Skin Lesion Detection and Classification using Deep Learning Techniques

Elaboration of Problem & Literature Review

Project Advisor:
Mr. Bismillah Jan

Group Members:

Zara Ahmad	15L-4351
Fatima Tahira	15L-4169
Hafiz Ahmad Hassan	15L-4359
Adeel Khan	15L-4006

National University Of Computer and Emerging Sciences
Department of Computer Science
Lahore, Pakistan

Table of Contents

LIST OF TABLES

1. ABSTRACT

Recent technologies of machine learning, especially deep neural networks have made many breakthrough in different fields. One important field where these techniques can be applied is digital image processing in medical domain. To detect and classify skin lesion in images using machine learning techniques is a crucial task to assist physician in decision making related to patient health.
This work focus on skin lesion classification and detection in images using deep learning techniques. In first stage Convolution Neural Network (CNN) is used to classify melanoma images of benign and malignant tumours. In second stage we will use the same technique to detect cancerous lesion in melanoma images. This Tumour detection model is basically an expert system built to detect skin cancer in patients. The model is trained on images using deep learning techniques. Initial results obtained using above mentioned technique has 78% accuracy. The aim is to improve accuracy from that already achieved by previous research works.

2. INTRODUCTION

Deep learning is a piece of a more extensive group of machine learning. It is based on learning data representations rather than task-specific algorithms. In Image Processing different algorithms are applied to images to analyse and extract the required results.
The purpose of this project is to build a model that aid doctors in diagnostics. It can be quite difficult to detect a disease as positive or negative and being humans, doctors naturally require an additional opinion at times to help or confirm their diagnostic. The model is hence intended to initially serve as a second opinion to a doctor's diagnosis.
To build our expert system, we will be using a Convolution Neural Network and some deep learning techniques. The model takes an image as input and detects skin lesion in the person. Upon positive results, the model further classifies the tumor as 'Benign', 'Malignant' or 'Unknown'.
A tumor may not be either of the two i.e. neither completely beinign nor completely malignant. The newer research being done is introducing newer types of classes. Hence to classify those images, we have used the label 'Unknown'.
Models detecting and classifying Skin Tumor detection have been built previously but we aim to use better and more efficient algorithms in our model to improve the accuracy that has been achieved up till now.
The dataset that is used for classification is taken from ISIC Archive [6]. Models detecting and classifying Skin Tumour detection have been built previously but we aim to use better and more efficient algorithms in our model to improve the accuracy that has been achieved up till now.

3. GOALS AND OBJECTIVES

1. Classify positive skin cancer results as 'Benign', 'Malignant' or 'Unknown' using Deep Learning Models
2. Skin Lesion Detection using CNN
3. Make our own neural network using hyperparameters

4. PROJECT SCOPE

In first phase we will train convolution neural network to classify skin melanoma images into ‘Benign’, ‘Malignant’ or ‘Unknown. In second phase our model will detect skin cancerous lesion using the same approach, to train the convolution neural network. In the end we will try to make our own neural network using hyperparameter.

5. WORK DONE SO FAR

So far we have studied various research papers on the subject and taken online tutorials on Machine Learning and Deep learning. The research papers we read had used insufficient data which is not enough for an unbiased analysis and can lead to model under fitting. In paper [1] we found out that some machine learning models implemented before have classification accuracy of 85% using Support Vector Machine and 95% using Back-Propagation Neural Network. They have trained the model on only 100 images and their accuracy results too are based on a very small chunk of test data i.e. one hundred images.

In research paper [1] hybrid (combination of BNN and SVM) model approach has been used to classify skin cancer as benign or malignant only. They have not handled unlabelled during training and testing phase. By using hybrid approach and increasing feature vector size twice to improve accuracy from 70% to 95%, they have also increased the complexity of the Neural Network Model. Under fitting has also resulted due to small size of dataset used for training and testing. Moreover, it was also stated in the paper that these results are initial and not final.

Although a high accuracy has been achieved in [1], paper [3] mentions that accuracy may not always be a good measure of evaluating a models performance. Other factors need to be considered too. A model for example trained on uneven data with a majority of one class maybe correctly able to identify only that class. Now when the test data containing a higher proportionality of the same class comes, with minimal or no test data belonging to other classes, the model will give a very high accuracy rate since the test data is mostly composed of the class it classifies correctly.

In [4] Image processing, techniques are applied in which Images were enhanced using Gabor filter and FFT. Then it was segmented using threshold and Watershed segmentation approach. As a result, features are extracted using binarization and masking approach, the true acceptance rate (TAR) of this method is (85.7%) and false acceptance rate (FAR) is (14.3%) without using any deep learning techniques.

Using deep learning techniques we aim to achieve more than 90% accuracy with models that are defined in project scope. We have also built a first draft version of our model which classifies the tumor with an accuracy of 78%. In the upcoming weeks we will work on training and testing our neural network with complete dataset from [6], improving its accuracy, handle unlabeled data by better pre-processing and extracting useful features. We will improve our algorithm to classify more efficiently.

6. ELABORATION OF PROBLEM

Our problem is to train a model to detect skin cancerous lesions and to further classify them into three classes, ‘Benign’, ‘Malignant’ or ‘Unknown’. Benign represents a class of tumor composed of non-cancerous cells. Malignant represents tumor composed of cancerous cells. They are invasive, composed of abnormal cells and do not stop growing. Unknown represents class of tumors images that cannot be classified as either benign or malignant. The problem is broken into sub problems listed below:

1. **Dataset**
   Data is acquired from ISIC [6] which contains 23887 images on which the model is trained. These are skin melanoma images.

2. **Image Pre-processing**
   Dataset acquired consists of images of differing dimensions so all of them are resized to have the same dimensions 300*300. The dimension size is chosen keeping in view the limited runtime space available with Google Colab.
   Also, the RGB images are converted to gray level images.

3. **Image Segmentation**
   The lesion in the images is segmented from the surrounding skin. This is done using thresholding since there is a clear distinction of color between the lesion and the surrounding skin

4. **Convolution Neural Network (CNN)**
   Model is built using CNN which is trained on the processed dataset. The trained model then classifies the images as ‘Benign’, ‘Malignant’ or ‘Unknown’.

5. **Evaluation Protocols**
   Result are evaluated based upon the features (weights) extracted using CNN.
   For classification, cross entropy function will be used and for detection Jaccard index will be used as evaluation function.
   Cross Entropy is used to quantify the difference between two probability distributions. Cross entropy loss determines how far away the prediction is from true distribution. Following is the formula [9] to compute cross entropy:

$$\mathbf{H}(\mathbf{p},\mathbf{q}) = -\sum_{\boldsymbol{x}} \boldsymbol{p}(\boldsymbol{x}) \log \boldsymbol{q}(\boldsymbol{x})$$

   Jaccard index computes the overlapping region between two sets, in our case these will be predicted segmented lesion and actual threshold lesion.
   Finally several calculations are also done to further evaluate a model. They include accuracy, precision, recall and specificity. These terms are elaborated in the upcoming section.

7. LITERATURE REVIEW

7.1 INTRODUCTION

Literature review is the first step in any research based project, helping us identify the issues with the current methods, paving way for future work and possible problems that researchers can face. A significant amount of research is being done every year in this domain. Various ISIC challenges on tumor detection and classification are held every year but there is yet to develop a model that accurately classifies the tumors given a large dataset, given that it's also tested using suitable accuracy measures. If done properly, this could save millions of doctors and patients around the world valuable resources and time.

To gain a deeper insight into our problem we analyzed [1][2][3][4][5][7][8] papers, from as early as 2007 to as recent as 2018 , which include three papers related to lung cancer and three related to skin cancer tumor detection and classification. This literature review focuses on the of pros and cons of various machine learning approaches and image processing techniques used for classification and detection analyzing them from different angles including unlabeled data handling, dataset size, biasness in data and the accuracy measures employed to test theirs results.

7.2 DETAILED LITERATURE REVIEW

This section provides a detailed review of the literature we read to gain a deeper insight into the subject and related work.

7.2.1 Definitions

- **Lesion:** Infected area of skin
- **Benign:** Tumor composed of non-cancerous cells
- **Malignant:** Tumor composed of cancerous cells
- **Thresholding:** Method of image segmentation
- **SVM:** Support Vector Machines which are used to analyze data for classification
- **Jaccard Index:** It is a statistic used to compute similarity/diversity between sample sets
- **TP, TN, FP, FN:** True Positive, True Negative, False Positive, False Negative

7.2.2 Classification

Various techniques are used to classify the tumor images.

- **Back Propagation Neural Network:** The input is propagated through the network and the error is calculated which is then propagated back into the network while adjusting the weights such that error is reduced. [1]

- **Support Vector Machine:** Through SVM, classification is performed by constructing an N-dimensional hyper-plane that optimally separates data into different categories. It separates clusters of vectors in such a way that different categories of target variable fall on different sides of the plane. [1]

- **Bayesian Classifier:** Paper [7] uses Bayesian classifier to estimate the probability of being a lesion pixel from all the pixels in a given image.

- **Convolution Neural Network:** In paper [8], Convolution Neural Network is used for skin lesion segmentation. The architecture uses 27 convolution blocks, 5 up-convolution layers, 5 max-pooling layers and 1 output layer.

7.2.3 Summary of Papers

Different researchers have applied different techniques to classify the tumor images. These papers are summarized below

1. **Paper:** A Novel Hybrid System for Skin Lesion Detection [1]
   Authors: Andy Chiem, Adel Al-Jumaily, Rami N. Khushaba
   Classification Techniques: Back Propagation using WPT features extraction Neural Network and Support Vector Machine
   Accuracy: 95.1% (Neural Network) and 85% (SVM)

   In this paper, a method based upon mixture of methods for classifying benign and malignant melanoma lesions is implemented. It consists of four stages, namely pre-processing stage, segmentation stage, feature extraction stage and classification stage.
   Primarily, the images are acquired using the Dermatoscopy technique. It is a non-invasive imaging technique that works by applying soaking oil onto the skin lesions to picture subsurface skin structures by making it translucent. Secondly, the images gathered are pre-processed using a median filter which ensures that unwanted structures are excluded from the image such as fine hair, noise and air bubbles etc. This process is followed by a contrast enhanced skin which make the edges of the lesions prominent. Thirdly, image segmentation is performed that divides the lesion from the skin around it using the thresholding method followed by a boundary tracing algorithm that validates the division. Fourthly, the important features of the image that plays an important role in the distinction of malignant and benign lesions are extracted with the help of wavelet algorithm. This algorithm works by providing detailed information about an image on different scale stressing upon different features in each scale. It presents information such as the texture and granularity of the image. Finally, the images are classified as either malignant or benign lesions.

Two algorithms used for this classification are the Black-propagation Neural Network and Support Vector Machine giving the accuracy of 95% and 85% respectively.

2. **Paper:** Skin Lesion Segmentation and Classification for ISIC 2018 Using Traditional Classifiers with Hand-Crafted Features [7]
Authors: Russell C. Hardie, Redha Ali, Manawaduge Supun De Silva, and Temesguen Messay Kebede
Detection Techniques: RGB Bayes Classifier guided by regression network
Jaccard Index: 0.701
Classification Technique: Hand crafted features
Recall (Sensitivity): 0.7303

This paper provides with the methodology and the results obtained in Skin Lesion Segmentation and disease classification. The approach taken in this paper is to uses handcrafted features along with traditional classifiers. In Phase 1 for skin lesion detections, they processed the images in RGB space, using RGB color vectors to distinguish between normal and lesion tissue. Gaussian mixture models (GMMs) are used to estimate the probability density functions for tissues and a Bayesian classifier is used to estimate the posterior probability of being a lesion pixel for all pixels in a given image. A threshold for image segmentation is chosen using SVM based on which value maximizes the overlap score (Jaccard Index). Morphological operations are used to fine tune the segmentation. In Phase 2, disease classification is done by SVM using 200 hand crafted features. Features are computed from the RGB image and the segmentation mask obtained in phase 1.
Phase 1, Lesion Detection, results in a mean overlap score of 0.701. In phase 2, Lesion Classification, the average class recall (or sensitivity) using 5-fold cross-validation on the provided training imagery is 0.7303.

3. **Paper:** Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks [8]
Authors: Hongming Xu and Tae Hyun Hwang
Classification Techniques: Skin Lesion segmentation using Convolution Network
Jaccard index: 0.7344

This paper implements segmentation of skin lesions from surrounding normal regions. A deep fully convolutional network that maps the input dermoscopic image into a posterior probability map is trained. The architecture of deep learning model has 27 convolution blocks, 5 up-convolution layers, 5 max-pooling layers and 1 output layer (1x1 convolution operation followed by sigmoid activation). The network is trained using Adam optimization with a learning rate of 10^{-4}**.** Loss function used incorporates binary cross entropy and Dice coefficient. After obtaining the posterior map of validation image from trained model, dual-threshold method is used to generate a binary skin. The average segmentation score obtained is 0.7344.

The following Papers are based on lung cancer dataset. Although they are not on Skin Lesion but they play an important role in our research based project by giving an insight on the different methods implemented to process images or detecting lung cancer which are very much related to our field of research.

4. **Paper:** Lung Cancer Detection Using Image Processing Techniques [4]
   Authors: Mokhled S. AL-TARAWNEH
   Techniques: Gabor Filter, Segmentation, Normality comparison

   Image processing techniques have been implemented in this paper. This research was divided into three stages namely image enhancement, image segmentation and feature extraction. Image enhancement techniques include spatial domain methods and frequency domain methods, with the aim to improve the interpretability of images for human viewers. It utilizes Gabor filter, Fast Fourier transform and auto enhancement. Gabor filter is a linear filter, defined by harmonic function multiplied by Gaussian function. Whereas Fourier transform works on transform of the image in frequency domain. Image segmentation part is used to locate boundaries and to simplify the image with the goal of assigning labels to every pixel in an image, resulting in a set of segments that cover the entire image. Segmentation algorithms are based on the discontinuity and similarity properties of image intensities. Image segmentation can be done through Thresholding approach or Marker-Controlled Watershed Segmentation approach. Thresholding is a non-linear function which converts an image into a binary image where levels are assigned to pixels above or below a threshold. Marker-driven segmentation technique extracts seeds that indicate presence of objects or background at specific image locations. In feature extraction phase, algorithms and techniques are used to isolate desired portions of an image, usually done by Binarization or Masking approach.
   The main detected features for accurate image comparison are pixel percentage and mask labeling.

5. **Paper:** Lung Nodule Classification Combining Rule-based and SVM [5]
   Authors: Zhang Jing, Li Bin, Tian Lianfang
   Classification Techniques: Combination of Rule Based and SVM
   Accuracy: 84.39%

   In order to intelligently identity the lung nodules, the paper proposes an approach combining rule-based and SVM to classify the lung nodules. To do this there are three possible methods, first is rule based, second being SVM and third is the combination of both.
   Initially the features are calculated. Total of thirteen features are extracted including seven shape features: area, perimeter, rectangle degree, ellipticity, circularity and slenderness, two gray features: gray means and gray variance and four texture parameters: energy, contrast, entropy and adverse moment.
   Secondly, if all the 13 features are taken into consideration then the computation can be wrong and some abundant blood vessels may take longer time in

classification thus decreasing the efficiency of SVM. As a result some non-nodule candidates can be ignored by following the simple rules which are:

i. Rule 1: if the rectangle degree R >= 1, then the ROI is considered to be the blood vessel and excluded;
ii. Rule 2: if the slenderness S <= 0.4, then the ROI IS considered to be the blood vessel and excluded;
iii. Rule 3: if the ellipticity e >= 4, then the ROI IS considered to be the blood vessel and excluded.

Finally all the extracted features are taken as an input to the SVM in order to classify the candidate. The input data of SVM are normalized to [0, 1], and the output of SVM is divided into two classes: nodule (out = 1) and non-nodule (out = -1).
The accuracy obtained by combining Rule Based and SVM is 84.39%.

6. **Paper:** Performance comparison of artificial neural network and logistic regression model for differentiating lung nodules on CT scans [2]
Authors: Hui Chen, Jing Zhang, Yan Xu, Budong Chen, Kuan Zhang
Classification Technique 1: Artificial Neural Network
Accuracy: 90%
Classification Technique 2: Logistic Regression
Accuracy: 86.9%

This paper analyzes ANN and LR model and their differences when it comes to detection. The database analyzed in this study includes 135 malignant and 65 benign nodules. In this study, an LR model with seven independent variables with seven selected features is developed with classification threshold set to 0.5 and a three layer feed forward ANN with a 12-5-1 nodal architecture is constructed and the performance of the two models analyzed. In this study three aspects of model performance are considered; calibration (prediction accuracy over an entire range), clinical usefulness and discrimination (accuracy when distinguishing between outcomes, in this case benign and malignant). For an unbiased analysis there was insufficient data for training and testing. Therefore bootstrap resampling approach was used. ANNs had higher discriminative performance than LR model. In addition to that the overall accuracy rate was also higher for ANN's. This study reveals that ANN's outperformed LR models in discrimination and clinical usefulness but did not outperform in calibration.

7. **Paper:** Skin Lesion Classification using Hybrid Deep Neural Networks [10]
Authors: Amirreza Mahbod, Rupert Ecker, Isabella Ellinger
Classification Technique: Convolutional Neural Network
Accuracy: 84.7%

In this paper, they have used the pre-trained CNN model for classification of three classes melanoma, seborrheic keratosis and benign. They have trained the model with 2000 images from dataset provided by ISIC. They have used features for classification like texture structure, color and other morphological features. After training they have tested the model with around 150 images and evaluate the result using evaluation method AOC is 84.8% and the result of binary classification for Melanoma and seborrheic keratosis is 93.6%.

The dataset composed of 2000 color dermoscopic skin images with corresponding labels. Three diff erent skin lesion types are found in this dataset including 374 melanoma images, 254 seborrheic keratosis and 1372 benign nevi images with various sizes (from 1022×767 to 6748×4499), After that they have applied the pre-processing techniques like RGB to grayscale and Resizing the images. From 2000 images they flip the each image 0, 90, 180,270 and create a artificial dataset for training model.

After that they have use AlexNet architecture and VGG-16 architecture as feature extractor. They have used 2 fully connected layers and one detector layer with dimension 1000.After feature extraction they have used SVM for multiclass classification(three classes).For evaluating SVM result they have mapped the probabilities from SVM to logistic regression.

8. **Paper:** Skin Lesion Classification Via Combining Deep Learning Features and Clinical Criteria Representations [11]
Authors: Xiaoxiao Li , Junyan Wu, , Hongda Jiang , Eric Z. Chen4 , Xu Dong5 and Ruichen Rong
Classification Technique: Convolutional Neural Network
Accuracy: 85.3%

In this paper novel method based on deep convolutional neural networks are used to solve skin lesion analysis towards melanoma detection problem. Fine-tuned state-of-the-art image classification networks − ResNet50 and Dense201 to encode the image features. Accurate segmentation can benefit recognizing the skin lesion classes. UNet model is used to generate lesion masks.
Feature Fusion LightGBM to combine the traditional features and CNN features. It provides multiple hyper-parameters for achieving best performance.
Dataset containing 10015 images are classified into seven different classes
Then LightGBM combining the features obtained after applying pre-trained models and selected the best ensemble parameters used to train the LightGBM.
Evaluation measures are applied on training dataset. Afterwards ResNet, ResNet +Crop, DenseNet +Crop, Clinical Feature and AllFusion model results are obtained on validation and training images.
Then it is analyzed that fusion model achieved better and more stable performance on normalized multi-class accuracy i.e 0.853

7.2.4 Content Representation

Upon obtaining positive results of skin lesion, we further classify the image into two types:

- **Benign:** Tumor composed of non-cancerous cells
- **Malignant:** Tumor composed of cancerous cells
- **Unknown:** Tumors images that cannot be classified as either benign or malignant

7.2.5 Evaluation Measures

Different performance measuring metrics exist and based on the model, the most suitable measure is chosen. Some of the ways to evaluate performance are discussed below [3].

i. **Jaccard Index:**
It is a statistic used to compute similarity/diversity between sample sets.

ii. **Confusion Matrix:**
It is used to determine the accuracy and correctness of the model and is used in classification problems which have outputs with two or more types of classes. Confusion Matrix is a matrix with predicted and actual results across rows and columns. The table below is a good representation of confusion matrix.

	Disease	**No Disease**
Positive Test Result	True Positive (TP)	False Positive (FP)
Negative Test Result	False Negative (FN)	True Negative (TN)

Table 1: Confusion Matrix [3]

False positive and false positives are the incorrect result and it is our goal to minimize these to make our model a better classifier.

iii. Accuracy:

It is the correct predictions made by our model over all the predictions made.

$$\text{Accuracy [3]} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+FP+FN+TN}}$$

iv. Precision:

It tells us the proportion of objects we classified correctly over the total number of objects that belonged to that class. Gives model's performance with respect to false positives i.e. how many did we caught.

$$\text{Precision [3]} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}}$$

v. Recall or Sensitivity:

Recall tells what proportion of objects that belonged to a particular class were classified correctly by our model. It gives model's performance with respect to false negatives i.e. how many did we miss.

$$\text{Recall [3]} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}}$$

vi. Specificity:

Specificity is the exact opposite of recall. It tells what proportion of objects that did not belong to a particular class were predicted correctly as not belonging to that class by our model.

$$\text{Specificity [3]} = \frac{\boldsymbol{TN}}{\boldsymbol{TN+FP}}$$

vii. F1 Score:

This represents both Precision (P) and Recall (R). One way of getting this score is through taking the arithmetic mean of both the values **(P+R)/2.** But harmonic mean is a better way of computing the score as it gives better performance measures in most cases. The formula is written below.

$$\text{F1 Score [3]} = \frac{\boldsymbol{2*P*R}}{\boldsymbol{P+R}}$$

7.3 LITERATURE REVIEW SUMMARY

Paper	Author	Approach	Score	Description
A Novel Hybrid System for Skin Lesion Detection [1]	Andy Chiem, Adel Al-Jumaily, Rami N. Khushaba	Classification using Back-Propagation Neural Network	**Accuracy:** 95.1%	Through WPT, features are extracted and fed into Back-Propagation Neural Network with 2 hidden layers are used to classify images as either benign or malignant
A Novel Hybrid System for Skin Lesion Detection [1]	Andy Chiem, Adel Al-Jumaily, Rami N. Khushaba	Classification using Support Vector Machine	**Accuracy:** 85%	For classification into benign or malignant N-dimensional hyperplane is constructed using SVM. SVM models in such a way that different categories of target variable fall on different side of the plane in clusters.
Skin Lesion Segmentation and Classification for ISIC 2018 Using Traditional Classifiers with Hand-Crafted Features [7]	Russell C. Hardie, Redha Ali, Manawaduge Supun De Silva, and Temesguen Messay Kebede	**i. Detection Techniques:** RGB Bayes Classifier guided by regression network **ii. Classification Technique:** Hand crafted features	**i. Jaccard Index:** 0.701 **ii. Recall/ Sensitivity:** 0.7303	First segmentation is done to detect skin lesion area then SVM is used for classification
Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks[8]	Hongming Xu and Tae Hyun Hwang	Skin Lesion segmentation using Convolution Network	**Jaccard index:** 0.7344	Test problem involved segmenting the image to detect skin lesion

Paper	Author	Approach	Score	Description
Skin Lesion Classification using Hybrid Deep Neural Networks [10]	Amirreza Mahbod, Rupert Ecker, Isabella Ellinger	Convolutional Neural Network	84.7%	In this paper, they have used the pre-trained CNN model for classification of three classes melanoma, seborrheic keratosis and benign.
Skin Lesion Classification Via Combining Deep Learning Features and Clinical Criteria Representations [11]	Xiaoxiao Li, Junyan Wu, Hongda Jiang, Eric Z. Chen4, Xu Dong5 and Ruichen Rong	Convolutional Neural Network	85.3%	Deep convolutional neural networks are used to solve skin lesion analysis towards melanoma detection problem.

Table 2: Summary of Research Papers I

Papers not directly related to our research project but which certainly play an important role in extending our research are summarized below.

Paper	Author	Approach	Score	Description
Lung Cancer Detection Using Image Processing Techniques [4]	Mokhled S. AL-TARAWNEH	Gabor filter, Segmentation using watershed, normality comparison	**Accuracy:** 85%	Image processing techniques have been researched upon in this paper which are the first step in medical diagnosis
Lung Nodule Classification Combining Rule-based and SVM [5]	Zhang Jing, Li Bin, Tian Lianfang	Combination of Rule Based and SVM	**Accuracy:** 84.39%	The paper proposes an approach combining rule-based and SVM to classify lung nodules. Some nodules are ignore on basis of rules. Finally features calculated are inputted to SVM and results obtained.
Performance comparison of artificial neural network and logistic [2]	Hui Chen, Jing Zhang, Yan Xu, Budong Chen, Kuan Zhang	i. Artificial Neural Network ii. Logistic Regression	i. **Accuracy 1:** 90% ii. **Accuracy 2 :** 86.9%	This paper analyzes ANN and LR models for classifying lung nodules. Three aspects of model performance are considered; calibration, clinical usefulness and discrimination. Study reveals that ANN's outperformed LR models in discrimination and clinical usefulness but did not outperform in calibration.

Table 3: Summary of Research Papers II

7.4 CONCLUSION

The work that we have analyzed so far employed various techniques including back-propagation neural network, SVM, ANN. Giving accuracy of 95% with back-propagation neural network [1], 85% with SVM [1] and about 90% with ANN [2]. We have found that most papers did not handle unlabeled data in ISIC archive. In addition to this small chunk of data is used for training and testing mostly 200 images. The studies conducted so far lacked proper accuracy measures. In paper [2] unbalanced data has been used to train and test their model, using 135 malignant and 65 benign, putting a question mark on their accuracy claims. In reference paper [1] it is mentioned that their results are initial and not final. Our aim is to achieve more than 90% accuracy using CNN while handling unlabeled data using a minimum of 6000 images.

The model could be extended in future by further classifying the unlabeled data into distinct types using unsupervised or semi-supervised learning approach.

7.5 APPENDIX

Paper Number	Training data Size	Testing	Data Source	Year	Conference / Publisher
[1]	100 total	Not specified	De Department of Dermatology University lowa College of Medicine and Derm Net	2007	Intelligent Sensors, Sensor Networks and Information, 2007. ISSNIP 2007. 3rd International Conference on, 2007, pp. 567–572.
[4]	254 total	Not specified	3A grade hospital in Guangzhou.	2010	Bio-Inspired Computing: Theories and Applications (BIC-TA), 2010 IEEE Fifth International Conference on, 2010, pp. 1033–1036.
[2]	200	200	Not specified	2012	Leonardo Electron. J. Pract. Technol., vol. 20, pp. 147–58.
[5]	254 total	Not specified	Not specified	2012	Expert Syst. Appl., vol. 39, no. 13, pp. 11503–11509, 2012.

[8]	2590	100 validation, 1000 for test	ISIC[6]	2017	IEEE Transactions on Medical Imaging (Volume: 36 , Issue: 9 , Sept. 2017)
[7]	not specified	not specified	ISIC[6]	2018	ArXiv
[10]	2000	150	ISIC[6]	2017	ArXiv
[11]	70011	validation 1001 Testing 2003	ISIC[6]	2018	BiorXiv

Table 4

8. REFERENCES

[1] A. Chiem, A. Al-Jumaily and R. N. Khushaba, "A Novel Hybrid System for Skin Lesion Detection," IEEE, Melbourne, Qld., Australia, 2007.

[2] H. Chen and J. Zhang, "Performance comparison of artificial neural network and logistic regression model for differentiating lung nodules on CT scans," sciencedirect, Expert Syst. Appl., 2012.

[3] M. Sunasra, "Performance Metrics for Classification problems in Machine Learning," 11 11 2017. [Online]. Available: https://medium.com/greyatom/performance-metrics-for-classification-problems-in-machine-learning-part-i-b085d432082b. [Accessed 13 09 2018].

[4] M. S. AL-TARAWNEH, "Lung Cancer Detection Using Image Processing," Academic Direct, 2012.

[5] Z. Jing, L. Bin and T. Lianfang, "Lung nodule classification combining rule-based and SVM," IEEE, Changsha, China, 2010.

[6] "ISIC Archive," [Online]. Available: https://www.isic-archive.com. [Accessed 2018].

[7] C. R. Hardie, "Skin Lesion Segmentation and Classification for ISIC 2018 Using Traditional Classifiers with Hand-Crafted Features," 2018.

[8] Y. Yuan, M. Chao and Y.-C. Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks With Jaccard Distance," *IEEE Transactions on Medical Imaging*, p. 99, 2017.

[9] "Wikipedia," [Online]. Available: https://en.wikipedia.org/wiki/Cross_entropy. [Accessed October 2018].

[10] A. Mahbod, R. Ecker and I. Ellinger, "Skin Lesion Classification Using Hybrid Deep Neural Networks," 2017.

[11] X. Li, J. Wu, H. Jiang, E. Z. Chen, X. Dong and R. Rong, "Skin Lesion Classification Via Combining Deep Learning Features and Clinical Criteria Representations".