CLASSIFICATION OF SKIN CANCER IMAGES USING CONVOLUTIONAL NEURAL NETWORKS

Ritik Gulati, Vayam Bhatt, Tushar Tyagi, Indu Khatri
1,2,3 B.tech Student, 4Assistant Professor
Department of Computer Science Engineering
Mahavir Swami Institute of technology, Sonipat

Abstract: - Skin cancer is the most prevalent human malignancy [1], and it is primarily diagnosed visually, beginning with a clinical examination and perhaps followed by dermoscopic (skin-related) analysis, a biopsy, and histological investigation. Skin cancer is caused by mistakes (mutations) in the DNA of skin cells. The mutations lead the cells to expand out of control, resulting in a tumor. The goal of this study was to use convolutional neural networks to classify photos of skin lesions. Deep neural networks have enormous potential for image classification while taking into consideration the environment's high unpredictability. Here, we used pixel values to train photos and then used disease labels to classify them. The dataset was obtained from the ISIC(International Skin Imaging Collaboration)[3] Archive, which was obtained from an Open Source Kaggle Repository[2]. Multiple models were used in the training accompanied with Transfer Learning. The highest model accuracy achieved was over 86.65%. The dataset used is publicly available to ensure credibility and reproducibility of the aforementioned result.

Keywords: Convolutional Neural Network, Machine Learning, Malignant, Benign, Transfer Learning, Computer Vision, Skin Cancer, Confusion Matrix, Deep Learning, Gradient Class Activation Maps

1. INTRODUCTION

The skin is the body's largest organ, protecting all of the internal organs from the hostile environment outside. It aids in temperature regulation and infection protection. There are three layers to the skin:

The epidermis, dermis, and hypodermis are the three layers of the skin.

Cancer is diagnosed when a mass of healthy skin cells alters its properties and begins to grow uncontrollably, becoming a tumour. This tumour can be benign, meaning it causes no harm to the patient and does not develop or spread to other parts of the body, or it can be malignant, meaning it disrupts the natural flow of things and can grow and spread to other parts of the body.

The most prevalent type of cancer is skin cancer, which is caused by excessive exposure to the sun or other high-energy particles that disrupt the genetic makeup of the cell.

As the world enters a new millennium, pollution in the air is increasing at an exponential rate, allowing skin disorders such as skin cancer to thrive. The only protection from the sun's dangerous UV rays, which are the leading cause of skin cancer, is the Ozone Layer. This layer is decreasing geometrically as a result of rising levels of air pollution, and consequently the number of skin cancer cases is increasing. Melanoma (malignant) and non-melanoma (benign) skin cancers are the two most common types of skin cancer (Benign). Melanoma is one of the most lethal cancers. Early identification of this cancer, on the other hand, can be beneficial. Early identification of this malignancy, on the other hand, can improve the chances of survival.

Before deciding on the actual functions for this thesis, various methodologies were investigated and studied. We used the greatest and most advanced machine learning resources we could find. Skin cancer detection through histopathology can be achieved by augmenting the power of deep learning. The classification of skin cancer Using customised convolutional neural networks to detect and possibly cure disease in its early stages can aid the medical community. Discoloration or inflammation of the skin, itching and bleeding of skin patches, moles or red, waxy bumps on the skin caused by malignant cells are some of the early symptoms associated with this cancer.

The data was taken from an open source Kaggle repository [2], which was part of the ISIC (International Skin Image Collaboration). The photos were separated into test and train portions after being retrieved. The dataset was expanded with changed copies of photographs using an appropriate level of image augmentation. The training dataset had over 2900 photos divided into two categories: “benign” and “malignant.” The test set had 350 images and the training dataset had over 2900 images divided into two categories: "benign" and "malignant."

We discovered that classification using K-nearest Neighbours, Support Vector Machines, or even Decision Trees yielded low precision and accuracies within the scope of our study. Further research into the mathematics underpinning categorization revealed that using Deep Learning models was the most complex way for achieving the desired results. We tried a variety of mathematical models, both with and without Transfer Learning, but in the end, it was determined that the depth and quality of activation provided by pretrained models were insufficient, so we combined our mathematical knowledge and created a Dense Convolutional Network model with an accuracy of over 86.6 percent.
Deep Learning, a comprehensive and powerful offshoot of AI, supplied us with the capabilities we needed to undertake effective and accurate image classification. It had the same structure and function as a human brain, with neurons firing across the brain, conveying information, classifying data, forming inferences, and creating outcomes. Deep Learning employs a Neural Network structure, which is made up of a stack of layers. As the name implies, neural networks function similarly to neurons in the brain, finding patterns and making predictions.

The study's goal is to develop a skin cancer screening technology that is both speedier and more accessible. The earlier it is discovered, the more likely the victim will recover. Skin cancer is highly treatable if detected early, according to the American Academy of Dermatology Association. A skin biopsy is performed after the model has been taught. A skin biopsy performed by a dermatologist is the only technique to confirm the presence of skin cancer.

1.1 Related Works
Many additional articles and publications were referenced to and consulted in order to maintain the paper as current as feasible. With his use of mathematical models on an enlarged dataset of 3000 photos, the author [5] achieved a maximum accuracy of 0.74. In his paper, the author also hopes to increase early melanoma diagnosis through the creation of machine learning algorithms.

A few additional publications [7] were reviewed, and the authors determined that his proposed system had an 81 percent accuracy. His proposed approach used non-dermoscopic images of lesions from which he isolated lesion patches and computed the findings based on colour and texture.

2. MATERIALS AND METHODS

2.1 Dataset
A dataset of 2947 histopathology images were considered for this study. The dataset was acquired from an open source Kaggle Repository [2], a subset of the ISIC (International Skin Imaging Collaboration) archive [3]. The dataset consists of two classes namely, benign and malignant. The test dataset comprises 350 images taken from the Kaggle repository [2]. The dataset chosen is publicly distributed and open source to ensure better credibility of the models.

2.2 Methodology
The aim of the study is to determine the efficiency and the accuracy of different models in-order to classify images between being benign and malignant. The study takes into account multiple Transfer Learning models paired with a set of Dense Layers to predict the class of the image.

2.3 Augmentation
The Dataset that was obtained was sufficiently supplemented. Exaggerating the dataset to create unpredictability and increase the data so that a more robust and trustworthy model can be trained is known as data augmentation. We chose the augmentation settings very carefully so as not to affect the image properties that are important in determining the results. • The photos were re-scaled to a re-scaling ratio of 1:2.55:0.5 as part of the augmentation.
• The photos were sheared at an angle of 0.2 degrees, which is the anticlockwise shear angle.
• To maintain the lesion markings and create fresh synthetic data, the photos were randomly flipped horizontally and vertically.
• A zoom augmentation is implemented, which randomly zooms in on the image and either adds new pixel values or interpolates pixel values.

All the images generated were passed through a generator and were passed into the model as a whole with a height and width of (176, 176). All the images were passed through all three of the channels, i.e., red, green and blue (R, G, B).

2.4 Convolutional Neural Network
Convolutional Neural Networks are a type of deep learning network that is used to analyse picture or visual data. A Convolutional Neuron is made up of weights and biases that can be learned and modified to provide desired results. A Convolutional Neural Network, or CNN, is made up of thousands or millions of these neurons arranged in a specified order with other neurons. CNNs are mostly used to classify and organise visual data, cluster them together if they appear to be similar, and then recognise objects.

Filters or kernels are used to convolve data or images. Filters are little matrices that we use with a sliding window to engage over the dataset. The image’s depth is the same as the input; for example, if the depth of a coloured RBG image is 4, a filter with a depth of 4 would be applied to it.
The next step is to use a rectifier function (usually known as a Rectified Linear Unit) to boost the Network's non-linearity. Finally, the features are down sampled using a pooling layer that is applied through the image volume. Finally, we utilise a fully connected Layer to flatten the entire 3d feature matrix into a single column, which is then supplied to the neural network for further processing. Convolutional Layers, Activation Layers, Pooling Layers, and Fully Connected Layers are all coupled to form a Convolutional Neural Network, as previously indicated.

Formulating the convolution function in a neural network, the inputs of layers may change. Internal covariate shift, or changes in the distribution of inputs to a layer for each mini batch, can result as a result of this.

For a neural network, batch normalization entails scaling the layer's output by normalizing the activations of each input variable each mini batch. Batch normalization makes the model more robust and less susceptible to hyper parameter tuning impacts.

2.8 Flatten
Flatten is a technique for converting data into a 1D array that can be used as an input for the next layer. The output vector, or feature matrix map, is flattened to create a single feature vector that may be sent into the next layer for additional processing.

2.9 Pooling
One of the most important layers of a deep neural network is pooling. The pooling layer is responsible for reducing a feature matrix's dimensionality or spatial size. The pooling layer reduces the amount of computing resources needed to train a neural network. It's also important for extracting the data's most important features.

Max Pooling and Average Pooling are the two main types of pooling layers that can be used in a deep neural network.

The maximum value for the picture area covered by the kernel is given by Max Pooling, as the name suggests. The average of the values for the picture region covered by the kernel is returned by Average Pooling. We utilised Max Pooling instead of Average Pooling in our study since it helps to eliminate noise or noisy features while also reducing dimensionality, whereas Average Pooling just helps with dimensionality reduction. As a result, the pooling layer aids in the extraction of useful information from a sea of data, which is critical for assisting a model in learning the features in a dataset. The removal of noise and the presentation of just useful features aids in the prevention of overfitting and accelerates the computation.

2.10 Dropout
Dropout Layers are employed in big networks to reduce the amount of overfitting. Simply put, a dropout layer will discard a certain amount of random Neurons during training to avoid overfitting and bias in the model's output.

2.11 Categorical Crossentropy
Categorical Cross Entropy is a loss function employed in a Network where there is a multiclass classification.
This function is an excellent method of differentiating between different classes or discrete probability distributions.

2.12 Optimizer and Metrics
Optimizers assist in fine-tuning the weights so that the loss function is as little as possible. Their goal is to cut down on losses and improve performance. For training the model with a variable learning rate, the 'adam' optimizer was employed. After the validation-loss crossed a particular patience threshold with a decrease factor of 10-1, the learning rate was adjusted, or more precisely, reduced, to a value of 0.0001. The model's metric was 'accuracy,' and early halting was used to avoid the model from overfitting or overtraining. The validation-loss with a high yield patience level was again the monitored hyper-parameter.

Activation Functions are commonly used to convert a network's output to a probability distribution over the output classes. SoftMax was used as the activation function for the model's feature extraction layer.

2.13 Models Used
The study necessitated a thorough evaluation of a variety of transfer learning models. For this study, DenseNet, XceptionNet, ResNet, and MobileNet were all examined as transfer learning models. Because of its strong gradient flow, improved back propagation, and smooth decision boundaries, DenseNet201 was chosen. Xception Net is a modified depth-wise separable convolution that has the same number of parameters as InceptionV3 but is more efficient in terms of computation. Resnet50 demonstrates the notion of skip connection, which helps the model learn the identity function appropriately for better results and avoids the problem of vanishing gradients (which occurs when training very deep networks). MobileNetV2 additionally includes depth-wise separable convolution, which reduces the complexity and size of the network.

2.14 Gradient Class Activation Maps
Grad CAM is a technique for analysis of the Activations generated by the trained model. Grad CAM simply is a heatmap generated by the Convolutional Layers depicting the importance of a certain feature or the part of an image. It encapsulates the area or region.

2.15 Training
After completing a training and validation split, the Model was trained on the previously given dataset. The aforementioned models were trained with 60 epochs, albeit early-stopping was used due to validation-loss, as previously noted. The number of steps per epoch was calculated by dividing the batch size by the total number of photos in the database. Varied models produced different outcomes, which are depicted in the graph below. Earlystopping was used to avoid overtraining, and it tracked validation loss as a parameter. When the model's performance deteriorated, the previous weights were restored. We also used a conditional reduction in learning rate of 1/10 to ensure that when the function's rate of convergence plateaued, the learning rate was reduced for a more exact reduction in the hyper-slope.

3. RESULTS
A total of 350 photos were used to test the models, including 150 from the malignant class and 200 from the benign class. For each model, the accuracy, precision, recall, and F1-score were calculated to gain a better understanding of its efficiency and operability. Based on the test data, the confusion matrix was drawn for each model, and the metrics were generated using the formulas below.
3.1 DenseNet201
DenseNet201 did exceptionally well in both the Validation and the Final Test splits. The combination of a Neural Network and Dense Layers, which were then combined with a Leaky ReLU, resulted in excellent accuracies and an F1 score. DenseNet assisted in the solution of the vanishing-gradients problem and facilitated feature reuse and propagation, resulting in significant results and accuracy. The Confusion Matrix obtained from the test set, as well as a graphical representation of the metrics utilised during the training and validation processes, are provided below for visual insights.

The accuracy recorded over the test set for this model was 86%.

The activation supplied by the last layer is shown as an overlayed heat map in the class activation map for this model. The red-yellow colour shows the region of interest, whereas the violet-blue colour indicates low activation areas.

3.2 ResNet50
The ResNet50 architecture has skip connections, which means that the input to one layer can be sent directly to another, enhancing the performance of this neural network over others. It also aids in the solution of the problem of vanishing gradients via identity mapping. To make comparisons with the other models in this study, the confusion matrix was shown together with the metrics utilised throughout the training and validation processes. On the test data, the model had an accuracy of over 86.57 percent.

The activation supplied by the last layer is shown as an overlayed heat map in the class activation map for this model. The red-yellow colour shows the region of interest, whereas the violet-blue colour indicates low activation areas.

3.3 XceptionNet
The feature extraction backbone of the Xception-Net architecture is comprised of 36 convolutional layers. In linear geometry, the 36 convolutional layers are assembled into 14 distinct modules. This setup allowed us to achieve incredible precision in the subsequent organisation of the dense and...
batch normalisation layers. The model was trained using the same collection of training data, and the model's performance was evaluated using the same set of validation photos. For the test dataset, the validation graph and training graph were plotted, and the confusion matrix was examined to gain a better understanding of the results and to assess the model's performance. The model had an accuracy of more than 82.5 percent, as seen in the graphs below.

3.4 MobileNet
MobileNet is a simpler and lighter model that is typically trained for mobile devices with lower computational capacity. The goal of using MobileNet was to make it possible to scan for malignant or benign lesions using globally accessible devices like cellphones. This model was created primarily to bring the model's power to the general public. Considering its efficiency-oriented orientation, the model worked admirably. This architecture was also trained throughout the complete dataset.

According to the tabulated data, the model had an accuracy of over 80.8 percent.

The activation supplied by the last layer is shown as an overlayed heat map in the class activation map for this model. The red-yellow colour shows the area of interest, whereas the violet-blue colour represents low activation areas.

3.5 Comparisons
The metrics were plotted after compiling the findings from all of the models. Based on the outcomes of the test set prediction, some metrics were calculated. To compare the model performance of all the chosen architectures, the metrics Accuracy, Precision, Recall, and F1-Score were utilised.
4. CONCLUSION AND OUTLOOK

Skin Cancer and similar disorders may be diagnosed and classified instantly using Machine Learning and Artificial Intelligence-based classification. The models we created demonstrate that cost-effective, user-friendly, and non-invasive Machine Learning and Artificial Intelligence-based solutions can be developed as a first step in fighting Skin Cancer and other disorders. Each model we trained during the research process was a successor to the previous one, and every new model we built was based on the prior one's flaws. All of the information gleaned from the precursing models was used to increase the accuracy of the new models to come.

The final models were trained and analysed in the study to increase the efficiency and reliability of the approaches used to identify skin cancer. The comparison of multiple models assisted in identifying the optimum model that may be used in real-world applications for disease early detection. The models utilised in the study have varied approaches to diagnosing the condition, each with its own set of procedures and attributes. The findings of this study have been given to the scientific community in both written and visual form (through graphs and confusion matrices) for their consideration.

Furthermore, we hope that our findings will aid in the continuing crisis and provide a valuable contribution to the virus's fight. Human judgment, on the other hand, is required. As a result, we are not attempting to replace a Dermatologist with our model, but rather to assist with the forecasts.

4.1 Future Scope

This research can be comprehended and used as a starting point for further research. The preliminary findings presented here may only be the start of a more in-depth examination of Machine Learning's utility in cancer research. Such research may one day assist the general public in quickly detecting the presence of any malignant tumour and receiving timely treatment.

REFERENCES

- American Cancer Society. Cancer facts & figures 2016. Atlanta, American Cancer Society 2016.
- http://www.cancer.org/acs/groups/content/@research/documents/document/acspc-047079.pdf
- https://www.kaggle.com/fanconic/skin-cancer-malignant-vs-benign
- https://www.isicarchive.com/#/topWithHeader/wideContentTop/main
- https://www.aad.org/public/diseases/skin-cancer/find/now-how
- https://theses.org/Downloads/Volume10No7/Paper_46_Convolutional_Neural_Network.pdf
- https://www.researchgate.net/publication/334751850_Skin_Cancer_Detection_Using_Convolutional_Neural_Network#:~:text=This%20paper%20proposed%20an%20artificial,images%20using%20feature%20extraction%20technique.&text=Flow%20chart%20for%20the%20system%20using%20convolutional%20neural%20network.
- Giotis, I., Molders, N., Land, S., Biehl, M., Jonkman, M. F., & Petkov, N. (2015). MED-NODE: a computer-assisted melanoma diagnosis system using non-dermoscopic images. Expert systems with applications, 42(19), 6578-6585
- IRJET-V4I4702-with-cover-page-v2.pdf (d1wqtxlxzle7.cloudfront.net)
- 169_Melanoma_Skin_Cancer_Detection_using_Image_Processing_and_Machine_Learning20190703-23098-l08hr8x-with-cover-page-v2.pdf (d1wqtxlxzle7.cloudfront.net)