Enhanced Skin Cancer Classification using Pre-trained CNN Models and Transfer Learning:

A Clinical Decision Support System for Dermatologists

Team 1:

Camellia Marwan

Mahmoud Mohamed

Farah Ossama

Marwan Osama

Reem Adel

Under the supervision of: Dr. Eman Ayman

Table of contents:

Introduction:

Skin cancer is one of the most common and wide-spread cancers worldwide. It significantly affects the quality of the lives of many. The most frequent cause is the over exposure of skin to ultraviolet radiations coming from the sun. The rate of being intensely affected when exposed to UV radiations is higher and more risky in fair skinned, who are more sun-sensitive people than in dark skinned, who are less sun-sensitive people. Melanoma, which is a dangerous type of skin cancer that affects the melanocytes (squamous cell layer), accounts for 1% of all skin malignancies, with the other 99% being basal cell carcinoma or squamous cell carcinoma. Melanoma has become the most severe and ferocious skin cancer and is responsible for around 75% of all skin cancer mortality.

The best method to control and tame skin cancer is its early detection and prevention . Awareness and gaining knowledge of new or changing skin spots or growths, particularly those that look extraordinary, should be examined carefully. Any new lesions, or progressive change in a lesion's appearance (size, shape, or color), should be evaluated by a clinician.

Doctors use plenty of techniques and procedures to detect skin cancer. An expert dermatologist usually starts with naked-eye recognition of suspicious tumours, then proceeds to dermoscopy, and finally a biopsy . As observed the correct diagnosis in that case is highly dependent on the clinician's abilities and there is a big room for error in the manual identification of skin diseases. This is where the urgent need for computer-assisted diagnosis became very clear and obvious at that point as it helps the medical experts in analyzing the dermoscopy procedures in case of a lack of expertise in the diagnostic process and lack of availability of a professional.

That is when innovation took place with the usage of artificial intelligence and deep learning in medical diagnostics. In this day and age, with these new technologies capabilities the efficiency of predicting a result increases exponentially as compared to the dependency on a visual diagnostic. The convolutional neural network (CNN) is a very crucial artificial intelligence algorithm in feature selection and object classification. Deep convolutional neural networks aid tremendously in classifying skin lesions into seven diagnostic categories, namely melanocytic nevi, melanoma, benign keratosis-like lesions, basal cell carcinoma, actinic keratosis, vascular lesions, and dermatofibroma, with the help of their dermoscopic images, covering all the lesions found in skin cancer identification. Deep learning algorithms backed by high performing GPUs in computing large datasets have shown better performance than humans in skin cancer identification.

The dataset used for this paper work is based on the HAM10000 dataset which consists of 10015 images

This dataset used for classification was not so recent and not sufficient enough to identify and analyze all types of lesions . By keeping all this in mind, three objectives were demanded:

- To classify the images from HAM10000 dataset into seven different types of skin cancer.
- To use transfer learning nets for feature selection and classification so as to identify all types of lesions found in skin cancer.
- · To properly balance the dataset using replication on only training data and perform a detailed analysis using different transfer learning models.

All the three transfer learning nets used were compared, and their training and validation loss, training and validation accuracy, along with their individual confusion matrices, were plotted. A comparative analysis of accuracy was then performed for all these learning nets and concluded with the model, which gave the highest accuracy in identifying all the lesions.

Literature Review:

- Various techniques have been introduced before for melanoma classification in the previous few decennaries. Like method [1]. Where DCNN and transfer learning methods are used for the classification between benign and malignant skin cells. The data required for such model should be large amount of data. And testing accuracy was 90.16%.

- Method [2] model was pre-trained with the ImageNet dataset, and afterwards used various architectures for the fine-tuning of the model Hence utilizing the low-resolution highly imbalanced, grayscale HAM10000 skin cancer dataset into several pre-trained network architectures and its accuracy for testing of the first and second steps were 85% and 75%.

- Method [3] used a pre-trained architecture named ResNet50 CNN to classify the skin lesion as melanoma or nevi. The proposed model achieved 77.9% and 82.3% ratios for sensitivity and specificity, respectively. The classifier's performance was established on a test-set disjunct from the training and validation set.

- Method [4] the study, part of the 2017 ISBI challenge hosted by ISIC, explores various pre-trained network architectures for fine-tuning on the HAM10000 skin cancer dataset. The dataset's characteristics, including low resolution and high imbalance, pose significant challenges. The authors achieved a testing accuracy of 85% and 75% in the first and second steps, respectively. The research sheds light on the complexities of real-world datasets in dermatology and underscores the importance of adapting pre-trained models to address the nuances of skin cancer detection.

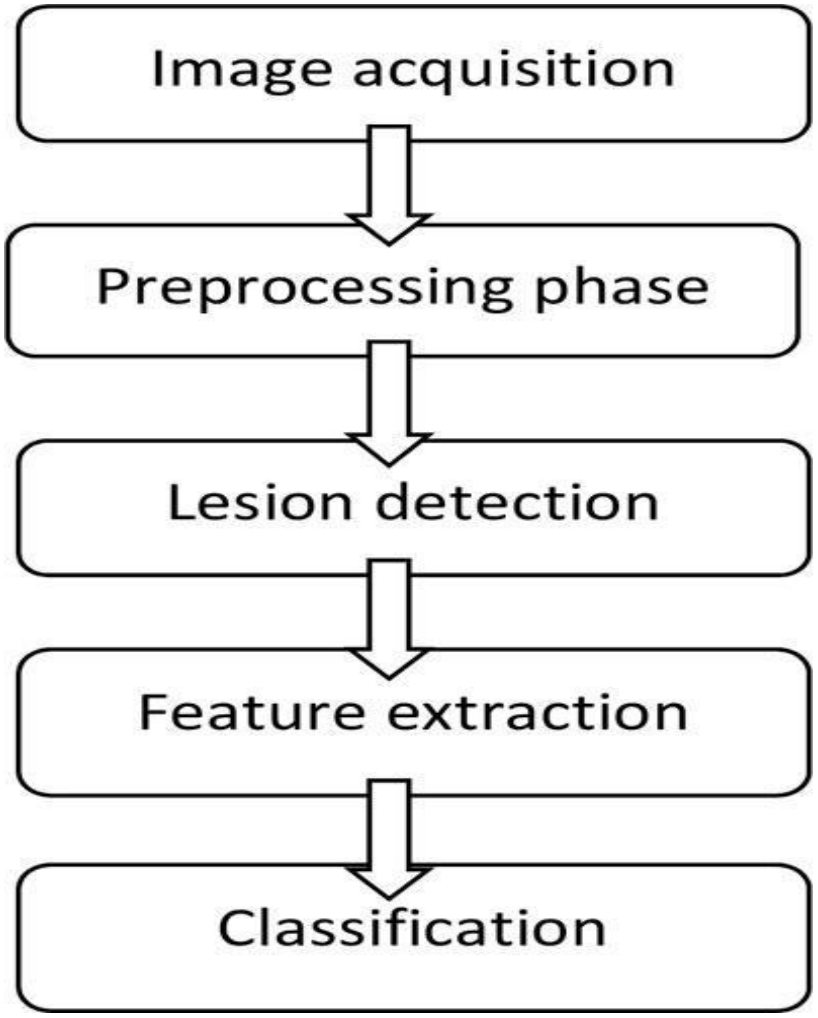
Challenges Of Skin cancer Detection:

- Skin cancer detection poses significant challenges, as traditional methods, such as skin self-examination and clinical evaluation, are resource-intensive and time-consuming, requiring the expertise of a professional physician. Alternatively, AI-based diagnostic tools for skin cancer leverage both shallow and deep AI methodologies. These approaches involve customizing computer algorithms through a training process to learn from data characterized by predefined features. Our chosen approach, specifically transfer learning, aligns with deep AI methodologies, utilizing a large dataset and multiple layers with machine learning techniques.

- Despite the promise of AI, a notable challenge in skin cancer detection is the scarcity of high-quality medical imaging datasets for training. This scarcity is often attributed to the lack of annotated images for abnormal classes. Addressing the imbalanced nature of the dataset, we implemented a strategy involving the adjustment of weights assigned to each class. Specifically, higher weights were assigned to classes with fewer data instances, thereby rectifying the imbalance in class distribution. Additionally, the dataset presented redundancy and duplicates, necessitating thorough pre-processing to ensure data readiness for analysis. These considerations underscore the complexities and nuances involved in developing effective AI solutions for skin cancer detection.

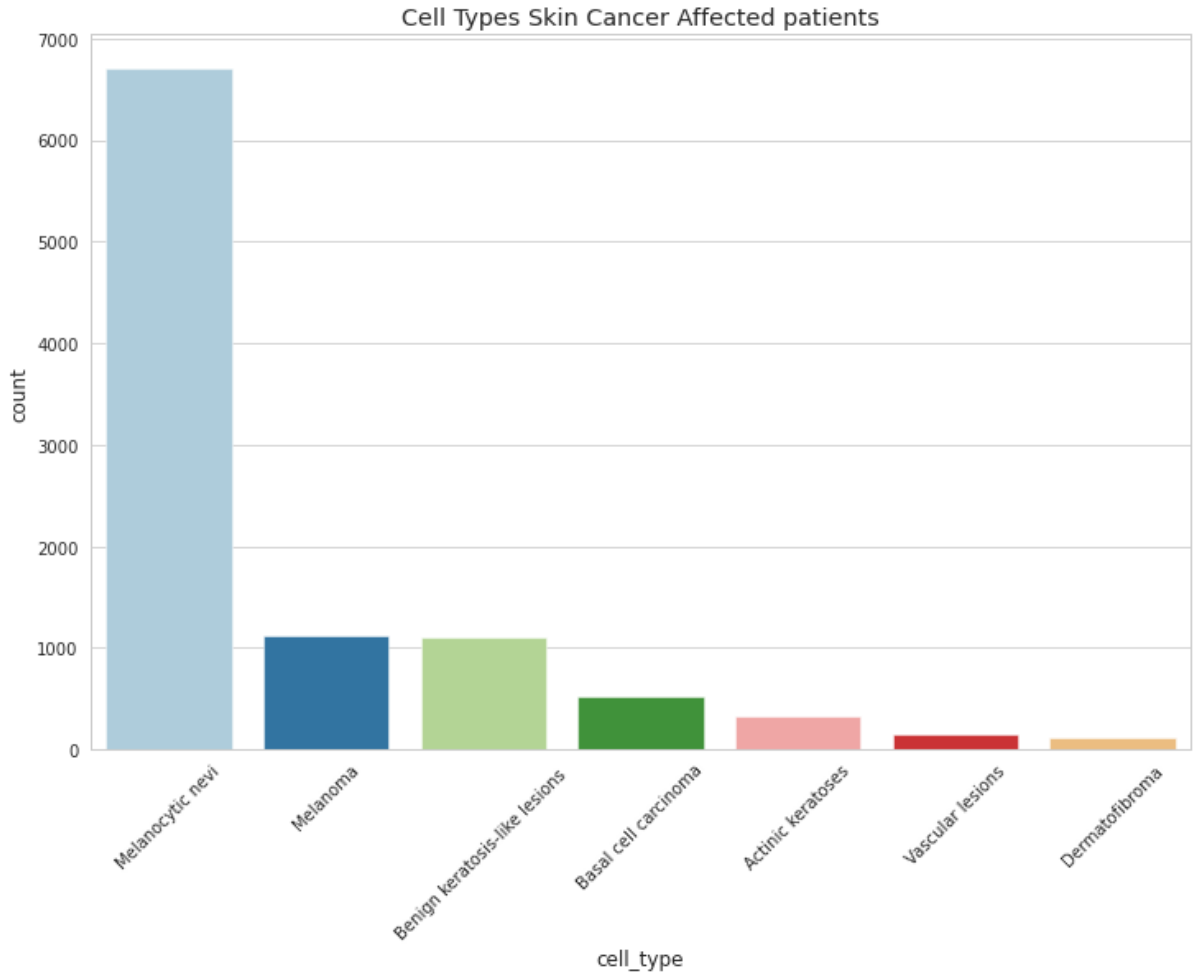
Material and Methods:

In order to better understand our suggested skin tumour classification, this part covers the pre-processing image pipeline. These are the general steps in our methodology: The initial stage, gathering dermoscopic images, then data preprocessing which includes resizing, balanced sampling, and augmentation can be applied which solves for the issue of imbalanced data. Following this is lesion detection, feature extraction and classification After that, utilizing the CNN model along with pre-trained models such as VGG16, DenseNet201, MobileNetV2 and ResNet50 with modifications in model architectures, using transfer learning to train the HAM10000 dataset.



With a resolution of 600 x 450 pixels in the the dataset, it includes 10015 dermatoscopic images representing seven different categories of skin cancer. The HAM10000 training dataset contains the following seven classes: Actinic keratosis (AKIEC) is a benign condition that can progress to a malignant tumor. Melanoma (MEL) and basal cell carcinoma (BCC) are examples of malignant malignancies. Benign Keratosis, Melanocytic Nevi, vascular lesions, and Dermatofibroma are examples of benign lesions. Three portions of this dataset

were split into: training set, validation set and testing set. The distribution of images in each class is shown With 6705 photos in the largest class, NV, and just 115 in the smallest, DF, the dataset exhibits a severe class imbalance.



Data Pre-processing:

- Data Resizing:

The HAM10000 dataset includes the images in 600 × 450 dimensions. The dataset has been resized and scaled down to 100 × 100. The model's complexity and processing time as a result will be significantly reduced.

- Data Augmentation:

The process of adding slightly changed copies of existing data without actually gathering new data from training sets is known as data augmentation.We applied different augmentation techniques, including rotation with a range of 0 to 10 degrees, zooming, and horizontal and vertical flipping.

- Imbalance Handling:

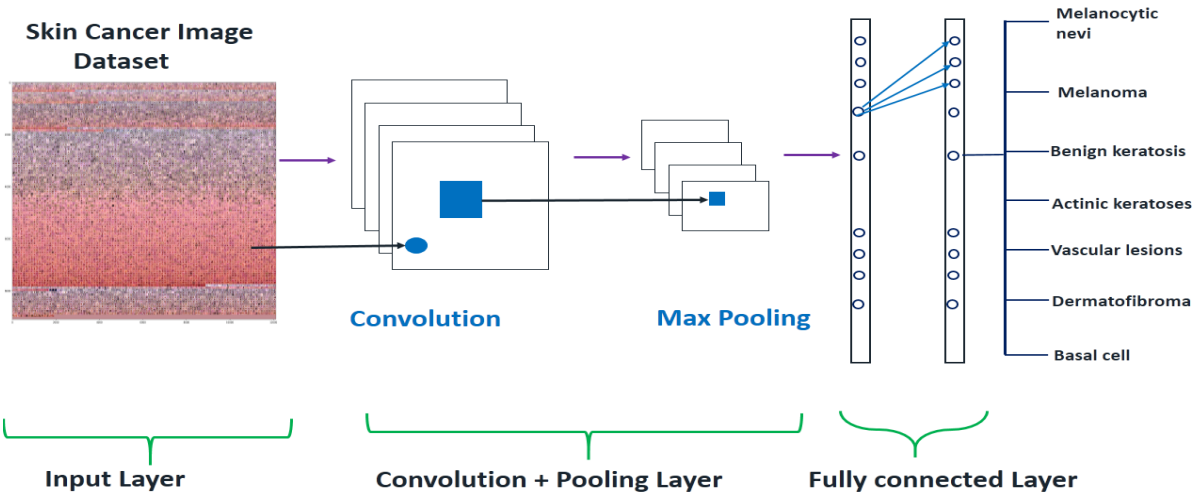
The HAM10000 dataset has a significantly unbalanced class distribution. If this imbalance is not addressed, the model may exhibit bias towards the majority class and underperform on the minority classes. Problems with imbalanced data are common in deep learning. An efficient way to reduce the impact of this problem during training and prioritize the minority classes was to utilize a weighted loss function.Based on the quantity of samples in each class, class weights can be computed and utilized to modify the model's loss function. The majority class's weight is reduced while the minority class's misclassification is penalized by an increase in class weight, which forces the model to give these classes more weight. While for the second

Transfer Learning:

Transfer learning is the technique of applying prior information from models trained on the ImageNet dataset to improve learning in a new task. The main goal of using transfer learning is to save time and money by eliminating the need to train numerous models from beginning for tasks that are similar. Furthermore, by using pretrained models, transfer learning can assist in overcoming the problem of a lack of labeled training data .This study classified skin cancers into seven classifications using four pre-trained models: DenseNet201, MobileNetV2, ResNet50 and VGG16. The HAM10000 dataset was used for this purpose. To attain excellent performance, both transfer learning and fine-tuning were used.

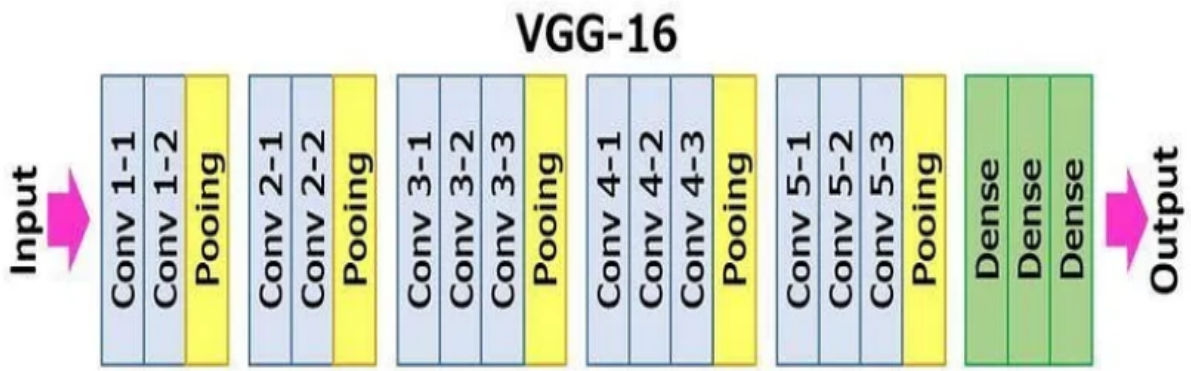
First and second CNN models:

A sequence of convolutional layers consists of filters of varying sizes—96 filters of size 11x11 with a stride of 4 in the initial layer, followed by 256 filters of size 5x5, and subsequently 384 filters of size 3x3. These layers utilize rectified linear unit (ReLU) activation functions followed by batch normalization to enhance model stability max pooling layers to down-sample feature maps, facilitating spatial summarization while preserving essential information.



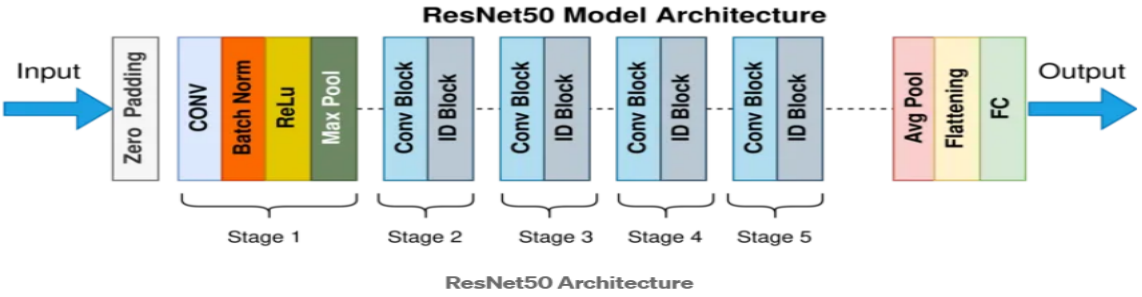
VGG16:

This network architecture comprises 16 Convolutional layers along with 3 fully connected layers. It uses the ReLU activation function and on the final layers, it applies softmax as the classifier.



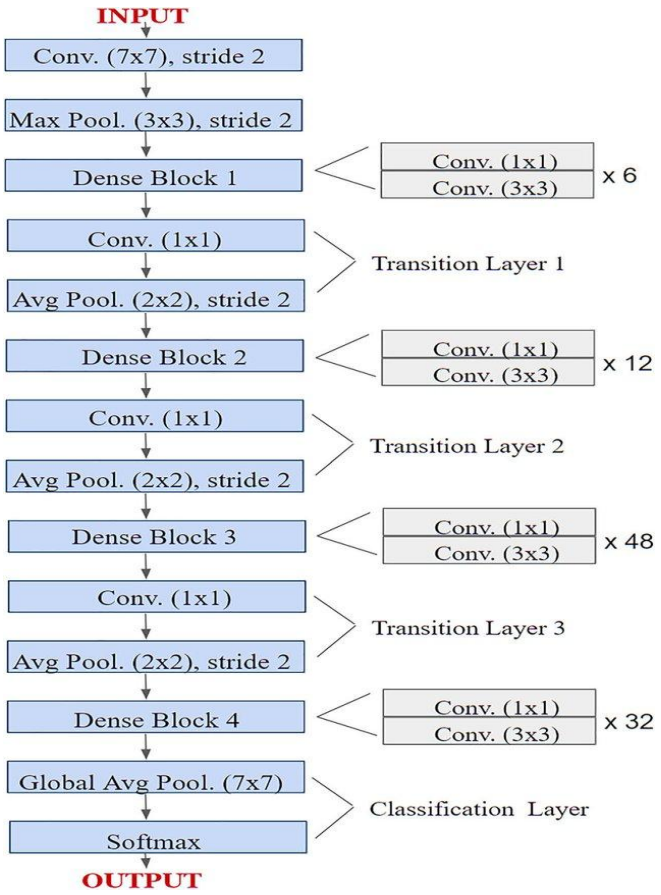
ResNet50:

The convolutional layers in ResNet50 consist of several convolutional layers followed by batch normalization and ReLU activation. These layers are in charge of taking features like edges, textures, and shapes out of the input image. Max pooling layers, which decrease the spatial dimensions of the feature maps while maintaining the most significant features, come after the convolutional layers.



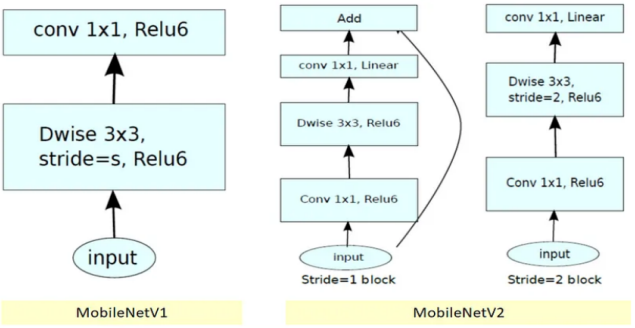
DenseNet201:

DenseNet201 employs a sequence of convolutional layers enriched by batch normalization and Rectified Linear Unit (ReLU) activation. Followed by max pooling layers to down-sample feature maps.



MobileNetV2:

a design for a convolutional neural network.The first fully convolution layer with 32 filters makes up the entirety of MobileNetV2's architecture. followed by batch normalization and ReLU activation

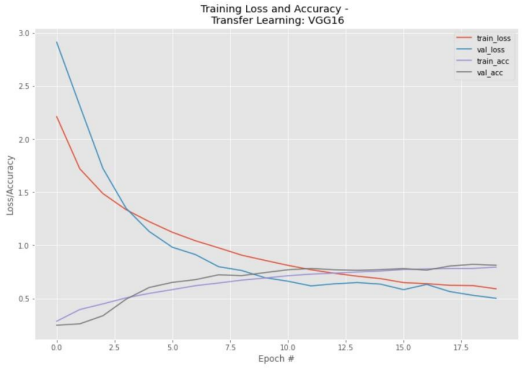


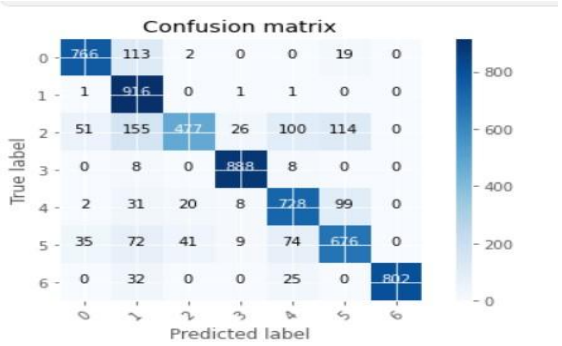
Results & Discussion:

VGG16:

First VGG model confusion matrix, and accuracy test.

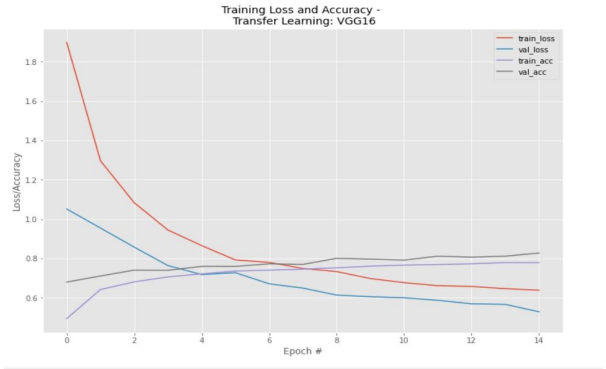
stages	Balanced	Epochs	hidden layers	data augmentation	batch size	accuracy	validation accuracy	loss	validation loss
1	True	20	512, 128 256, 256,64	True	1000	0.7938	0.8119	0.5899	0.5004

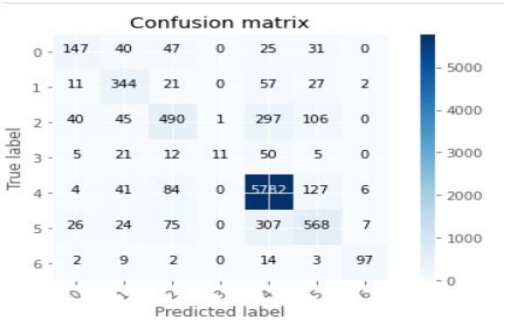




Second VGG model confusion matrix and accuracy test:

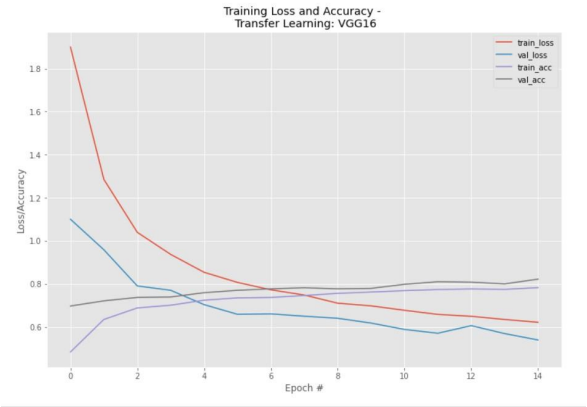
stages	Balanced	Epochs	hidden layers	data augmentation	batch size	accuracy	validation accuracy	loss	validation loss
1	False	15	512, 128	True	100	0.7786	0.8273	0.6383	0.5286

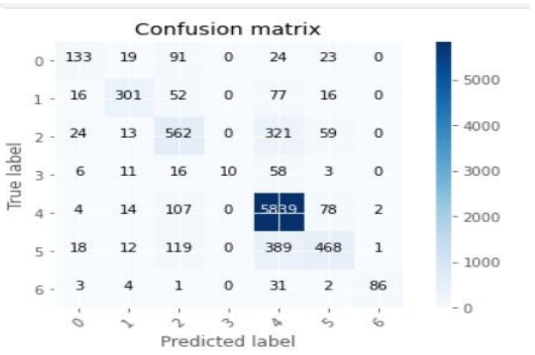




Third VGG model confusion matrix and accuracy test.

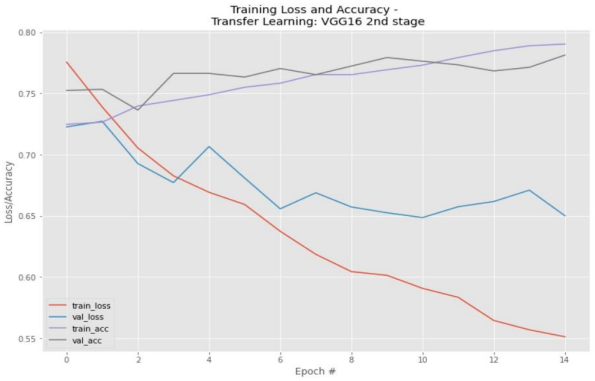
stages	Balanced	Epochs	hidden layers	data augmentation	batch size	accuracy	validation accuracy	loss	validation loss
1	False	15	512, 128	False	100	0.7831	0.8224	0.6221	0.5400

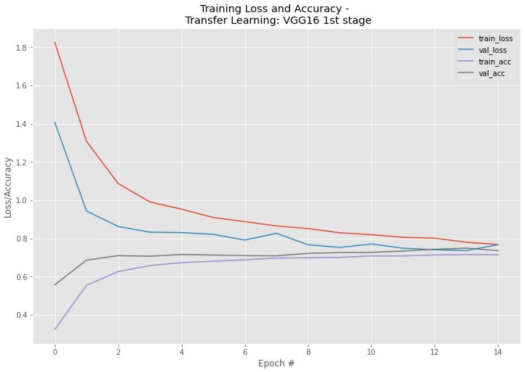


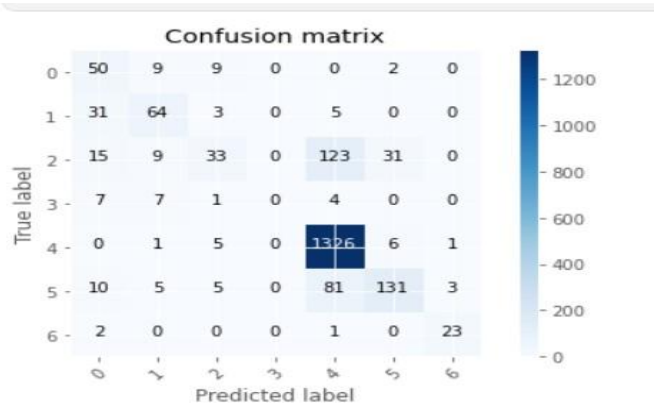


Fourth VGG model confusion matrix and accuracy test

stages	Balanced	Epoch	hidden layers	data augmentation	batch size	accuracy	validation accuracy	loss	validation loss
2	False	15	[256,256,64]	False	100	0.7904	0.7814	0.5512	0.6501

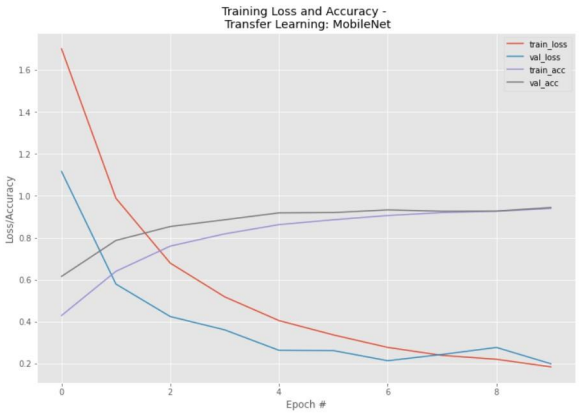


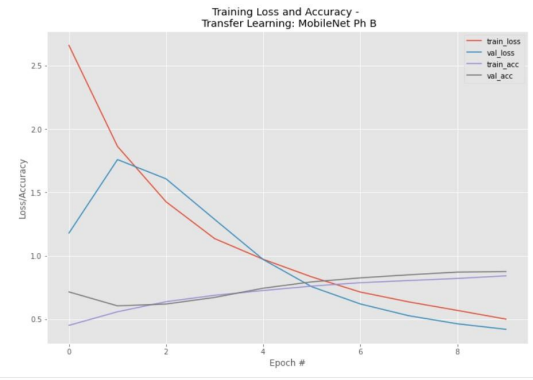


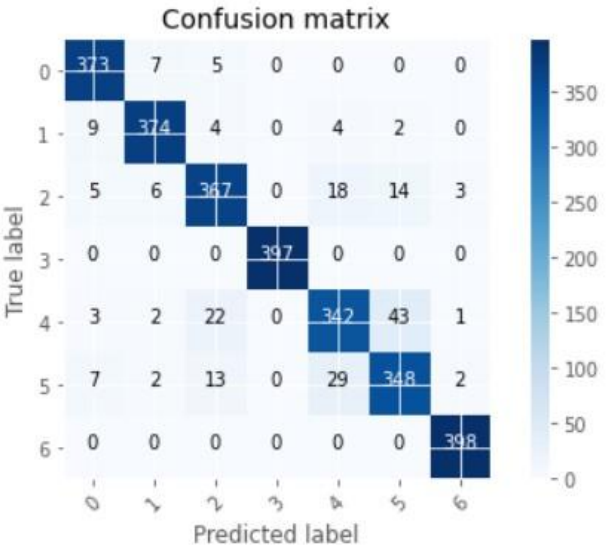


MobileNet:

stages	Balanced	Epochs	hidden layers	data augmentation	batch size	accuracy	validation accuracy	loss	validation loss
2	True	10	[256,256,64]	True	200	0.8411	0.8750	0.4997	0.4192



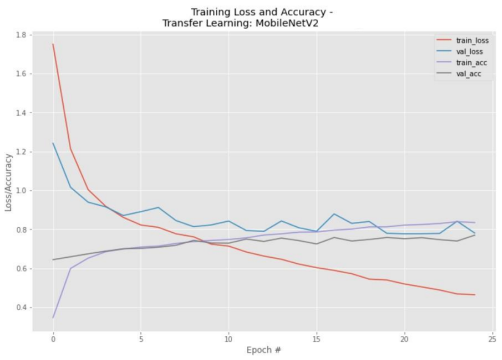


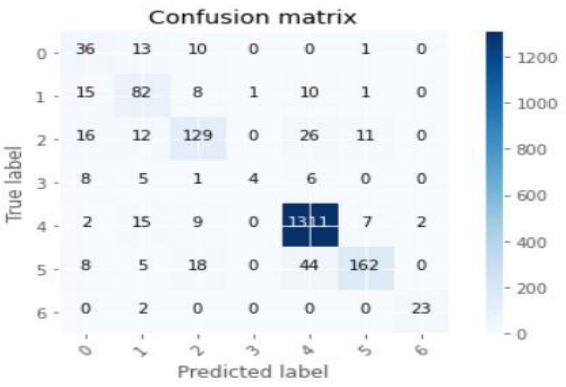


MobileNetV2:

First Model's confusion matrix and accuracy test:

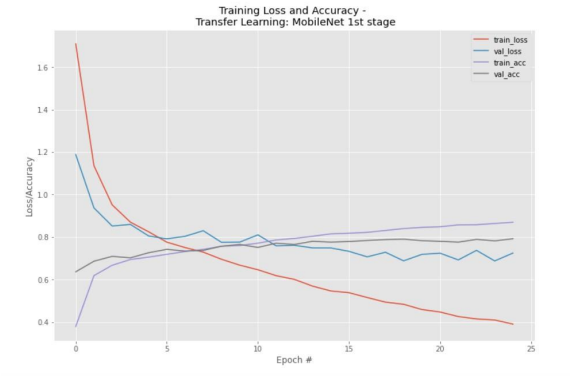
stages	Balanced	Epochs	hidden layers	data augmentation	batch size	accuracy	validation accuracy	loss	validation loss
1	False	25	[512,128] [256,256,64]	False	100	0.8342	0.7695	0.4641	0.7804

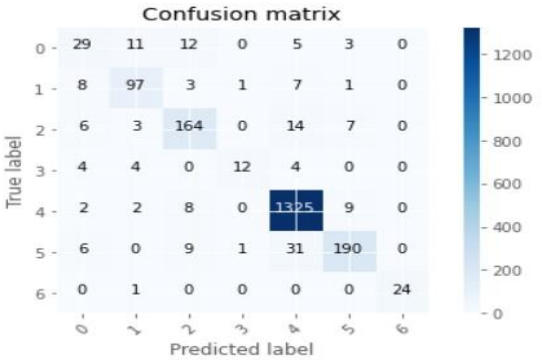




Second Model's confusion matrix and accuracy test:

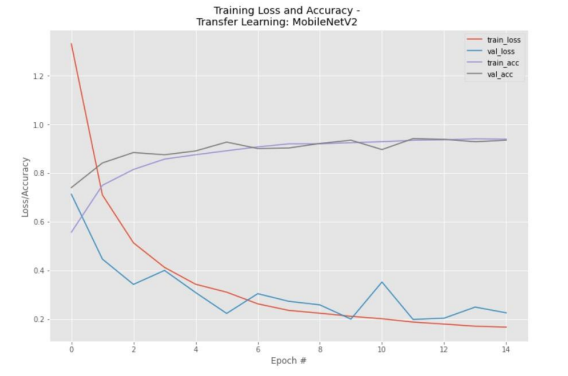
stages	Balanced	Epochs	hidden layers	data augmentation	batch size	accuracy	validation accuracy	loss	validation loss
1	False	30	[512,128]	False	1000	0.8692	0.7914	0.3895	0.7240

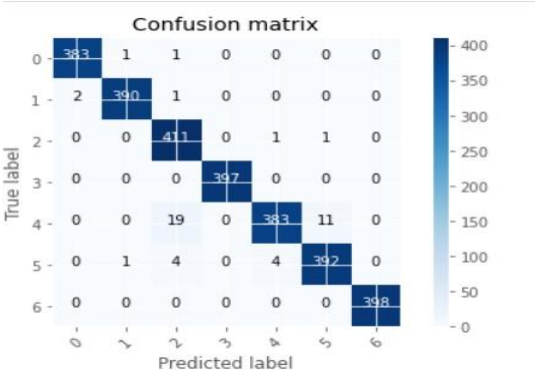




Third Model's confusion matrix and accuracy test:

stages	Balanced	Epochs	hidden layers	data augmentation	batch size	accuracy	validation accuracy	loss	validation loss
1	True	15	[512,128] [256,256,64]	True	100	0.9394	0.9350	0.1666	0.2253

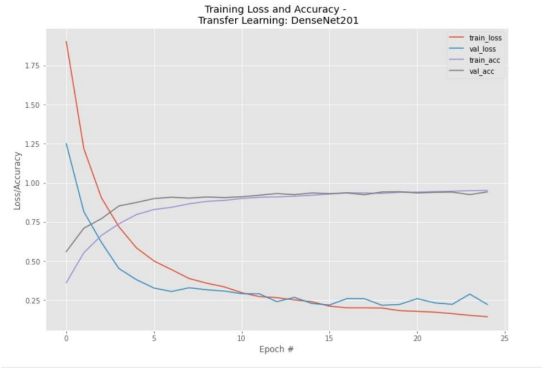


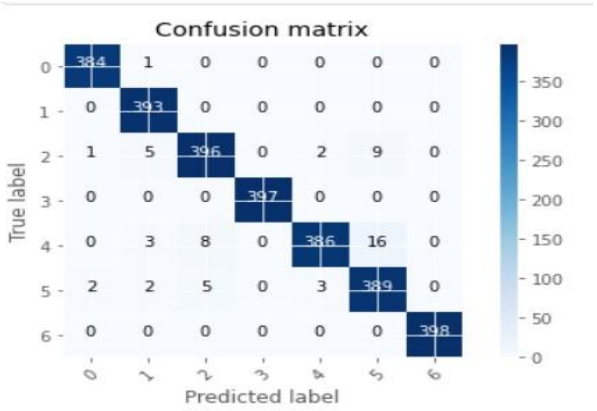


DenseNet:

First Model’s confusion matrix and accuracy test.

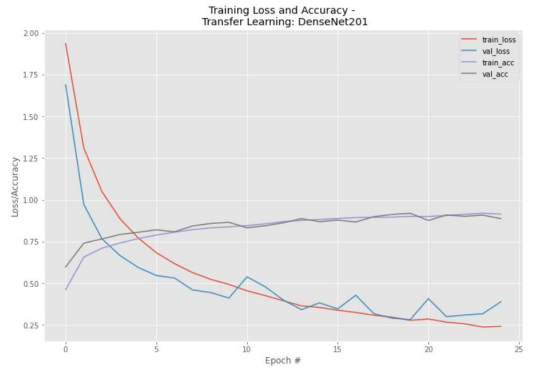
stages	Balanced	Epochs	hidden layers	data augmentation	batch size	accuracy	validation accuracy	loss	validation loss
1	True	25	[512,128] [256,256,64]	Trye	100	0.9516	0.9421	0.1448	0.2231

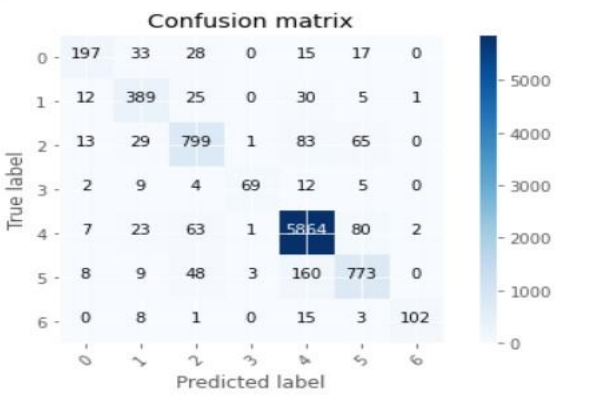




Second Model’s confusion matrix and accuracy test:

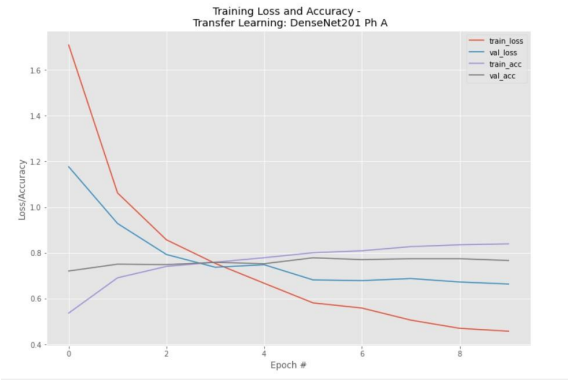
stages	Balanced	Epochs	hidden layers	data augmentation	batch size	accuracy	validation accuracy	loss	validation loss
1	False	25	512,128	True	100	0.9140	0.8872	0.2421	0.3896

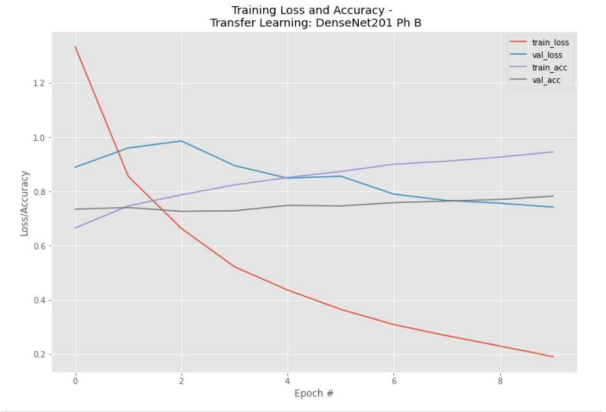


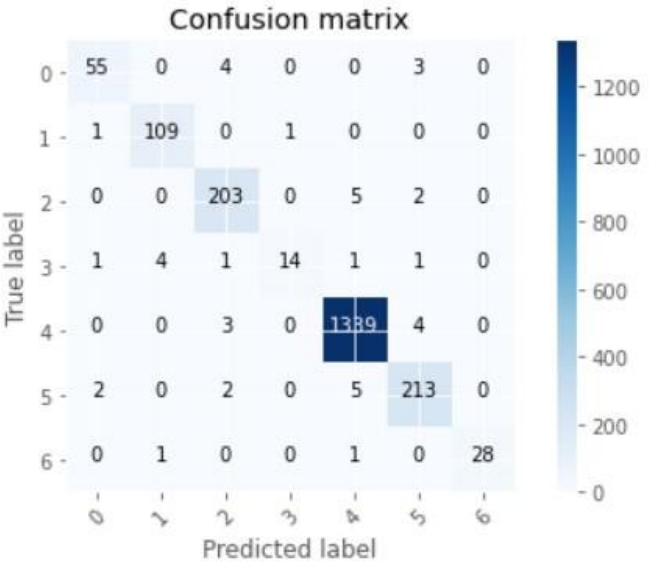


Third Model's confusion matrix and accuracy test.

stages	Balanced	Epochs	hidden layers	data augmentation	batch size	accuracy	validation accuracy	loss	validation loss
2	False	10	512,128	True	128	0.9457	0.7824	0.1906	0.7419







Conclusion:

Melanoma is the worst time of skin cancer, hence early detection of it is a necessity of preventing a life-threatening scenario. Therefore it is critical to employ supportive imaging modalities that have been proven to help with diagnosis. In our model we provided a transfer learning model by using CNN then several architectures as MobileNetV2, ,ResNet50, VGG16 and DenseNet201 that can be used to investigate any suspicious lesion. This method is applied to a dataset ham10000 of skin cancer disorders. We obtained adequate response in testing accuracy and training.In addition the imbalances between the classes in our dataset hinders the model from acquiring better accuracy. So as a result we concatenated the data set to make it more balanced, as well as setting the weights for classes.

Future work:

In the future, we will work on a dataset with better labeled skin lesions and images and contains less imbalances between its classes to acquire the best testing accuracy obtainable. As well as we will try using machine learning algorithms after transfer learning to notice if it gets better results.

References:

1- https://www.sciencedirect.com/science/article/pii/S2666827021000177#sec2

2- https://link.springer.com/chapter/10.1007/978-981-16-2422-3_26

3- https://www.sciencedirect.com/science/article/pii/S0959804919303491

4- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7006718/

5- https://www.who.int/news-room/questions-and-answers/item/radiation-ultraviolet-(uv)-radiation-and-skin-cancer

6- https://www.sciencedirect.com/science/article/pii/S0959804919303491

7- https://www.analyticsvidhya.com/blog/2021/10/understanding-transfer-learning-for-deep-learning/

8- https://www.mdpi.com/1424-8220/23/2/570

9- https://datagen.tech/guides/computer-vision/vgg16/

10- https://arxiv.org/ftp/arxiv/papers/2303/2303.07520.pdf

11- https://www.mdpi.com/2075-4418/13/11/1911