Image Classification of Melanoma Skin Cancer using Deep Learning Approach.

MSc Research Project
Data Analytics

Piyush Kishor Dhande
Student ID: x20115725

School of Computing
National College of Ireland

Supervisor: Prof. Noel Cosgrave

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Piyush Kishor Dhande
Student ID:	X20115725
Programme:	Data Analytics
Year:	2021-2022
Module:	MSc Research Project
Supervisor:	Prof. Noel Cosgrave
Submission Due Date:	16/12/2021
Project Title:	Image Classification of Melanoma Skin Cancer using Deep Learning Approach.
Word Count:	4695
Page Count:	15

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	16th December 2021

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Image Classification of Melanoma Skin Cancer using Deep Learning Approach.

Piyush Kishor Dhande
X20115725

Abstract

Skin Cancer is one of the deadliest diseases in the world and, its diagnosis includes clinical screening, examination, and biopsy. It is difficult to diagnose skin cancer in early-stage even for a medical professional. The issue with Melanoma skin cancer is the possibility of spreading it to other body parts.In recent years, various Deep learning techniques took a leap in development of Computer Aided Design (CAD) systems to detect cancer using dermatoscopic images. In this research, the state-of-the-art Capsule Neural Network is used with pre-trained VGG-16 model to detect skin cancer. The results obtained from the VGG16-CapsNet architecture are promising when compared to baseline models with accuracy of 0.992, sensitivity of 0.99 and specificity of 1.

1 Introduction

Several illnesses have burdened the healthcare system in recent years, and skin cancer is one of them, which is on the rise and targets a significant portion of the population. This condition occurs when genetic material gets damaged, leading to the multiplication of cells into lesions and it turns into a malignant melanoma, which is the main reason for skin cancer. The occurrence of skin cancer is influenced by both regional and age-related factors. Cancer treatments need thorough evaluation and supervision by qualified doctors. Skin cancer can spread to other parts of the body if it is malignant or restricted to a specific part, then it is benign. In this study, there is a classification of skin cancer which is defined below:

- **Non-melanoma skin cancer (NMSC):** This type is most common in skin cancer as they usually grow on the skin area exposed to the sun. This type is mainly found in older people. The most occurring NMSC are Basal Cell Carcinoma (BCC) & Squamous Cell Carcinoma (SCC).
- **Melanoma Skin cancer:** This type of cancer is rare, yet it is one of the deadliest forms of skin cancer. It is deadly due to its tendency to spread across other parts of the body, leading to severe issues. Globally, skin cancer is expanding by around 2 million cases (NMSC) and about 2,00,000 melanoma cases each year.

Melanoma Skin cancer This type of cancer is rare, yet it is one of the deadliest forms of skin cancer. It is deadly due to its tendency to spread across other parts of the body,

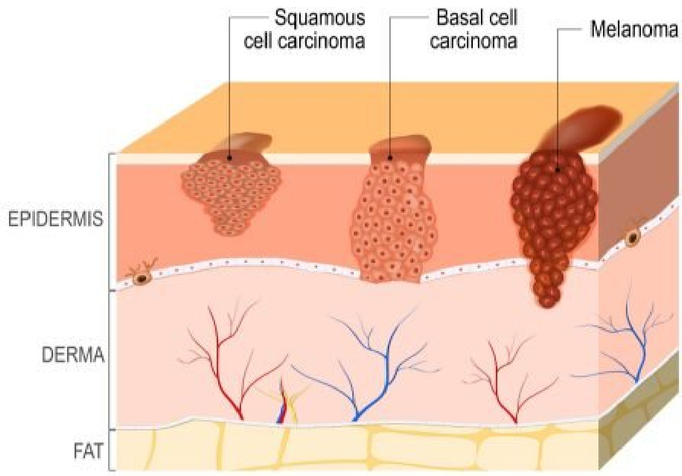


Figure 1: Skin cancer type cancer research uk (n.d.)

leading to severe issues. Globally, skin cancer is expanding by around 2 million cases (NMSC) and about 2,00,000 melanoma cases each year[1].

1.1 Background

Skin cancer is caused by a genetic flaw that permits skin cells to proliferate and reproduce uncontrollably. This is caused primarily due to ultraviolet radiation (UV), a destructive lifestyle, and DNA mutation. According to World Health Organization, one out of three cancer is diagnosed as skin cancer, and early detection increases the chances of the disease not spreading to other parts of the body. Skin lesions could be noticed without the use of clinical devices. One of the best methods to diagnose cancer is histopathology; in this process, a biopsy is carried out to analyze it using a microscope. However, distinguishing between malignant and benign lesions is a complicated task due to the similarity in their appearances. This complication leads to unnecessary biopsies; according to Kondaveeti and Edupuganti (2020), about 10,00,000 biopsies were carried out to diagnose 2000 melanomas. In the present situation, classification takes a long time, affecting the fatality rate.

1.2 Motivation

The primary goal is to reduce deaths caused due to skin cancer. This research aims to use computer vision and deep learning for classification to detect the cancer type from dermatoscopic images, reducing patients' suffering. There is demand for quicker and better solutions to replace outdated systems and reduce the gap between cancer diagnosis and its treatment. Unnecessary biopsies will be avoided if cancer is detected in the early stages by applying the model with improved accuracy and efficiency. This research will use state-of-the-art Capsule Neural networks to classify skin cancer . The

[1] https://www.who.int/

significant difference between CapsNet and convolutional models is the ability of CapsNet to evaluate the spatial relationship in data. Unlike Convolutional Network, Capsule Neural Network stores information in the form of vector instead of using scalar value (Goceri; 2021). In CapsNet dynamic routing is applied for training between layers of the capsule from a low-level capsule to a higher level.

1.3 Research Question

How effective will be the VGG16-Capsule Neural Network classify skin cancer than traditional convolutional networks on dermatoscopic images?

1.4 Research Objective

- Implementation of state-of-the-art Capsule Neural Network with VGG-16 for classification of melanoma.
- Comparison of transfer learning models with VGG16-CapsNet on same dataset to assess the performance.
- Evaluation and comparision of models using Precision, F1 score, Accuracy, Sensitivity, Specificity.

2 Related Work

In the medical field, early detection of disease is essential. For early diagnosis, many researchers are designing tools that could help in the real-time detection of diseases. Methodologies based on deep learning give out better classification results and assist health care professionals in uncovering data's hidden potential, which could lead them to make better medical decisions.In this section, various approaches related to cancer classification are discussed.

2.1 Review of cancer diagnosis methods based on dermatoscopic images utilizing deep learning

Ismail and Sovuthy (2019) used a deep learning approach for the detection of normal and abnormal tumors in the breast. IRMA dataset is used here, where the image is of size 128x128 pixels and has 584 malignant images. Pre-processing was carried out to resize the image to 224x224 to fit into the network's layer. The classification is done using ResNet and VGG-16 and evaluated using accuracy, recall, and precision. Overall, VGG performed better, and the key finding in this is that processing time is affected due to a large number of layers in the architecture, and it could also cause overfitting to some extent. Here CENGİL and ÇINAR (2018) proposed a 3D Convolutional Neural Network (CNN) for lung cancer classification. In this network, 4D tensor is used for the input layer, and there is an additional dimension added named Depth along with height, width, and no. of channels. Dataset used in this research contains 70 CT scan images which are divided into 40 for training and 30 for testing. The 3D CNN model gave an accuracy of 0.70, and it is evaluated using the confusion matrix. In this research, the model's success increased due to the addition of layers in the network. Here performance

can be improved by using a larger dataset and improving the architecture.
Another approach by Aburaed et al. (2020) made use of Deep CNN and VGG-19 for Skin cancer classification. The dataset used was highly imbalanced, to balance it augmentation process was carried out and total image size was 8000. To evaluate the model weighted F1 , accuracy , recall and precision were used and DCNN model outperformed VGG-19.
Li (2021) proposed a classification method to detect breast cancer. The dataset used for this research was ultrasound images of tumors totaling 145, of which 74 were malignant. For image enhancement, pre-processing was carried out, including image median filtering and histogram equalization to enhance the tumor part (ROI). To reduce the noise Butterworth filter was applied and finally performed binarization. Here, two kinds of images were used for classification, coronal and region of interest (ROI). After pre-processing model used in this process is CNN which is evaluated using Area Under the Curve (AUC).In this research, the key finding is that the methodology used has the potential to solve problems on 3-Dimensional data classification.

Chakraborty (2021) in this research, two different approaches are considered, including supervised methods and deep learning methods. The dataset contains images with portable grey map format collected from elastography and ultrasound to detect carcinoma cells in the breast. After this, the noise level was reduced in pre-processing, and contrast was adjusted, temperature calibration was done. Image segmentation was performed to partition it into kinds of fragments. In final stage the implementation of the model in which SVM was able to identify minor lumps along with this random forest, which failed to classify between healthy and infected cells and was not able to detect any kinds of lumps. In the case of the deep learning model, D-CNN outperformed all these models and gave 0.97 accuracies. It comprises 10 million neurons that correctly detect the edge of lumps and their location and nature.
Kaur et al. (2017) propose a method to detect ovarian cancer. The dataset contains cancer images with a scale of 512x512 size. Image was converted to gray scale from RGB. Further in this, edge detection is applied, which can help detect edges with any kind of value and remove distortion was done to reduce image noise. Feature extraction was carried out using SIFT algorithm. After this stage, CNN and SVM modeling with evaluation matrix sensitivity, specificity, and accuracy were used. CNN (0.98) gave better result than SVM (0.85) The issues in past detection systems were that the working accuracy was between 90 to 95 with proper optimization, it could work more efficiently. The key finding in this research is the need for more data for training, and in the future, if fuzzy logic is used, it will give a high response time and less error rate.

2.2 Use of Deep learning in the classification of Skin Cancer

Barata and Marques (2019) proposed architecture using a deep learning approach to perform a diagnosis based on the lesion. In this dataset (ISIC 2017), all types of lesions are considered for classification. Two different transformations are applied lesion segmentation along with colour normalization, and the images were resized to 299x299. Dataset is unbalanced, so augmentation on small classes was performed. The model used for this is CNN with evaluation metrics as Area under the curve (AUC) and Balanced accuracy. The primary outcome of this research is that results are improved by using colour normalization, and this approach should be applied to larger datasets to validate the outcome.
Younis et al. (2019) used a similar dataset and the pre-processing technique as Barata

and Marques (2019). Use of transfer learning was applied in which pre-trained model Mobile-Net was tuned and used on ISIC 2017 dataset. The images used were down-scaled to (224,224) pixels. To balance the dataset, augmentation methods were applied on minority classes and made some changes in the Mobile-Net model by removing the last 5 layers from 93 layers and using the softmax activation function. Metrics used for evaluation were precision, recall and BACC. This central finding in this research was the time taken to classify cancer with reasonable accuracy.

2.3 Initial research of image classification using CapsNet

Ding et al. (2019) proposed a method to overcome problems faced in original Capsule network. This network consists of neurons and instead of using max-pooling use of routing by agreement was done to map lower level capsules to upper level. The dataset used for classification is CIFAR-10/100 Street view house number (SVHN) and augmented data using horizontal flipping. In this proposed method routing by agreement method is modified and this modified version performed faster and cost of computation was less. The primary finding in this research was to prevent the model from overfitting capsule max-pooling should be used. This model performed better than the convolutional model.

Long et al. (2021) proposed a model using CapsNet to classify blood cells and name it as BloodCaps. In this dataset, eight different categories of images were there according to blood cells. In the model, the use of a convolutional layer was done to get a feature map which was further utilized in the primary layer of the capsule. The dataset contains low-resolution images, and the size was also small. This model was compared with Alex-Net and VGG-16, where the proposed model outperformed other models. To check the performance of model, other datasets were also considered, such as cell vision, IDB2 and found this model has the potential to classify better than other models.

A similar approach was used in this research where Panigrahi et al. (2020) used CapsNet to classify Oral squamous cell carcinoma. Cancer could be detected accurately if the main features are extracted in this case; the region of interest is keratin pearl and some difference in the epithelial layer. The image was resized to 256x256 pixels, and due to class imbalance, augmentation was used on minority classes to balance them. To assess the performance of the model, 10 fold cross-validation was used, and it gave an accuracy of 0.97 and to evaluate the performance metrics used were sensitivity, specificity, and F1 score. All of them were good enough, which makes the proposed model has the capability to capture spatial relation and could easily classify between benign and malignant cancer.

In this Iesmantas and Alzbutas (2018) proposed a CapsNet model with some modifications in the primary layer. It is used to classify breast cancer, and the final class is divided into four types. The image size was reduced to 256x256 from (2048x1536). Due to the difference in image shades, Reinhard's method was used on all images, and image rotations were performed randomly. In this model, an additional five convolutional layers were added before the primary capsule. The proposed model was applied, and the cross-validation accuracy was 0.87. The primary finding of this research is model is powerful in classification and performs better than traditional models.

2.4 Conclusion

After reviewing all the research, the deep learning approach will be of great use in the image classification process. Convolutional Networks is viewed as an efficient feature extractor. Yet, it has some disadvantages like the use of pooling layers which makes the loss of spatial relation. A capsule neural network is the network that could overcome this issue is used for the classification process. In the case of classification of disease convolutional network could detect features of the disease and its existence while CapsNet model also detects the ailment and connection between features.

3 Methodology

In any research, there is a structured framework requirement starting from data collection to final results. This section explains how the research has been carried out step-by-step manner. In this research of Skin cancer classification, the methodology used is CRISP-DM (Cross-industry standard process for Data Mining).

3.1 Business Understanding

Cancer is classified as a severe health issue. It could occur in any form; all of it depends on where it originated from. To detect them, various imaging techniques are used such as radiology and oncological imaging. In these present technique, various details are left unattended for cancer classification. These tests require medical professionals to examine, and this process is lengthy. Computer-aided design software has drastically changed the diagnostic process to overcome this tedious process. All these systems need huge investments. Researchers are trying to build a cost-effective method to help classify cancer in a less-time period.

3.2 Data Understanding

The dermatoscopic image is collected from International Skin Imaging Collaboration (ISIC) archive[2], and the dataset is publicly available. The ISIC 2020 data set contains a total of 33,126 training images of the benign and malignant lesion, where the benign class has 32,542 images and the malignant class has 584 images with image of size 6000x4000 pixels. To understand this data further in-depth, we carried out Exploratory data analysis

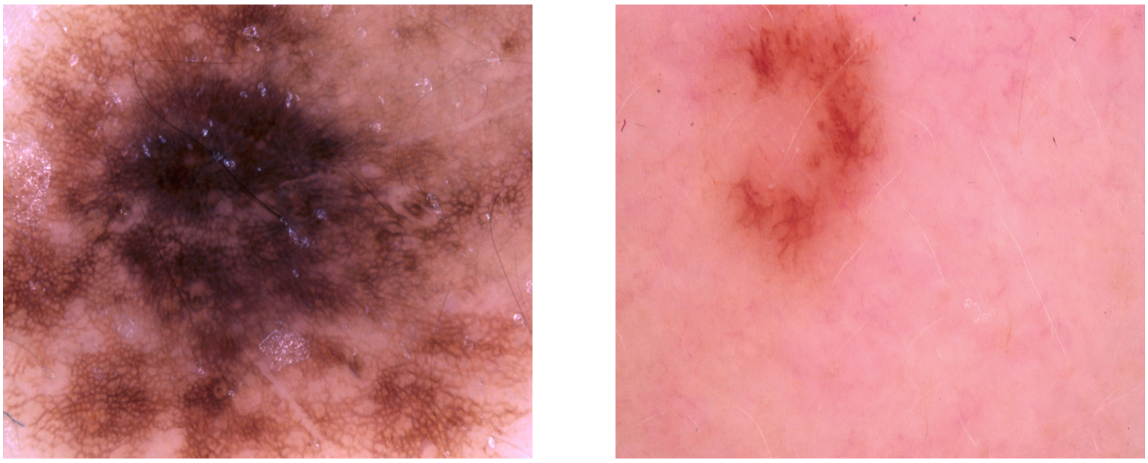
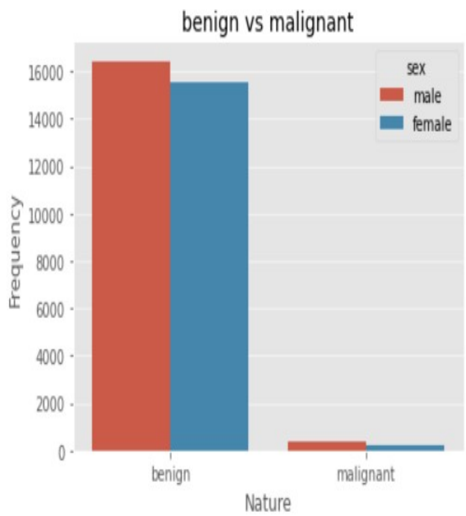


Figure 2: Malignant Lesion(Left) & Benign Lesion(Right)

[2]Dataset: https://challenge2020.isic-archive.com/

of all the data gathered, including the metadata. The research goal is binary classification and the images should be mapped to metadata labels. In the EDA of ISIC, the dataset found the number of cases registered against the possibility of having cancer (benign lesion = 0, malignant lesion = 1) gender-wise. From Figure 3 (left), it is visible that males are more susceptible to cancer. In Figure 3 (right) are the regions that are affected



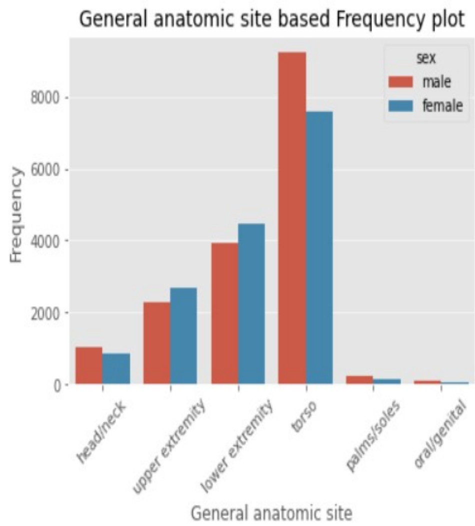


Figure 3: Benign & Malignant count w.r.t. Gender(Left) & Cancer affected parts w.r.t. Gender(Right)

by cancer gender-wise in this female are more susceptible to cancer on the lower and upper extremity while the torso is highly affected area as compared to other areas. The lesser affected areas are palm/soles & oral/genital as compared to the other areas of body.

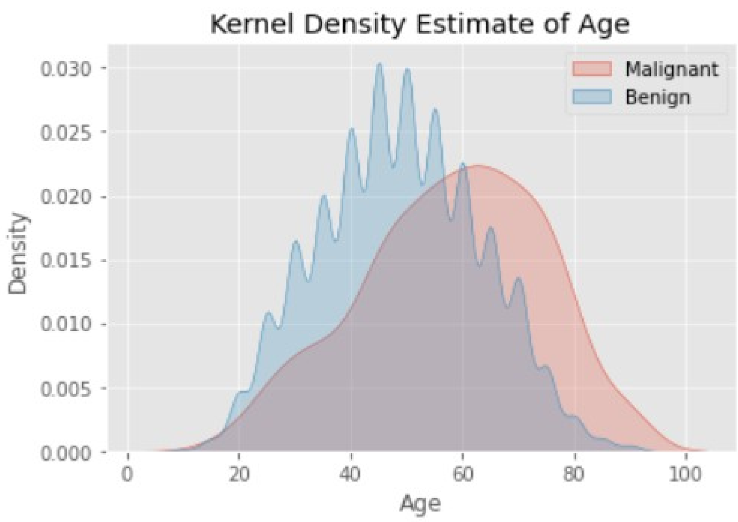


Figure 4: Cancer type w.r.t. age

The primary risk factor of any disease starts with ageing; according to Figure 4 the age ranging from 45 to 80 are diagnosed with melanoma cancer. In the case of benign type, the majority of people come in the range of 40 to 75. This dataset is highly imbalanced, and it was balanced by using augmentation methods on minority classes which is malignant.

3.3 Data Preparation

Before the training phase, the raw data needs to be cleaned and transformed properly, which will make the model's training to be successful. Before training, the data should be adequate and balanced. To make the imbalanced data set balanced, the use of image augmentation is done. Total images in the dataset are 33,126, out of which the malignant class has 584 images. Image augmentation techniques were applied to the malignant class to balance both classes. From keras library the ImageDataGenerator[3] is used for augmentation, and the benefit of using this class is at each epoch is that the model receives a new set of images different from the original set of images and requires low memory for processing, as the images are loaded in batches. The augmentation techniques used are rotation, width shift, height shift, shear range, horizontal vertical flip.After augmentation, the dataset is split into the train (9676 images), validation (2628 images), test(1942 image).

3.4 Modeling

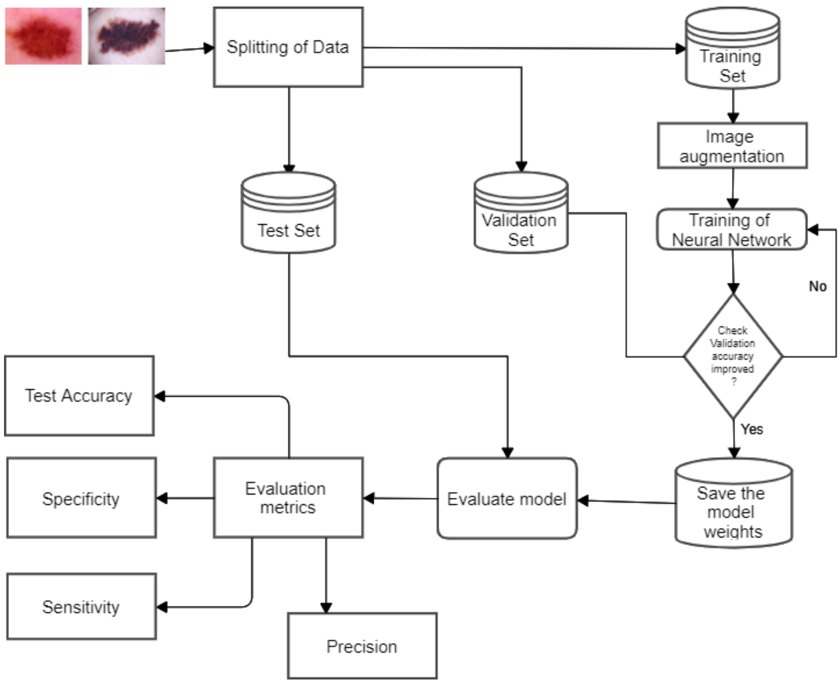


Figure 5: Process-flow diagram

This section contains the model workflow, and explanation of technical processes sequence-wise as illustrated in Figure 5.

1. The downloaded images are unzipped to apply exploratory data analysis techniques so that the data can be studied thoroughly.

[3] `https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/image/ImageDataGenerator`

2. The data was found to be highly imbalanced in our study. Before performing augmentation on the dataset it was split into train, validation and test set.

3. After segregating the data, augmentation techniques are applied on the minority class in the train set.

4. During the training of the model, training and validation datasets are utilized. We use the accuracy metric to evaluate model performance . During the training phase, when validation loss does not improve, the next training epoch terminates. The model weights with negligible loss value are saved.

5. After the training phase, model performance is evaluated using the test dataset and various evaluation metrics are calculated.

3.5 Comparison & Evaluation

Transfer learning technique is one of the best techniques in deep learning. These models are trained on huge datasets and contains information that can be used for similar kind of tasks. The main advantage of using a transfer learning technique is it reduces the training time and improves image feature extraction. In this research, VGG-16 and ResNet50[4] trained on ImageNet will be used for comparision with our proposed model. For the purpose of the evaluation of our model, we use several metrics and those are Accuracy, Loss, Sensitivity, Specificity, Precision, and F1 score (Jaiswal; 2021). We use these metrics because accuracy alone would not be enough as our data is highly imbalanced. The goal in such models in disease detection should be too achieve minimum false negatives, thus the other metrics will help us to get a clearer picture of the model performance.

4 Design Specification

4.1 VGG16-CapsNet

Capsule neural networks are the type of deep learning network that perform computation in their input, and the results are in the form of informative vectors. The proposed method is capable of identifying the texture and orientation of features. Capsule Networks can detect the presence of features even if the viewing angle is changed. The combination of a convolutional layer with a Capsule neural network makes a new architecture Conv-Caps where convolutional layers are added on top of the capsule network. The skin cancer images are passed into a convolutional layer first, and it is integrated via max-pooling to scale the sample space and capture features. The captured feature is given as an input to the primary capsule, and the output comes in vector form. This feature information is passed from low-level capsule to upper-level capsule, i.e., the input is sent from the low-level capsule, and all of it is possible due to the dynamic routing algorithm Sabour et al. (2017).The architecture used for this research is based on Palaniappan et al. (2017) The network proposed in this research comprises of VGG-16 and CapsNet. The VGG-16 model will help find robust features, and the relation between features will be detected using CapsNet. The extraction of high-level features will be done using the VGG network. The reason behind the implementation is the model's simplicity and robustness

[4] `https://keras.io/api/applications/resnet/`

and VGG-16 consistency in outperforming different base-line models; it is one of the most used models in image recognition.

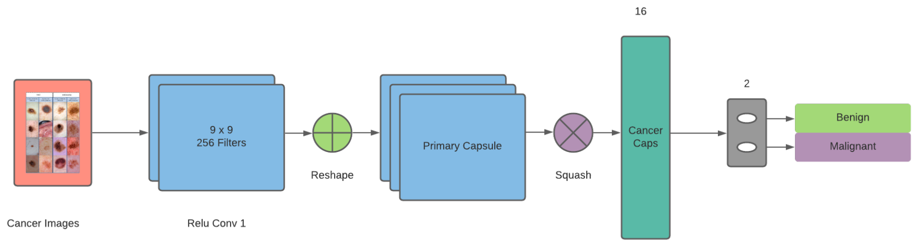


Figure 6: VGG16-CapsNet architecture

1. The VGG-16[5] model is trained using Imagenet[6] weights that help in accelerating the training process.

2. The first experiment involves freezing all the layers of VGG16 so that the weights do not change during training. We add three more fully connected layers, which are actually responsible for fine-tuning the feature output from the VGG16 network.

3. We enable the last three layers of VGG16, which in turn provides six trainable layers in our second approach. This arrangement is trained till ten epochs, after which we found no significant change in validation loss. This technique fine-tuned the weights of the last three layers of VGG16 to better identify features related to Melanoma images.

4. We then extract the output of the trained VGG16 from our second experiment and feed it to the CapsNet network, followed by a final output layer.

Description of CapsNet architecture from Figure 6 is explained below Bharati et al. (2020):

1. The extraction of feature space is done on images using max pooling layer on applying sampling technique in convolutional layer.

2. The image pushed in convolutional layer that is of size 299x299.After that output is reshaped and converted 3-Dimensional output into 1D array.

3. The second layer in architecture is layer of primary capsule which consist of 32 capsule of and dimension is of size 8.

4. In the third layer we have a secondary capsule. Here the number of capsules should be equivalent to the prediction of final class in this case we have two (Malignant,Benign)

5. The fourth layer is generated due to output of the previous layer and to calculate vector size implementation of squash softmax function is used.

6. Classification of images are based on marginal loss and it should always be minimum for actual classification.

[5] `https://keras.io/api/applications/vgg/`
[6] `https://image-net.org/`

5 Implementation

5.1 Implementation of VGG16-CapsNet

This section gives an overview of step by step execution of our model described in subsection 4. The model training and implementation was done on Kaggle notebooks which gives dedicated GPU usage. The dataset was stored on kaggle's cloud storage is split into train, validation and test set ,i.e, 60% for train set, 20% for test set and 20% of data as validation set. The size and length of the dataset is mentioned in subsection 3.2. A callback function for early-stopping is implemented so that training stops when there is no improvement in validation loss. Considering the patience value, if there is no change in loss value, training stops and the improved weights are stored. To balance our data, we include synthetic data by using ImageDataGenerator, which creates augmented images by varying parameters like zoom, shear and rotation. The batch size is set at 64. The batch size, and the datasets (i.e., train, validation, test) remain same across all experiments.

5.2 Implementation of ResNet-50

ResNet50 is a 50-layer Convolutional neural network consisting of pre-trained weights from ImageNet (Peng et al.; 2019). It provides the functionality to use these pre-trained weights for training the ResNet50 model. All the required libraries and functions were imported along with Keras and TensorFlow at the first step. Input shape is set to image size (128,128,3), where 3 is for RGB value. Further, the output of the Resnet50 is flattened at last layer and the output is in binary form and the activation function used is SoftMax. In this process the optimization function used is Adam. Checkpoints and early stopping were also defined with a patience level of 5 epochs. An exponential decay function was developed to give variable inputs to the learning rate functions. Finally, the model was fitted on training, validation sets is used to check the performance of trained model. The best performing model is saved in .h5 format. In the testing phase, the saved model is loaded and test generator is applied to convert image into (128, 128) pixel.

5.3 Implementation of VGG-16

VGG-16 is a 16-layer Convolutional neural network which is known for its robustness and simplicity VGG-16 is pre-trained on ImageNet. This model gives the ability to train the VGG16 model with these pre-trained weight.This network contains Convolutional layer of 3x3 filter and the strides used is 1. In this process same padding is used with max pooling layer of 2x2 filter. Input shape is set to image size (224,224,3). Here the output of VGG-16 is flattened at last layer and output is generated in binary form and the activation function used is SoftMax. The optimization function used in this process is Adam. Considering these configuration the model is trained on training set and validated on validation set. Checkpoints and early stopping were defined with the patience level of 5. The best performing model is saved as vgg_16.h5 format.

6 Evaluation

In this the performance of the model is validated on the basis of various evaluation metrics. The metrics used are Accuracy, Loss, Sensitivity, Specificity, Precision, Recall, and F1 score (Jaiswal; 2021).

6.1 Evaluation of VGG16-CapsNet model

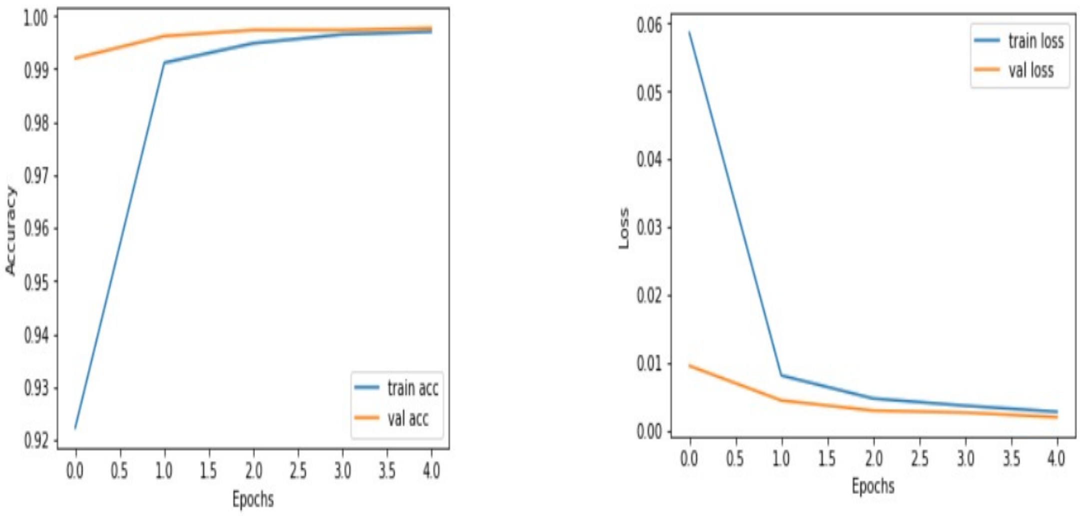


Figure 7: Learning curve of VGG16-CapsNet

Table 1 shows the evaluation metrics compared against the base-line models. The specificity and sensitivity of the proposed model is better than baseline models. The specificity and sensitivity value should be close to 1 meaning that the model will perform better than baseline models.In case of precision the VGG16-CapsNet model performs better meaning it shows VGG16-CapsNet is more likely to predict postive case than baseline models. In case of F1-score our model VGG16-CapsNet perform better than baseline models as it should be closer to 1.

Model	Accuracy	Loss	F1 Score	Precision	Specificity	Sensitivity
VGG16-CapsNet	0.99	0.0019	0.95	0.95	1	0.99
VGG-16	0.98	0.053	0.69	0.79	0.16	0.79
ResNet-50	0.93	0.21	0.71	0.80	0.22	0.80

Table 1: Comparision of evaluation metrics

In deep learning models learning curve is a source to check the model's underfitting or overfitting. From Figure 7 we could check the accuracy curve and loss curve running parallel till last epoch and there is no evidence for overfitting.

	Benign (Predicted)	Malignant (Predicted)
Benign (Actual)	TN=341	FN=0
Malignant (Actual)	FP=3	TP=1598

Table 2: VGG16-CapsNet confusion matrix

The Table 2 shows the confusion matrix of VGG16-CapsNet model where it is tested on 1942 test images with prediction of melanoma skin cancer as malignant or benign.Only 3 times it was predicted as False positive other-wise proposed model performed well in terms of evaluation metrics , accuracy and loss.

6.2 Discussion

In this research three models are applied and primary goal of our research is to assess the performance of model VGG16-CapsNet to other models applied for Classification of Skin Cancer. VGG-16 and ResNet-50 models were applied by Ding et al. (2019) for Melanoma classification and got better results in terms of test accuracy,sensitivity and F1 score. The detailed comparision is shown in Table 1. The best performance of classification with 99 percent of accuracy was achieved with newly designed model and the sensitivity is 99% with minimal loss value and 95% F1 score. After this model VGG-16 achieved second highest accuracy with 98 percent accuracy but in terms of sensitivity and specificity VGG 16 still lags behind ResNet 50.

7 Conclusion and Future Work

Skin cancers are one of those widespread diseases that turn out to be endangered diseases for the human race. Many people are getting diagnosed with various types of skin cancer. As a result of these, there is a need for carrying out research based on skin cancer. The objective of our research was to classify and identify Melanoma skin cancer using histopathological images at an early stage. A deep learning approach was used to reach the goals of the research. After gathering the image data, the exploratory data analysis was carried out to study the insights of the data. The image augmentation process was carried out on the training set. VGG16-CapsNet, VGG-16, ResNet-50 are some of the deep learning models implemented in the research. To evaluate the result of this model's performance, techniques like Accuracy, F1 Score, Precision, Specificity, Sensitivity were used. VGG-16, ResNet-50 are used as baseline models in the research. The proposed model, i.e., VGG16-CapsNet, has shown better accuracy in melanoma classification than the other two base models. Values of other evaluation parameters like precision, F1 score are also close to one, supporting the results obtained from the accuracy of VGG16-CapsNet. Overall, VGG16-CapsNet has helped the research achieve the research objective. For future work, we intend to use CapsNet with other CNN models to check the performance for new hybrid model. In the present network we have scope to increase the number of capsule layer.

References

Aburaed, N., Panthakkan, A., Al-Saad, M., Amin, S. A. and Mansoor, W. (2020). Deep convolutional neural network (dcnn) for skin cancer classification, pp. 1–4.

Barata, C. and Marques, J. S. (2019). Deep learning for skin cancer diagnosis with hierarchical architectures, pp. 841–845.

Bharati, S., Podder, P. and Mondal, M. R. H. (2020). Hybrid deep learning for detecting

lung diseases from x-ray images, *Informatics in Medicine Unlocked* **20**: 100391.
URL: *https://www.sciencedirect.com/science/article/pii/S2352914820300290*

cancer research uk (n.d.).

CENGİL, E. and ÇINAR, A. (2018). A deep learning based approach to lung cancer identification, pp. 1–5.

Chakraborty, D. (2021). A deep convolutional neural network based framework for breast cancer detection.

Ding, X., Wang, N., Gao, X., Li, J. and Wang, X. (2019). Group reconstruction and max-pooling residual capsule network, pp. 2237–2243.
URL: *https://doi.org/10.24963/ijcai.2019/310*

Goceri, E. (2021). Analysis of capsule neural networks for image classification.

Iesmantas, T. and Alzbutas, R. (2018). Convolutional capsule network for classification of breast cancer histology images.

Ismail, N. S. and Sovuthy, C. (2019). Breast cancer detection based on deep learning technique, pp. 89–92.

Jaiswal, G. (2021). Performance analysis of incremental learning strategy in image classification, pp. 427–432.

Kaur, B., Mann, K. S. and Grewal, M. K. (2017). Ovarian cancer stage based detection on convolutional neural network, pp. 855–859.

Kondaveeti, H. K. and Edupuganti, P. (2020). Skin cancer classification using transfer learning, pp. 1–4.

Li, M. (2021). Research on the detection method of breast cancer deep convolutional neural network based on computer aid, pp. 536–540.

Long, F., Peng, J.-J., Song, W., Xia, X. and Sang, J. (2021). Bloodcaps: A capsule network based model for the multiclassification of human peripheral blood cells, *Computer Methods and Programs in Biomedicine* **202**: 105972.
URL: *https://www.sciencedirect.com/science/article/pii/S016926072100047X*

Palaniappan, M., Aravindan, C., Gokul, V. and Sakthi, S. (2017). Deep learning for skin lesion classification.

Panigrahi, S., Das, J. and Swarnkar, T. (2020). Capsule network based analysis of histopathological images of oral squamous cell carcinoma, *Journal of King Saud University - Computer and Information Sciences* .
URL: *https://www.sciencedirect.com/science/article/pii/S1319157820305280*

Peng, J., Kang, S., Ning, Z., Deng, H., Shen, J., Xu, Y., Zhang, J., Zhao, W., Li, X., Gong, W., Huang, J. and Liu, L. (2019). Residual convolutional neural network for predicting response of transarterial chemoembolization in hepatocellular carcinoma from ct imaging, *European Radiology* **30**: 1–12.

Sabour, S., Frosst, N. and Hinton, G. E. (2017). Dynamic routing between capsules.

Younis, H., Bhatti, M. H. and Azeem, M. (2019). Classification of skin cancer dermoscopy images using transfer learning, pp. 1–4.