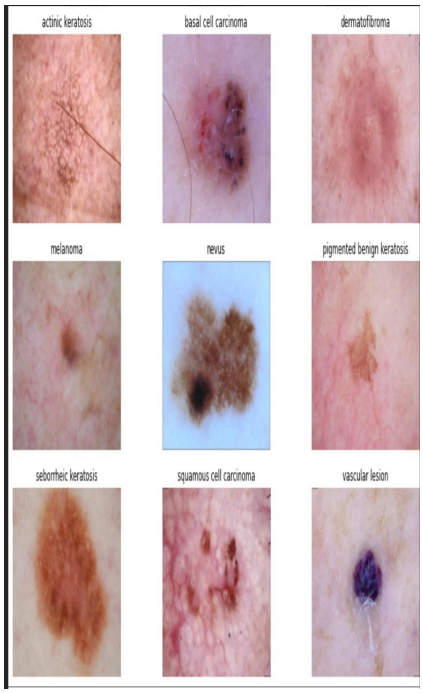
Define the problem

Skin cancer is a dangerous and widespread disease [1] Each year there are approximately 5.4 million new cases of skin cancer are recorded in the USA alone [2], [4]. The global statistics are equally alarming [3][5]. Recent reports show that from 2008 to 2018, there has been a 53% increase in new melanoma cases diagnosed annually [1,4]. The mortality rate of this disease is expected to rise in the next decade. The survival rate is less than 14% if diagnosed in later stages.

Dataset

The dataset was generated by the International Skin Imaging Collaboration (ISIC) and images are from the following sources: Hospital Clínic de Barcelona, Medical University of Vienna, Memorial Sloan Kettering Cancer Center, Melanoma Institute Australia, University of Queensland, and the University of Athens Medical School.



Input & Output

Developing a DeepLearningmodel to classify skin cancer images into 9 different categories with high enough accuracy to be reliable. Those categories are

- actinic keratosis
- basalcellcarcinoma
- dermatofibroma
- melanoma
- nevus
- Pigmentedbenignkeratosis
- seborrheic keratosis
- squamouscellcarcinoma
- vascular lesion

Motivation

If skin cancer is detected at early stages then the survival rate is nearly 97% [3]. This demands the early detection of skin cancer. This paper addresses the issue of early diagnosis, with improved accuracy.
It is found that a skilled dermatologist usually follows a series of steps, starting with the naked-eye observation of suspected lesions, then dermoscopy (magnifying lesions microscopically), and followed by a biopsy. This would consume time and the patient may advance to later stages. Moreover, accurate diagnosis is subjective, depending on the skill of the clinician. It is found that the best dermatologist has an accuracy of less than 80% incorrectly diagnosing skin cancer. Adding to these difficulties, there are not many skilled dermatologists available globally in public healthcare.

To diagnose skin cancer speedily at the earliest stage and solve some of the aforementioned problems, there have been extensive research solutions by developing computer image analysis algorithms. The majority of these algorithmic solutions were parametric, meaning that they required data to be normally distributed. As the nature of data cannot be controlled, these methods would be insufficient to accurately diagnose the disease. However, non-parametric solutions do not rely on the constraint that the data is in normal distribution form.
Our Model augmented assistance to the dermatologist is provided using deep learning. The essence of the approach is that a computer is trained to determine the problem by analyzing the skin cancer images.

What makes it difficult?

- Working with Multi-Class is not easy for getting good accuracy.
- Not Available Large Datasets.
- Datasets are available imbalanced.
- The distribution of data is not good.

Related work:
What has been done? The paper we used relied on VGG-16 which resulted in an accuracy of 69.57% while the theVGG-19 had a better accuracy of 71.19% The paper also did some preprocessing on the images such as Augmentationsetting Range

rotation range	90
shear range	0.1
zoom range	0.14
horizontal flip	True

What are the problems?

- they both resulted in a model of a bigger size and used a padding 0
- which resulted in an image with fewer features in this particular model.
- They are more attention to models than distribution data.

Working about data:

- Datasets 4600 images, which is very small for 9 class, in addition, data is imbalanced and have very bad distribution. So, We working to solve this.
- Firstly, we split data into two partitions final set & train set, then worked in data training to increase images and slove distribution.
- We making Augmentation using Python package is known as Augmentor (https://augmentor.readthedocs.io/en/m to add more samples across all classes so that none of the classes have very few samples

- Instantiate a Pipeline object pointing to a directory containing your initial image data set. Define several operations to perform on this data set using your Pipeline object. Execute these operations by calling the Pipeline's sample() method.
- Finally, we split data training to train & validation, we worried that this approach makes overfitting, so, we split from data partition to data set

Algorithm

- We Used Xception & inceptionv3 because the two algorithms use padding the same as it is suitable for our data and utils suitable resources, not as VGG.
- Our data is not very large, so transfer learning is a very good option, we do not make unfreeze layer also this problem dissimilar problem, but we used it as an initial weight instead of We use Random initial weight, this option gives us a good result.
- Also, we added some layers
  - GlobalAveragePooling2D
  - Dropout
  - BatchNormalization
  - Dense
- Using optimizers Adam with learning_rate= 3e-4
- Using callbacks with
  - early stopping
  - ReduceLROnPlateau to reduce the learning rate the confusion metric has stopped improving.
- Xception accuracy with
  - Validation 94
  - final test 91
- inceptionv3 accuracy with
  - Validation 91
  - final test 88

  So, we select Xception

Evaluation

- We used the classification report to show the precision, recall, F1 Score, weighted average, and micro average.
- We used confusion_matrix in the final test to know the conflict with classes

Validation Result

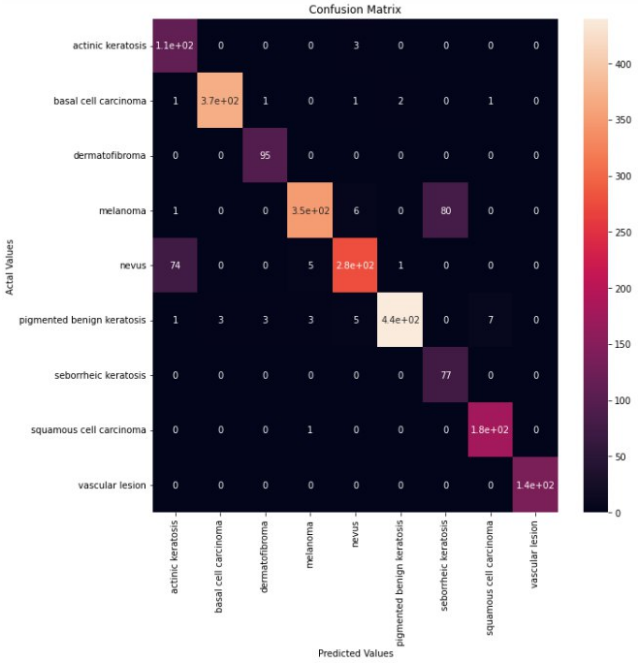
```
Testing:
57/57 [==============================] - 583s 10s/step - loss: 0.1653 - accuracy: 0.9472
```

	precision	recall	f1-score	support
actinic keratosis	0.82700	0.98000	0.89703	200
basal cell carcinoma	0.97512	0.98000	0.97756	200
dermatofibroma	0.99502	1.00000	0.99751	200
melanoma	0.97590	0.81000	0.88525	200
nevus	0.95758	0.79000	0.86575	200
pigmented benign keratosis	0.97938	0.95000	0.96447	200
seborrheic keratosis	0.84615	0.99000	0.91244	200
squamous cell carcinoma	0.99010	1.00000	0.99502	200
vascular lesion	1.00000	1.00000	1.00000	200
accuracy			0.94444	1800
macro avg	0.94959	0.94444	0.94389	1800
weighted avg	0.94959	0.94444	0.94389	1800

Final Test

	precision	recall	f1-score	support
actinic keratosis	0.59043	0.97368	0.73510	114
basal cell carcinoma	0.99196	0.98404	0.98798	376
dermatofibroma	0.95960	1.00000	0.97938	95
melanoma	0.97500	0.80137	0.87970	438
nevus	0.94863	0.77591	0.85362	357
pigmented benign keratosis	0.99323	0.95238	0.97238	462
seborrheic keratosis	0.49045	1.00000	0.65812	77
squamous cell carcinoma	0.95745	0.99448	0.97561	181
vascular lesion	1.00000	1.00000	1.00000	139
accuracy			0.91112	2239
macro avg	0.87853	0.94243	0.89354	2239
weighted avg	0.94064	0.91112	0.91732	2239

Confusion Matrix in the final test



Analysis

- Our results are very good with another solution, but not very good for the medical field.

We Noticed, Model, did not good t0 difference between:

- Actinic & Nevus
- seborrheic keratosis & melanoma

In our future, the plan is to work around data to solve this problem

- Our solution we working on using some preprocessing in data to solve this conflict.
- Another solution, make our system consist of two layers and merge Actinic & Nevus in one category and seborrheic keratosis & melanoma in one category, and layer two have two models of binary classification if the result is one of the merged categories, the disadvantage will resources usages

. Contribution

- **Mohamed Badr**-sec 7
- **Abdullah Alqasabi**-sec 5
- **Abdelrahman Khaled**-sec 4
- **Omar Adel Hassanen**-sec 5
- **Ibrahem Abd-Alatef**-sec 1