Journal of Namibian Studies, 34 S2(2023): 1930–1941 ISSN: 2197-5523 (online)

Melanoma Skin Cancer Detection by using Ensemble Model

Rosey Chauhan [1], Sunil Gupta [2]

[1] Jaipur National University, Jaipur, India
errose.it@gmail.com

[2] Jaipur National University, Jaipur, India
sunilg95@rediff.com

Abstract

Skin cancer is the abnormal development of skin cells that occurs most commonly on sun exposed skin. This type of cancer can also develop on parts of your skin that aren't often exposed to the sun. A convolutional neural network is a type of deep neural network that is most typically used to evaluate visual imagery in deep learning. The main goal of this model is to detect skin cancer at an early stage so that it may be treated. This model is designed to distinguish between the seven forms of skin cancer. The major goal of this model is to increase the accuracy of existing models, with a particular focus on melanoma skin cancer detection because it is the worst type of skin cancer and failing to detect it early can result in death. If one notice a change in the skin, it could be a new growth, a scar that refuses to heal, or an uneven mole. There is an ABCDE rule for melanoma detection that helps dermatologists as well as the general public identify it. In this model there are numerous models, including Inception, Dense Net, and others. These are pre-trained models, and by fine-tuning the top layers, as well as all of the layers, to see how accurate they are. Finally, there is a combination of the two best models into an ensemble model to improve the system's accuracy and range even further.

Keywords: Convolutional Neural Network(CNN), VGG 16, AlexNet, ResNet.

Introduction

There are seven categories of cancer that are named as Actinic keratoses and intraepithelial carcinoma / Bowen's disease, basal cell carcinoma, benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichenplanus-like keratoses), dermatofibroma, melanoma, melanocytic nevi, and vascular lesions are all represented in the dataset (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage).

To develop hierarchical representations of data, the Deep Learning approach combines many processing layers. It provides a technique to harness a huge amount of data with a small number of people involved in feature engineering. Starting with AlexNet in 2012, the Deep Learning approach has made significant advances and evolution in Computer Vision in recent years. The image classification challenge is a fairly broad problem that encompasses any work that requires discriminating between photographs of various items. Deep CNN is distinguished by the fact that the earliest layers of the network often learn relatively generic and "low-level" image features, while the very final levels of the network learn semantics and high-level features. As a result, Deep Convolutional Neural Networks trained on one dataset for image classification tasks can be utilised for image classification tasks on various datasets with fine-tuning. As a result, fine-tuning has become popular in computer vision research. In "Dermatologist – level classification of Skin Cancer with Deep Neural Networks," Esteva et al claimed that " By optimising InceptionV3, CNN demonstrates artificial intelligence capable of classifying skin cancer with a level of competence equivalent to dermatologists, performing on par with all tested professionals. There are 2032 skin disorders in total in this dataset, divided into nine skin disease partitions. Esteva et al attained up to 66 percent accuracy classification on nine classes by fine-tuning InceptionV3 on this dataset.

Cancer is a disorder when the body's cells proliferate abnormally and out-of-control. [1]Cancer forms when the body's normal regulatory system fails. Old cells do not perish; rather, they expand uncontrollably, giving rise to new cells (abnormal cells). These extra cells might come together to form a tissue mass known as a tumour.

Leukaemia is one type of cancer that only results in the growth of tumours[2]. Any area of the body may experience it. Even if there are many skilled physicians available, people suffer more in developing countries. [3] People often mistake any skin lesion for a common virus or anything else because they are ignorant of the condition. As a result, they overlook the lesion, which causes it to spread and potentially be fatal. Therefore, there are technologies like DCNN that can identify skin cancer just by providing the machine a picture of the condition [4].

Related Work

Hamd and Asa were able to forecast skin cancer by examining pigment. They used a computational method based on symmetric colour for skin lesions' pigments [5]. Artificial Intelligence and Signal Processing (AISP) was first used to identify and segment the lesion's margin. The symmetry level for each picture was evaluated in order to differentiate benign tumours. Then, symmetric parameters were chosen in order to calculate. In the images of skin lesions, melanoma, basal cell carcinoma, and squamous cell carcinoma were recognised. Two adaptation techniques for skin lesions were developed in this work. The earliest method of adaptation was using reddish, yellow, brown, and black colours or synthetic spectrums.

The second method included adjusting utilising a database of healthy and unhealthy colours. The comparison of the outcomes involved 40 photographs altogether. Classification results for the first and second approaches were 80 and 92.5 percent, respectively. The results will be better with more segmentation of skin lesions or with a database of pigments and spectra.

Skin cancer was identified by Abdol-al Jalil et al using a neural network [6]. They used software for image processing and artificial intelligence. In order to improve picture quality and reduce noise, the skin photographs underwent preprocessing. They were segmented following the thresholding process. The distinctive characteristics of skin cancer were extracted using the 2-D wavelet transform. The retrieved characteristics served as input for the neural network. A back-propagation neural network was used by the researchers to divide the dataset into cancer and non-cancer categories [12. The results generally indicated that the data classification accuracy was excellent. A technique based on dermoscopy was proposed by Celebi and associates to identify skin cancer [7]. The JSEG technique was originally applied in [8] to identify lesion borders. Colour and texture were extracted using GLCLM, and malignant lesions were subsequently identified using SVM classification. The results of this method's testing on a database of 655 dermoscopic pictures revealed that the system was 90% accurate.

A method that employs a quick median filter as a preprocessor was reported by Abdul Jaleel et al. in their study [9]. GLCM was used to extract contrast, correlation, and asymmetry properties after the lesion was removed from the picture using thresholding and entropy maximum methods. Finally, a forward feed neural network was used to perform a classification procedure. The accuracy of the algorithm was discovered to be 88 percent.

Barata et al. [10] suggested a method for identifying skin cancer based on regional and universal features. This study sought to identify the ideal diagnostic strategy (both local and global), as well as the ideal traits (colour and texture). Three distinct classifiers—KNN, SVM, and Adaboost—as well as different combinations of attributes were used in the research. The results showed that a simple classifier like KNN may be used to get the desired outcome. In addition, the colour feature operates better than the texture feature. Only a few characteristics are needed to classify with high accuracy and improve system generalisation. The results showed that both procedures could be used to get the intended outcome, however the local strategy required more time.

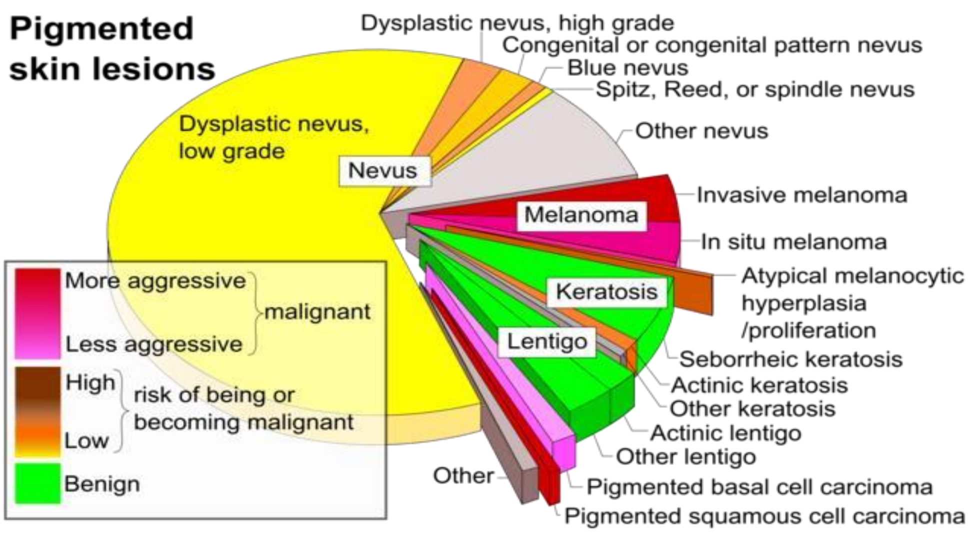
Proposed Work

The following problem statements in relation to skin cancer are the focus of this research: -Classification of skin cancer images into appropriate 7 categories: Melanocytic nevi, melanoma, benign keratosis-like lesions, basal cell carcinoma, actinic keratoses, vascular lesions, and dermatofibroma are all examples of melanocytic nevi. Specifications of the Dataset For the issue statement, HAM10000 and ISIC 2018 Dataset used. For the purpose of classification: The 10000 images are divided into 7210 training instances, 1803 validation examples, and 1002 test examples after being normalized by dividing by 255.

A. Dataset

HAM 10000 dataset and ISIC dataset used for the skin lesion images. 2000 cancer cases and 8000 benign samples make up the dataset. With almost 7000 occurrences, melanocytic nevi are overrepresented in the collection. As a result, the designed neural network model ought to attain an accuracy of more than 60% even in the worst-case scenario. Five samples of each type of skin lesion are shown in Figure 1. It's challenging for non-experts to differentiate between the two types. The original images are of large size (450 x 600), scaled them to 64x64 RGB images for the baseline model and 192x256 for fine-tuning models. The dataset is divided into 7210 training examples, 1803 validation examples, and 1002 test examples after being normalized by dividing by 255.

Fig1: Pie Chart of Pigmented skin lesion



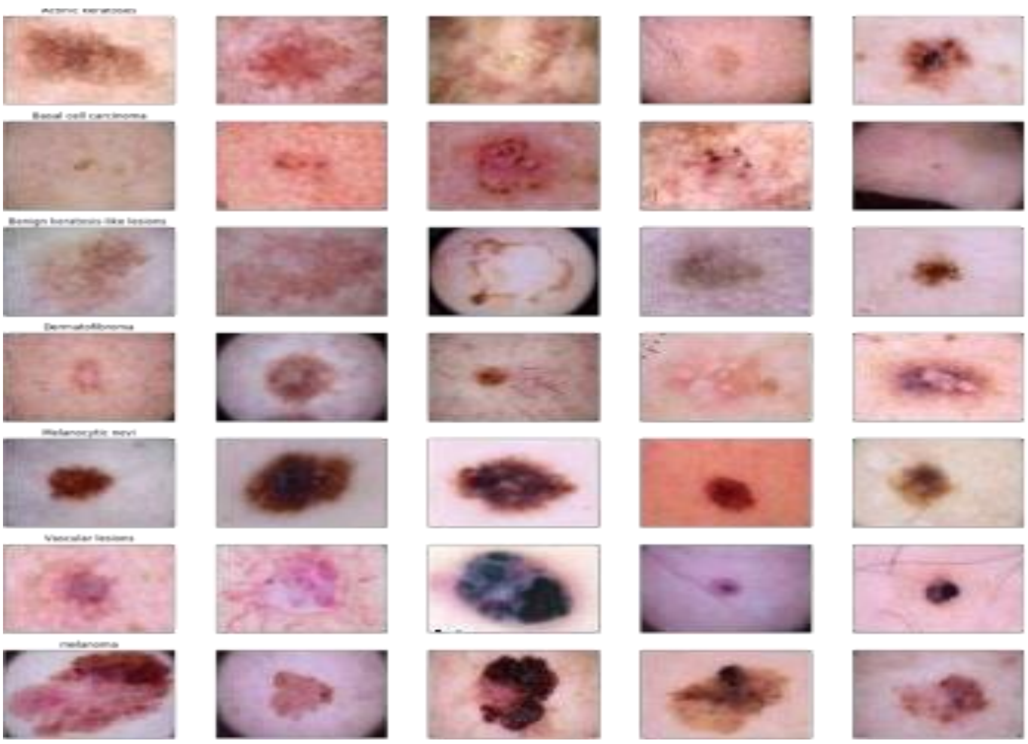
B. Baseline Model

Before fine-tuning DCNNs, create a small CNN to simulate the difficulties of classifying skin lesions. The developed CNN's architecture for this research is as follows: A convolutional layer with 16 kernels of size 3 was applied first, along with padding to maintain the same picture size. Secondly, a 2x2 window maximum pooling layer. As a result, the size of the spatial activation is reduced by 2x in the feature maps. In next step, a convolutional layer with padding to maintain the size uniform and 32 kernels of size 3 each. A maximum pooling layer with a 2x2 window comes in next stage. As a result, the size of the spatial activation is reduced by 2x in the feature maps.

The fifth step is a convolutional layer with 64 kernels, each of size 3, with padding to maintain uniform size. A max pooling layer with a 2x2 window makes up the sixth layer. As a result, the size of the spatial activation is reduced by 2x in the feature maps.[13]

Heuristics are the foundation on which this model is formed. I employ the shortest (3x3) convolutional layers and increase the number of filters in the output anytime the spatial activation size is reduced in order to maintain roughly constant hidden dimensions. In order to train this model, more data is needed. The goal of this approach is to ensure that the model never sees the same image again by slightly changing the training dataset at the beginning of each session. The Adam optimizer is used, and the learning rate is set at 0.01. Learning rate decay is used to cut the learning rate in half when the validation accuracy peaks after three epochs. The baseline model is trained for a total of 35 epochs.

Figure 2: Skin Cancer Images



C. VGG16

Despite the fact that several DCNNs models perform better on ImageNet than VGG16, previously chose to improve VGG16 because to its simplicity. The VGG16 D schema is shown in Figure 3. Similar to the top performing VGG16 net, the third, fourth, and fifth convolutional blocks all have four convolutional layers. On ImageNet, VGG16 has a top-5 accuracy of 0.901 and a top-1 accuracy of 0.713.

Figure 3: VGG16 Model

The top fully-connected layers of VGG16 are eliminated, and new fully-connected layers are added for our classification tasks (including one global max pooling layer, one fully connected layer with 512 units, one dropout layer with 0.5 rate, and one softmax activation layer for 7 types of skin lesions). To prevent the weights for these layers from being entirely random and the gradient from being too big when we start fine-tuning, first freeze all layers in VGG16 and do feature extraction for the just added FC layers. Unfreeze the last convolutional block of the VGG16 after 3 epochs of feature extraction, then start finetuning the model for 20 epochs. Adam optimizer and learning rate of 0.001 are applied throughout the training procedure. The baseline model's data augmentation and learning rate degradation strategies are used. VGG16 underwent 30 epochs of finetuning[14].

D. Inception

On ImageNet, Inception V3 had the best results, scoring 0.779 for top-1 and 0.937 for top-5 accuracy. The Inception modules, which are essentially little versions of the larger model, are what give the third version of Inception its name. The concept stems from the notion that each layer must decide what kind of convolution it wants to make: Do you want a 33? Or a 5×5? The notion is that you don't have to know in advance if it would be preferable to do, for instance, a 33 or a 55. Simply do all the convolutions and let the model to choose what is best. The model can recover both local characteristics through smaller convolutions and highly abstracted information through bigger convolutions thanks to its design. Because the bigger convolutions need more work, [4] advises using an 11 convolution to reduce the feature map's dimensionality, followed by sending the resultant feature map through a ReLU, before performing the larger convolution (in this example, a 55 or 33). The 11 convolution is crucial since it will be applied to lower the feature map's dimensionality.

Fig4: Inception V3

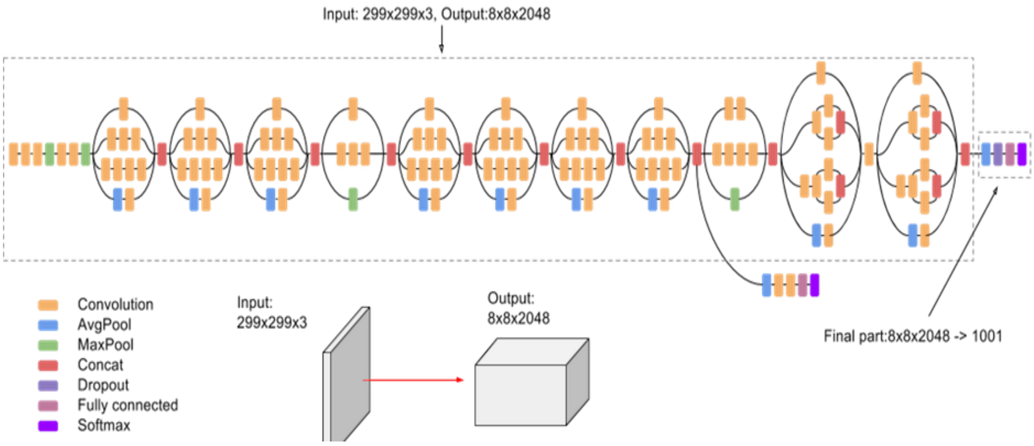
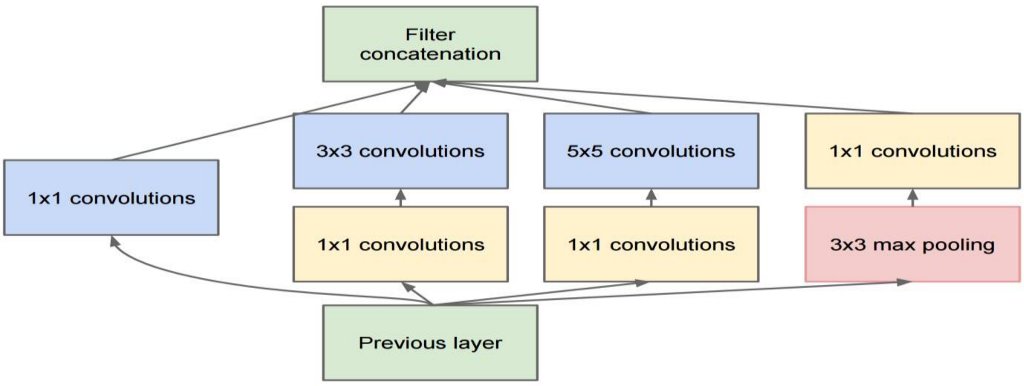


Fig5: Inception V3 Block Diagram



Perform two different types of experiments using Inception V3, which was learned on ImageNet using 11 inception blocks: Inception V3's last two inception blocks should be fine-tuned, followed by the entire pretrained model. The reason for this is because Keras has implemented Batch Norm. The following describes how Keras implemented Batch Norm. Whether the BN layer is frozen or not, the network will always utilize the mini-batch statistics during training, and it will also use the previously learnt statistics of the frozen BN layers during inference. The weights of the top layers will thus be updated to the mean/variance of the new dataset if fine-tune them. However, because the mean and variance of the original ImageNet dataset will be utilized, they will get values that are scaled differently during inference. Therefore, it will give extremely poor validation accuracy if it is followed precisely the fine-tuning VGG16 with InceptionV3.

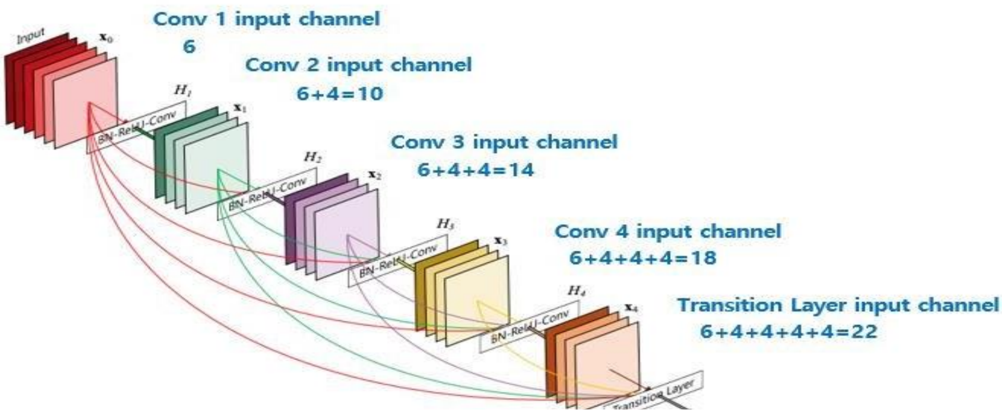
The statistics of the mini-batch from the training data set will be used by the batch normalization layers during inference as a temporary fix for this problem if all layers are set to trainable. To fine-tune Inception V3's top two blocks as well as all of its layers, with all of the model's batch normalization layers set to trainable. These 35 epochs make up this experiment. It took 20 epochs to perfect the whole Inception V3 system.

Inception- ResNet is another kind of Inception that excels on ImageNet. Residual connection, which is proven to be essentially important for training very deep neural networks, is incorporated into Inception ResNet. The overall architecture of this model is shown in figure 6, and the architecture of the Inception-ResNet-C module, which will be adjusted, is shown in figure 7. Inception V3, Inception-ResNet V2, and VGG16 were tuned using the same training method. Inception-ResNet V2's top layers were tweaked for 30 epochs.

E. DenseNet

Dense Net, a newly released DCNN architecture, ranks among the best on ImageNet with a top-1 ranking of 0.936 and a top-5 ranking of 0.773. Although Dense Net performs similarly to Inception V3 in terms of performance, it contains less parameters (20 million as opposed to 23 million for Inception V3). Figure 8 shows the overall layout of a dense block, one of the four dense blocks that make up Dense Net 201. In a dense block, all successive layers L - l receive the featuremaps from all preceding convolutional blocks as inputs, and the lth layer also receives its own feature-maps. Every layer writes to the layer after it and reads the state from the ones before it. It alters the situation while simultaneously transmitting important information. By concatenating features rather than summing them as in ResNet, Dense Net design explicitly distinguishes between information that is contributed to the network and information that is kept.

Figure 6: DenseNet Model



One layer in a dense block produces feature maps using a composite function that combines batch normalisation, ReLU, and a 3x3 convolution. Between each dense block, there is a layer of convolution and a layer of pooling, together known as a transition layer. Four dense blocks make up Dense Net 201, and it will use for two different kinds of experiments: [15]The final dense block, which contains 32 layers, should be fine-tuned first, followed by the entire network. Dense Net is tuned using the same training method as in earlier sections. DenseNet 201's top layers underwent a total of 27 epochs of fine-tuning, while the entire network underwent 20 epochs of fine-tuning. It initialised the weights as the original weights from the pretrained model on ImageNet when fine-tuning the whole Dense Net 201 and Inception V3 model[16].

F. Ensemble Model

While fine-tuning all layers takes less time overall than fine-tuning only the top layers, it produces results that are superior to those obtained when fine-tuning only the top layers. This is because the model only worked for 20 epochs while finetuning all layers, but for 30 epochs when fine-tuning the top layers. The findings wouldn't be as good if you merely finetuned the top layers for fewer than 30 epochs. This finding suggests that fine-tuning the entire model leads to better final results and speeds up convergence compared to only the top layers. Dense Net 201 provides the best single outcome in both scenarios, which is astounding considering that this model has even less parameters than Inception V3. thick Net is an extremely thick, deep model with minimal parameters, as stated in [5]. The effectiveness of DenseNet 201 in this experiment confirms the validity of employing DenseNet 201 for transfer learning on a dataset from a totally new domain that was pretrained on ImageNet. I used ensemble learning to generate an ensemble of the previously fully-tuned Inception V3 and DenseNet 201 models, and I got the best results with 88.8% accuracy on the validation set and 88.52% accuracy on the test set.

Performance Evaluation

Better results are obtained when all layers are fine-tuned than when only the top layers are, and fine-tuning all layers takes less time overall. The reason for this is that while fine-tuning all layers only lasts for 20 epochs, it lasts for 30 epochs when finetuning the top layers. The findings wouldn't be as good if you merely fine-tuned the top layers for fewer than 30 epochs. This finding suggests that fine-tuning the entire model will improve the final product and speed up convergence compared to merely fine-tuning the top layers. Dense Net 201 provides the best single outcome in both scenarios, which is astounding considering that this model has even less parameters than Inception V3. A particularly thick and deep model with few parameters is called a dense net. The effectiveness of DenseNet 201 in this study confirms the validity of utilizing DenseNet 201 for transfer learning on a dataset from a totally new domain that was pretrained on ImageNet. I used ensemble learning to construct an ensemble using the previously fully-tuned Inception V3 and DenseNet 201, and I got the best results with 88.8% for the validation set and 88.52% for the test set. By using the techniques of transfer learning and ensemble learning, it will be possible to put together a team of optimized DenseNet 201 and Inception V3 that obtained 88.52% accuracy on the test set and 88.8% accuracy on the validation set for HAM10000.

Table1: Comparision of results of Ensemble Model

Model	Validation	Test	Test Loss	Depth	# Params
Baseline Model	77.48%	76.54%	0.646671	11 layers	2,124,839
VGG16	79.82%	79.64%	0.708	23 layers	14,980,935
Inception V3	79.935%	79.94%	0.7482	315 layers	22,855,463
Inception ResNet V2	80.82%	82.53%	0.6691	784 layers	55,127,271
DenseNet 201	**85.8%**	**83.9%**	0.691	711 layers	19,309,127

Conclusion

Inception V3 and DenseNet 201 were combined in this study using the techniques of transfer learning and ensemble learning, and they were able to obtain 88.52% accuracy on the test set and 88.8% accuracy on the validation set for HAM10000 (a subset of ISIC 2018). It has also been discovered through tests that, for this dataset, fine-tuning the entire model results in superior final results and speeds up model convergence.

Overfitting is one severe issue that has been noted throughout training. The training data in every study trial was overfit by 10% to 13%. Many techniques are employed to reduce overfitting, but they were unable to reduce it any more. The models will improve when more is done in the future to prevent overfitting and develop better training methods.

Future Work

As a follow-up to the experiment, researchers investigated ways to more actively identify lesions in the image itself and do the categorization in real-time for easy usage by dermatologists and the individual who suspects they may have skin cancer. We suggest employing segmentation with masking for this, which may be done by wellknown pre-trained CNNs like AlexNet by instructing it to do segmentation tasks. Microsoft's proposal to apply MASK-RCNN on the COCO dataset yielded the best results.

Bibliography

1. “Computer aided Melanoma skin cancer detection using Image Processing” (Shivangi Jain , Vandana Jagtap, Nitin Pise) at International Conference on Intelligent Computing, Communication & Convergence (ICCC- 2015)
2. “Deep Learning Can Improve Early Skin Cancer Detection” (Abeer Mohamed, Wael A.Mohamed, and Abdel Halim Zekr) in 2018.

3. “The Melanoma Skin Cancer Detection using Ensemble Model using Support Vector Machine” (Hiam Alquran , Isam Abu Qasmieh, Ali Mohammad Alqudah, Sajidah Alhammouri, Esraa Alawneh, Ammar Abughazaleh, Firas Hasayen)in 2017.

4. “Skin Lesion Classification Using Hybrid Deep Neural Networks” (Amirreza Mahbod, Gerald Schaefer, Chunliang Wang, Rupert Ecker, Isabella Ellinger Institute for Pathophysiology and Allergy Research, Medical University of Vienna, Austria Department of Research and Development, TissueGnostics GmbH, Austria Department of Computer Science, Loughborough University, U.K. Department of Biomedical Engineering & Health Systems, KTH Royal Institute of Technology, Sweden) in 2019.

5. “Melanoma skin cancer detection using color and new texture features” (Farzam Kharaji Nezhadian, Saeid Rashidi Faculty of Biomedical Engineering Islamic Azad University, Science and Research branch Tehran, Iran) in 2017.

6. “Classification of Melanoma Skin Cancer using Convolutional Neural Network” (Rina Refianti, Achmad Benny Mutiara , Rachmadinna Poetri Priyandini Faculty of Computer Science and Information Technology, Gunadarma University)

7. “Malignant Melanoma Classification Using Cross- Platform Dataset with Deep Learning CNN Architecture” (Soumen Mukherjee, Arunabha Adhikari and Madhusudan Roy)

8. Dumitru Erhan, Christian Szegedy, Alexander Toshev, and Dragomir Anguelov. Scalable object detection using deep neural networks. In Computer Vision and Pattern Recognition, 2014. CVPR 2014. IEEE Conference on, 2014.

9. Andrew G. Howard. Some improvements on deep convolutional neural network-based image classification. CoRR, abs/1312.5402, 2013.

10. Elgamal, M. Automatic Skin Cancer Images Classification. IJACSA 2013, 4. [CrossRef] 4. Key Statistics for Melanoma Skin Cancer. Am. Cancer Soc. Available online: https://www.cancer.org/content/dam/CRC/PDF/ Public/ 8823.00.pdf (accessed on 8 February 2021).

11. Khan, M.Q.; Hussain, A.; Rehman, S.U.; Khan, U.; Maqsood, M.; Mehmood, K.; Khan, M.A. Classification of Melanoma and Nevus in Digital Images for Diagnosis of Skin Cancer. IEEE Access 2019, 7, 90132–90144. [CrossRef]

12. Premier Surgical Staff. What Is the Difference between Melanoma And non-Melanoma Skin Cancer? PSS. Available online: https://www. premiersurgical.com/01/whats-the-difference-between-melanoma-and-non-melanoma-skincancer/ (accessed on 6 February 2021)

13. W.H Starbuck, A. Schwab,” Why Baseline Modeling is Better Than Null-Hypothesis Testing: Examples from International Business Research” Advances in International Management, ResearchGate

14. Liang, Jiazhi. (2020).” Image classification based on RESNET”. Journal of Physics: Conference Series. 1634. 012110. 10.1088/1742-6596/1634/1/012110.

15. M. Shafiq and Zhaoquan Gu," Deep Residual Learning for Image Recognition: A Survey", 2022, 12(18), 8972; https://doi.org/10.3390/app12188972

16. Milton, M.A.A. Automated Skin Lesion Classification Using Ensemble of Deep Neural Networks in ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection Challenge. arXiv 2019, arXiv:1901.10802. Available online: http://arxiv.org/abs/1901.108 02 (accessed on 22 January 2021).