl, P., Codella, N., Akay, B., Argenziano, G., Braun, R., Cabo, H., Gutman, D., Halpern, A., Helba, B., Hofmann-Wellenhof, R., Lallas, A., Lapins, J., Longo, C., Malvehy, J., Marchetti, M., Marghoob, A., Menzies, S., Oakley, A., Paoli, J., Puig, S., Rinner, C., Rosendahl, C., Scope, A., Sinz, C., Soyer, H., Thomas, L., Zalaudek, I. and Kittler, H., 2019. Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. *The Lancet Oncology*, 20(7), pp.938-947.

9. Acronyms

AI – Artificial Intelligence

ML – Machine Learning

PICOT- Patient. Intervention. Comparison. Outcome. Time

CT - Computed Tomography

nv - Melanocytic nevi

mel - Melanoma

bkl - Benign Keratosis-like lesions

bcc - Basal cell carcinoma

akiec - Actinic Keratoses

vas - Vascular lesions

df – Dermatofibroma

CNN – Convolution Neural Networks

SVM – Support Vector Machines

QUADAS-2 - Quality Assessment of Diagnostic Accuracy Studies 2 tool