Abstract—Melanoma is one of the most deadly cancers on the planet, and it has the ability to spread to new places of the body if not discovered early enough. Dermoscopic pictures are routinely used to identify melanoma. Many previous studies, based on standard classification approaches and deep learning models, have been proposed for automated analysis of skin lesions. In traditional classification systems, handcrafted functions are used as input. However, due to the high visual similarity between different classes of skin lesions and complex skin diseases, hand-made features are not discriminating enough and fail in many circumstances. Convolutional networks with fewer connections between the input layers and those close to the output can be significantly deeper, more accurate and more effective for training, according to a recent study. In this study, we accept and offer the application of Hadoop Distributed Tight Convolution Network (HdiDenseNet) for melanoma skin cancer detection and study of failed lesions, which can be connected to the mode of transmission. You can use all feature maps at all input levels as input, and your own maps will be used as inputs in all poster headers. DenseNets has a number of tentative features, including the ability to overcome the gradient drift problem, improve feature propagation, increase feature reuse, and dramatically reduce the amount of money. DenseNets is superior to most currencies for the mayor’s office, which forces me to remember and process. The tests used a total of nearly 40,000 photographs gathered from various sources. The best performing Distributed Densely Connected Convolutional Network technique using Patient’s Metadata of
Dermoscopic Images for Early Melanoma Detection receives great accuracy and Area Under Curve scores, according to the results (AUC).

**Keywords**---dense block, transition layer, densenet, melanoma, conV, pooling.

**Introduction**

The extremely common category of cancer is skin cancer. Even though have being the smallest amount common skin cancer, melanoma is accountable for 85 percent of skin cancer deaths. According to the American Cancer Society, approximately 1 lakh new instances of melanoma will be detected in 2020 to 2022. Nearly 17,000 individuals are anticipated to die because of the sickness. Early and accurate detection possibly supported by data science may make it to treatment more successful, just as it can with other malignancies. Currently, physicians examine all a patient's moles to detect "ugly ducklings" or outlier lesions that are most likely to be melanoma. Existing AI techniques haven't taken this clinical context into account enough. You must detect melanoma in photographs of skin lesions in this competition. You'll utilise photos from the same patient to figure out which ones are more likely to have melanoma. Using patient-level contextual information in the creation of image analysis technologies could assist clinical dermatologists better support their patients. Melanoma is a fatal illness.

Skin cancer is an invasive disease caused by abnormal melanocyte cell growth in the body, which proliferates and spreads through lymph nodes, killing surrounding tissues [1]. Melanoma skin cancer must be detected early in order for patients to have a good chance of recovery [2]. Pale skin tone, sunburns, and past genetic reasons, as well as a compromised immune system and excessive exposure to ultraviolet light rays, are all risk factors for melanoma.

Melanoma has the ability to grow and spread over the first layer of skin, the epidermis, until it encounters a lymph vessel and, eventually, blood, if it is not detected early. Moles with uneven borders, forms, colour change, and a diameter greater than 6mm are some significant features for detecting skin cancer, and they appear as moles with uneven borders, forms, colour change, and a diameter greater than 6mm. Various non-invasive approaches for detecting cancer and determining if it is benign have been presented. [2]. Melanoma diagnosis is currently done by visual analysis by a specialist. This method, however, may take some time. Furthermore, due to the difficulty of making the diagnosis, it is possible to make a mistake [3].

Throughout history, several academics have worked to try to build an effective Melanoma exposure system established on machine learning (ML) [4,5]. Finding a decent diagnosis method in practice is difficult due to artifacts such as hair across or just as in the lesion, adjusting lesion size, colour, and shapes, the presence of blood basins, and other artifacts [6], as demonstrated in the accompanying Figures.
The writers were inspired to write because of these features. Even while most systems follow the same fundamental method, with pre-processing coming first, followed by segmentation, quality extraction, and classification, these characteristics prompted the scientists to expand their research.

**Materials and Methods**

The systems used in this study are AI-based, which implies they can learn from one or several DSs (both small and large ones). These DSs are made up of high-resolution, hand-picked photographs that have been analyzed, categorized, and maybe subdivided by medical professionals in the field. Our research aims to demonstrate the developing trend of automated systems capable of diagnosing, segmenting, or identifying SLs based on available literature (notably Melanoma). These DSs include PH2, ISIC 2016, 2017, 2018, and 2019 challenge DSs, HAM10000, DermNet Atlas, Dermatology Atlas, Dermis, and MED-NODE (table-1) [3]

**Table-1 : Melanoma frequently used data sets**

<table>
<thead>
<tr>
<th>Dataset name</th>
<th>References</th>
<th>Availability</th>
<th>SL</th>
<th>ME</th>
</tr>
</thead>
<tbody>
<tr>
<td>PH2</td>
<td>[15]</td>
<td>yes</td>
<td>200</td>
<td>40</td>
</tr>
<tr>
<td>ISIC 2018, HAM10000</td>
<td>[16][17]</td>
<td>yes</td>
<td>10,015</td>
<td>1113</td>
</tr>
<tr>
<td>ISIC 2019</td>
<td>[18,19,20]</td>
<td>yes</td>
<td>25,333</td>
<td>4522</td>
</tr>
<tr>
<td>ISIC 2020</td>
<td>[18]</td>
<td>yes</td>
<td>33,126</td>
<td>584</td>
</tr>
<tr>
<td>DERMQUEST</td>
<td>[21]</td>
<td>yes</td>
<td>126</td>
<td>66</td>
</tr>
<tr>
<td>MED-NODE</td>
<td>[22]</td>
<td>yes</td>
<td>170</td>
<td>100</td>
</tr>
<tr>
<td>DERMNET</td>
<td>[23]</td>
<td>yes</td>
<td>22,500</td>
<td>635</td>
</tr>
<tr>
<td>DERMOFIT</td>
<td>[24]</td>
<td>yes</td>
<td>1300</td>
<td>76</td>
</tr>
</tbody>
</table>
PH2 is one of the most commonly used dermoscopic catalogues in some articles (DB). This can be found in [14]. The images were chosen for their quality, resolution, and dermoscopic features. The ABCDE [25] criteria are used by clinical practitioners to visually inspect lesions, which is followed by histological examinations. To automate the classification process, several artificial intelligence-based techniques have been developed, including pre-processing, feature extraction, segmentation, and classification. Several classification algorithms [10,11] were mostly relied on constructed feature sets, which had limited generalization power for dermoscopic skin images, due to a full grasp of biological patterns.

Deep learning algorithms offer the advantage of being able to classify data without any pre-processing. Deep networks are more efficient at calculating detailed features for precise lesion categorization than shallow networks [4]. The final layer (Softmax layer) is used for categorization and has 1000 neurons (1000 classes).

**System Architecture**

![System Architecture Diagram]

**Dataset**

The International Skin Imaging Collaboration (ISIC) is an academic industry partnership aimed at making digital skin imaging more accessible to help reduce melanoma mortality. ISIC aspires to fulfill its goals by creating and supporting digital skin imaging standards, as well as partnering with dermatology and computer science groups to improve diagnosis. Skin lesion imaging's quality and utility are now jeopardized due to a lack of dermatologic imaging standards. ISIC is developing suggested standards to increase the quality, confidentiality, and interoperability of digital skin images. ISIC is working on tools for the dermatological and computer science communities, including a growing open-
source public access collection of skin photos. https://www.isic-archive.com/#/topWithHeader/onlyHeaderTop/apiDocumentation. The public can utilise this image collection for educational purposes, research, and the creation and testing of diagnostic AI algorithms. ISIC engages stakeholders through meetings, publications, conferences, and the sponsorship of artificial intelligence Grand Challenges.

**Hadoop Distributed Densely Connected Convolutional Network (HdiDenseNet)**

Transfer training is the application of a pre scientific model to a new problem. Because it allows you to train deep neural networks with relatively little data, it has become very popular in the deep learning industry. There are certain pre-trained machine learning models that have gained a lot of traction. Although the objects you have learned to identify are significantly different from photographs of skin diseases, the feature detection part of such a pretrained model is often reused to classify completely different images. DenseNet-161 was first chosen as a good reusable model because it outperforms the others on ImageNet and has an intriguing design, with each layer using all previous feature maps as input. Furthermore, despite the large number of layers, the time required to train an era is comparable to simpler pre-trained models. DenseNets is a type of Convolutional neural network that uses dense blocks to create tight connections between layers that connect all layers (with corresponding dimensions in the feature map) directly. To preserve the nature of redirection, each layer receives additional input from all previous levels and transmits its own performance maps to all subsequent layers. Tight connections, also known as fully connected connections, are a deep neural network method that uses a linear operation to connect each input to each output by weight.

To summarize, the DenseNet architecture maximizes the residual mechanism by connecting each layer (within the same dense block) to its succeeding levels. Because the acquired features are all shared through a common knowledge, the model's compactness makes them non-redundant. Because of an implicit deep supervision where the gradient flows back more quickly thanks to the short connections, it is also significantly easier to train deep networks with dense connections.

\[ x_l = H_l(x_{l-1}) \]

The ResNets continued this comportment including the skip connection, reformulating this equation into:

\[ x_l = H_l(x_{l-1}) + x_{l-1} \]

The initial distinction between DenseNets and ResNets is made here. DenseNets merges the layer's output feature maps with the input feature maps instead of adding them. As a result, the equation is reformulated
When the size of the feature maps differs, this clustering is impossible, which is the same difficulty we encountered with ResNets. It doesn’t matter if the pooling is a summation or a concatenation, if we make $H_l$ produce $k$ feature maps each time, we can summarize for the $l$th layer:

$$k_l = k_0 + k * (l - 1)$$

The hyper parameter $k$ represents the growth rate. The growth rate determines the amount of data that is added to the network at each level.

The composite function, for each Convolution block in the network representations in the article (and in the blog) corresponds to the operation of

$$x_l = H_l([x_0, x_1, ..., x_{l-1}])$$

Solid blocks are used to represent the concept of tight connections. There are $n$ thick layers in a dense block. The dense chain connects these dense layers with each dense layer, capturing characteristic maps of the previous layers and transmitting them to the following layers. In a solid block, the dimensions of the elements (width and height) remain constant.
The Dense Layer, every dense layer is made up of two Convolutional operations: 1 X 1 CONV (which stands for typical Convolutional function for obtaining features) and 3 X 3 CONV (which stands for reducing feature depth).

Transition Layer At the end of each dense block, the number of feature maps is accumulated to the value of the input features, by adding the number of dense layer multiples of the growth rate. The following components make up the transition layer:
For the transition layer/block, use 1 X 1 CONV and 2 X 2 AVG POOL operations. The procedure 1 X 1 CONV reduces the number of channels by half. The width and height of the features are downscaled using the 2 X 2 AVG POOL layer. The DenseNet-161 image classification model is component of the DenseNet family of prototypes. The size and accuracy of the densenet-121 prototype differ significantly from the densenet-121 prototype. The densenet-161 is significantly larger than the densenet-121, weighing in at 100MB versus the densenet-121’s 31MB.

**Results and Discussions**

In this era of research, propose the use of Hadoop Distributed Narrow Convolutional Network (HdiDenseNet) techniques to detect melanoma skin cancer and examine suspicious lesions linking each layer to the others in a prospective approach. Feature maps from all previous levels are used as inputs to each layer, and their own feature maps are used as input to all subsequent layers. The ability to overcome the vanishing gradient problem, improve function propagation, increase function reuse, and significantly reduce the number of parameters are all attractive aspects of DenseNets. In most of them, DenseNets exceeds the state of the art and consumes less memory and calculations.

<table>
<thead>
<tr>
<th>Model</th>
<th>Crop</th>
<th>Accuracy</th>
<th>ROC AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>DenseNet</td>
<td>1</td>
<td>0.755</td>
<td>0.909</td>
</tr>
<tr>
<td>Xception</td>
<td>1</td>
<td>0.798</td>
<td>0.908</td>
</tr>
<tr>
<td>PNASNet</td>
<td>1</td>
<td>0.782</td>
<td>0.921</td>
</tr>
<tr>
<td>DenseNet</td>
<td>5</td>
<td>0.757</td>
<td>0.911</td>
</tr>
<tr>
<td>Xception</td>
<td>5</td>
<td>0.817</td>
<td>0.912</td>
</tr>
<tr>
<td>PNASNet</td>
<td>5</td>
<td>0.783</td>
<td>0.925</td>
</tr>
<tr>
<td>DenseNet</td>
<td>10</td>
<td>0.755</td>
<td>0.914</td>
</tr>
<tr>
<td>Xception</td>
<td>10</td>
<td>0.817</td>
<td>0.913</td>
</tr>
<tr>
<td>PNASNet</td>
<td>10</td>
<td>0.783</td>
<td>0.925</td>
</tr>
<tr>
<td>DenseNet</td>
<td>20</td>
<td>0.753</td>
<td>0.914</td>
</tr>
<tr>
<td>Xception</td>
<td>20</td>
<td>0.822</td>
<td>0.911</td>
</tr>
<tr>
<td>PNASNet</td>
<td>20</td>
<td>0.785</td>
<td>0.927</td>
</tr>
</tbody>
</table>

Deploying the three models with varied crops [1, 5, 10, 20] in terms of accuracy, ROC AUC, and response time are shown in table (2): 7.36 milliseconds The mean ROC AUC score rises to 0.944, outperforming the first score of DenseNet's initial challenge of 20 crops, which was 0.914.
In comparison to other models, DenseNet achieved consistent ROC AUC accuracy on 20 crops, i.e. 0.914 in Figure (7). DenseNets have a variety of appealing qualities, including the capacity to explain the vanishing gradient challenge, develop quality transmission, recycle growth features, and diminish the number of parameters significantly. On most of them, DenseNets outperform the current state-of-the-art while consuming less memory and compute. A total of about 40,000 pictures were utilized in the testing, which were acquired from a variety of sources. According to the findings, the highest performing Distributed Densely Connected Convolutional Network approach for Early Melanoma Detection using Patient's Metadata of Dermoscopic Images wins high accuracy and Area Under Curve ratings (AUC).

**Conclusion**

This research study uses computer diagnostics to detect skin cancer and is an effective way to screen for skin cancer. Detecting skin cancer by hand is not only time consuming but also exhausting. Biopsy is a common method to detect skin cancer. As a result, computer-aided diagnostics are needed to deal with difficulties. The skin image is pre-processed in this system and then segmented using a segmentation technique followed by feature extraction. Convolutional networks with fewer connections between the input layers and those close to the output can be significantly deeper, more accurate and more effective for training, according to a recent study. To confirm this finding, this study presents the Hadoop Distributed Densely Connected Convolutional Network (HdiDenseNet) techniques for the detection of melanoma skin cancer and the study of suspicious lesions, which directly connect each layer to each other. The feature maps at all levels above are used as inputs to each layer, and their own feature maps are used as input to all subsequent layers. It creates direct links between every two layers of the same size in the feature map, demonstrating that DenseNets can scale up to hundreds of layers organically without the need for optimization. DenseNets significantly improves accuracy as the number of parameters in our tests increases with no symptoms of readjustment or loss of performance. Under varying circumstances, it has provided cutting-edge performance on a number of highly competitive data sets. Furthermore, DenseNets require far fewer parameters and computational resources to obtain state-of-the-art results. The
best-performing distributed tightly coupled convolutional network technique, which uses patient metadata for dermoscopic images for early detection of melanoma, obtained high precision and results for area under the curve, according to the findings (AUC). Future enhancements will include the introduction and combination of two or more models for greater accuracy than currently available, as well as the use of the Hadoop ecosystem and the AWS Cloud to automate the system.

References


Dan Popescu, Mohamed El-Khatib, Hassan El-Khatib and Loretta Ichim , New Trends in Melanoma Detection Using Neural Networks: A Systematic Review


Ms. Gaana M, Dr Shweta Gupta, Dr Narayana Swamy Ramaiah, Diagnosis of Skin Cancer Melanoma using Machine Learning.


