CLASSIFICATION OF SKIN LESIONS USING DEEP LEARNING AND IMAGE AUGMENTATION

[1]Dr. VINOD H C, [2]Dr. SWAMY L N, [3]Dr. PRADEEP S, [4]Dr. MADHU CHANDRA.G

[1]Associate Professor, Department of ISE, SJB Institute of Technology, Bangalore, Pin-560060

[2]Assistant Professor, Department of CSE (MCA Programme), VTU Post Graduate Studies – Mysuru Region.

[3]Assistant Professor, Dept of CSE, Government Engineering College, chamarajanagara- Pin-571313.

[4]Associate Professor, Department of Electronics and Communication Engineering, R L Jalappa Institute of Technology, Doddaballapur

Abstract: Due to the lack of annotated pictures and the imbalance of classes in image sets, it is difficult to classify skin lesions using data-hungry deep learning models. Overfitting may occur when there aren't enough labeled data points or the classes aren't spread uniformly. As an alternative to the current problem, image augmentation may be the answer. Overfitting and the creation of generalizing network models have been combated using picture upgrades and a variety of other strategies documented in the scientific literature. Even if it's not apparent how to choose the best decision, various current deep learning-based approaches for skin lesion categorization and picture enhancement are reviewed in this study.

Keywords: Convolutional neural networks, deep learning, automated diagnosis, skin lesion classification, image processing, data augmentation.

I. INTRODUCTION

Skin problems are often diagnosed by visual examinations of lesions. Due to suspicion, clinical analyses are carried out when the dermatologist cannot make up his or her mind. Relying only on human vision may be erroneous, non-objective and difficult to duplicate even among dermatologists with substantial training. Lesions of diverse shapes, sizes and borders are difficult to see in images because of the varying skin colours and tones [1].

Computer-Aided Diagnosis (CAD) has been used [2,3] by using traditional image processing techniques such as edges, shapes, and regions. Because they depend so largely on special features and prior knowledge, these solutions are unlikely to be widely used.

Deep convolutional neural networks (CNNs) have been utilised to automatically diagnose lesions from images with large data sets and high-level performances [4–7]. Deep network topologies, fully convolutional layers, advanced computations, and simple access to massive data are primarily to blame for this exceptional performance. To improve the depth of feature maps, convolutional layers are used to decrease the spatial resolution of images. With huge data sets, a fully convolutional network has the ability to train features that best suit both the picture's appearance and its meaning hierarchically. Due to a scarcity of data, it is difficult to extrapolate from these models.

Generalization refers to the difference between the performance of neural networks that have seen training images and those that haven't. To train their convolution models, deep convolutional networks need a large number of annotated image sets. A lack of sufficient training data sets that include all possible variables makes it difficult to meet the overfitting criteria of deep fully-convolutional network models [9,10], which leads to erroneous results (such as lesion forms, sizes, or patterns at various patients).

A deep learning model's training and validation errors may be used to determine the model's efficacy. As a result, monitoring the error rate is critical throughout the training and validation phases. Overfitting may be detected by graphing the accuracy of the model's predictions at each epoch of training and validation.

Fig. 1 illustrates the links between training and testing errors, epochs, and accuracy scores in a visual manner When training errors continue to reduce and validation errors begin to climb, this is seen in Fig. 1.a. Due to a decline in testing set performance and a rise in error during the training stage, overfitting develops.

As seen in Fig. 1.b, there is no evidence of the network model being overfitted, indicating that it is useful and that the testing and training errors are about equal. By displaying comprehensive data in pictures, the gap between validation and training sets may be narrowed through image augmentation.

The overfitting problem may be effectively addressed by increasing the size of the data sets by image augmentation. To artificially increase the number of pictures without affecting the semantics of the original images, a variety of techniques are performed [4].

Pictures may be enhanced in a variety of ways (such as by random cropping, scaling, and mirroring) (such as random colour jittering, rotation, shift, shear, flip and translation). Dropout and batch normalisation are two common methods for reducing or eliminating the problem of overfitting in data. This is owing to the fact that image enhancement may be considered a sort of regularisation

Determining which augmentation approach works best is a critical step in network architecture. There is a lack of a mechanism for picking an appropriate augmentation strategy. When it comes to skin lesion diagnosis, there is currently no published data on the most recent image augmentation techniques used. Thus, this paper describes the most recent ways for offering image improvements to help in the identification of skin lesions. Our study, we believe, would be very useful to other researchers in the same area.

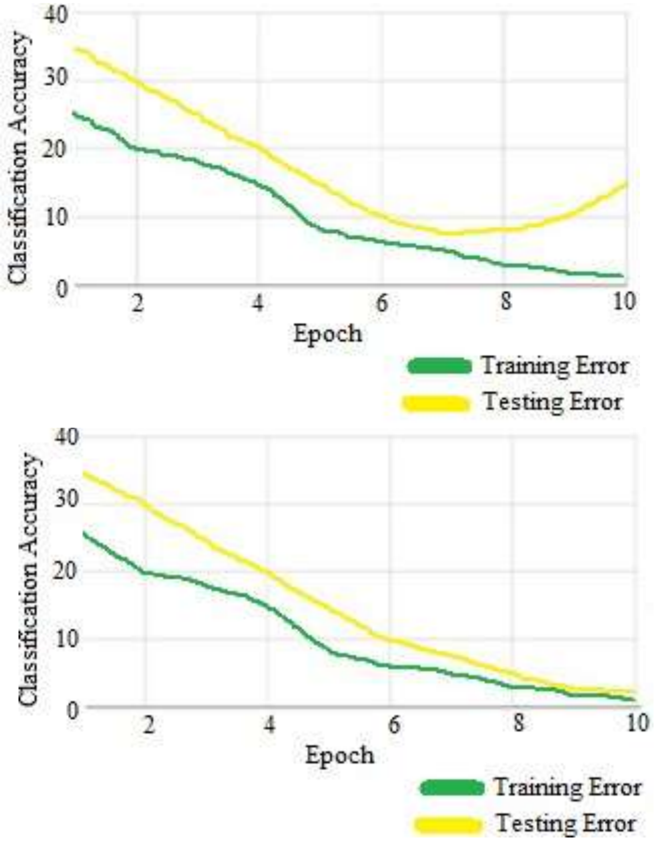


Figure1: Mistakes in training and testing are often intertwined. An indication of overfitting is an increase in validation errors despite a decrease in training errors (a). Testing and training errors in (b) are near to zero, therefore no overfitting is evident here.

The structure of the paper is laid down below. Overfitting problems may be overcome using augmentation techniques other than image augmentation, as mentioned in Section II. In Section III of this study, new image enhancement techniques for skin lesion categorization are discussed. Section IV is where the meat of the paper.

II. SOLUTIONS TO AVOID OVERFITTING PROBLEMS IN DEEP NEURAL NETWORK-BASED APPLICATIONS

To put it another way, there are many strategies to enhance deep neural networks' generalization performance by preventing or reducing overfitting. For the most part, these techniques focus on network design and complexity enhancement for AlexNet [4]-derived architectures, and include approaches like VGG-16[15], ResNet, sInception-V3, and DenseNet.

Transforming a large dataset into weight values and then feeding those weight values into another operation is an example of the transfer learning approach used by ImageNet. This technique is often used in recent classifications of cutaneous lesions. Researchers in examined InceptionV3, ResNet50, DenseNet201, and Inception-ResNetV2 to investigate whether transfer learning was feasible.In the classification of skin images. Skin lesions categorization was improved by the use of transfer learning.

Another way to regularise networks is by the use of "dropout," which involves randomly selecting neurons and setting their activation levels to 0. This is a good technique to pick up a wide range of capabilities. But although while various researchers developed novel dropout approaches for classifying images, they were seldom implemented in existing CNN architectures because of their small performance benefits in many situations . It is preferred to dropout because of the regularisation effect of batch normalisation. Neither the usual element-wise dropout technique nor the dropout of convolutional layers are successful . To avoid a decrease in accuracy, dropout should be used cautiously.

Activation layers may be normalised using the batch normalisation approach in order to boost network generalisation. The batch standard deviation is used to normalise each activation's variance from the batch mean. The authors of employed batch normalisation and dropout in a classification task. With caution, a recent study revealed that batch normalisation and dropout had an impact on CNN deep learning, but only when used carefully. If you have a question, you should utilise batch normalisation. During training, it has a negligible effect on accuracy. The first step in improving CNN performance is to apply batch normalisation.

It is possible to prevent overfitting by using a large dataset for pre-training. Weight values are transferred in the same way as new abilities are learned via transfer. However, unlike transfer learning, pre-training allows you to build your own network from the bottom up. The authors in demonstrate that pre-training a CNN with several image sets may be utilised in both unsupervised and supervised settings. It's common to apply pre-training in more recent studies .

Many techniques (such as image warping or oversampling) employ picture augmentation to solve overfitting challenges (without changing the semantic meaning of the labels). Image augmentation is often used in CNN-based systems (e.g., AlexNet , DenseNet, Inception and ResNet. To develop these network models, researchers drew on the ImageNet dataset, which contains millions of images organised into 1,000 separate categories.

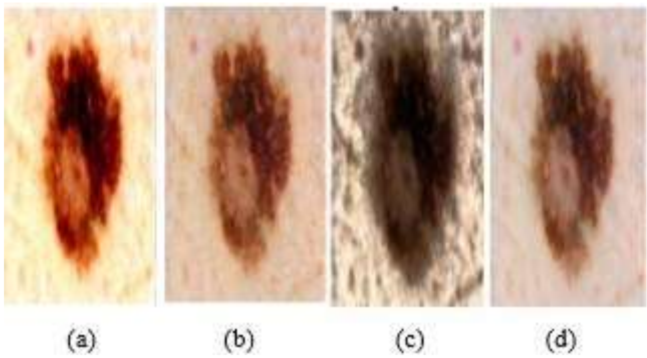
Techniques like as picture augmentation were used throughout the early phases of development. Regions in have been isolated by cropping from all four sides and the centre. Photos with varying pixel densities have also been cropped. These approaches have been used to boost accuracy despite their high production costs.

Public image data sets with annotations are much smaller than those of natural images when it comes to categorizing skin lesions. In the next part, we'll go over some of the image enhancement methods utilized in the deep learning-based classification of skin lesions.

III. IMAGE AUGMENTATIONS TO AVOID OVERFITTING IN SKIN LESION CLASSIFICATION

Image augmentations have been used extensively in deep learning-based skin lesion diagnoses. An example of this type of modification is changing the colour of something. Some of the most common ways for altering colour include boosting contrast and brightness, blurring and sharpening effects, histogram equalisation, and white-balancing. The results of applying colour adjustments to the same dermoscopy image are shown in Figure 2.

Figure 2. The same image after +20% contrast enhancement (a), sharpening (b), histogram equalization (c), white balancing (d)



Rotation, scaling, sharing, and reflection are some of the most often used picture-enhancing methods. Mixing various augmentations has been proven to increase classification performance, according to certain authors. Dermoscopy pictures have been enhanced by recognising cancerous or non-malignant ones using Inception V4. In, an example of a combination operation is shown. This binary categorization includes geometric, colour, and warping augmentations. Expert expertise has led to the employment of geometric augmentation and warping procedures like as cropping, scaling, distortion, and flipping (horizontally and vertically). independent of its location or size, a lesion retains its semantic importance. The bending and geometric augmentation have no effect on the final categorization of the lesion. Images acquired with different cameras and systems may be unified by employing colour augmentation. A high degree of accuracy was achieved by using 6762 training pictures and 600 testing photographs, however the sensitivity was just average. It is essential that researchers fine-tune the final layer of the network in order to utilise the network weights learned from more than one million images.

Using ResNet152, DenseNet161, and InceptionV4, researchers in investigated how to categorise melanomas using three alternative CNN architectures. If you want to change the colour, brightness, or contrast of an image, you may use affine transformations, random erasure, or turning it upside down. Random cropping, affine transformation, and flipping with additional augmentations such saturation adjustment, brightness, contrast, and hue have demonstrated the greatest effects when combined. These models' regularisation features reduced the impact of their improvements over Inception-v4-enhanced pictures compared to photos improved using ResNet152 or DenseNet161.

In order to produce better photographs, several techniques including colour improvement, warping while maintaining symmetries, and geometric modifications (such as rotation and turning) have been used. Tests of the system's performance were conducted on 75 cancerous photos and 304 non-malignant images using GoogleNet. Improved photos, according to the results, may help increase classifier invariance with respect to biological pattern.

Other image enhancement techniques like shifting and rescaling have been shown in to be used to approximate soft tissue oscillations by using a nonlinear distortion. Deep learning and traditional computer vision methods have been used to classify lesions. There were 379 testing photos and 900 training images. Analyzed by eight dermatologists, the classification approach was judged more accurate (76 percent vs. 70.5 percent)

Rotation and translation have been used to enhance the images . This project also makes use of transfer learning. Using a customised version of GoogleNet, melanoma was accurately identified 99.29% of the time, according to the study findings.

When employing GANs and the minimum-maximum approach, unsupervised picture generation may be achieved . GANs combine the Generator Gntr(z) and Discriminator Dscr(z) adversarial network models into a single model. You may create a photo-realistic picture with a generator. This technique is used to distinguish between actual and fabricated photos. Discrimination networks improve functions, while generators lower cost functions (F(Dscr,Gntr)) such as the greatest probability (F(Dscr,Gntr)).

The authors in developed a progressive GAN for image enhancement based on self-attention. Long-distance relationships may be described using the self-attention characteristic. According to the authors, the upgraded version of the test had a 70.1 percent accuracy rate, which is 2.8 percent higher than the baseline version (67.3 percent).

In order to improve data quality, recommends using GANs. After adjusting the generator's noise and style control layers, a GAN model generates pictures of skin lesions. As a part of the classification procedure, transfer learning was employed to build a pre-trained deep network. In order to generate fresh pictures for our training set, we used the GAN model. Sensitivity and specificity were improved as compared to the CNN design (95.2 percent, 83.2 percent, and 74.3 percent, respectively).

As an example, basal cell carcinoma, malignant melanoma, vascular lesions, and actinic keratosis were all categorised by into seven distinct types of lesions. The Rectified Linear Unit (Leaky-Rectified Linear Unit) has four deconvolutionary (convolutionary) layers that are used in this study's generator (discriminator). tanh is used to activate a deconvolutional layer in the discriminator and a full-connected softmax layer in the generator. The GAN model's images have undergone three typical adjustments (Gaussian blur, random cropping, and salt and pepper noise) in order to improve generalisation. It was discovered that the average F1-score was 0.83. The highest possible F1 rating. Melanocytes and actinic keratosis have the greatest and lowest F1-scores, respectively, on the skin (F1- score:0.71).

The effects of picture augmentation and up-sampling were studied by researchers in using MobileNet. Without the aid of up-sampling or augmentation, this network was formed After then, up-sampling and augmentation were used to bring the network up to date. Adding a picture enhancing step really enhances accuracy, contrary to common assumption.

IV. CONCLUSIONS

The lack of annotated images makes it difficult to use deep learning to classify skin lesions. Efforts to reduce overfitting are essential in this business. Some frequent picture editing techniques include cropping, rotating, scaling, mirroring, and shimmying. Other affinity conversions, such as colour alterations, are also often used. You may quickly and easily increase the number of photos in your training materials. However, the problem is that they effectively rehash the functions that already exist. Thus, they are unable to develop new learning tools. Models of networks and algorithms cannot generalise without these tools. Adding and removing random white noise from photos may also help networks become more resilient. These approaches alter the look of the picture to make it more readable to the human eye. Some new knowledge may be generated by a competent deep learning model, although this does not always happen.

Even while GANs seem to be effective in expanding the amount of training picture sets, it is unclear how generator models are quantitatively assessed. Using cost functions may have had a detrimental or positive effect on the examples' quality in this study. Non-realism may result from this, according . GANs have the potential to lengthen the time it takes to analyse data. Image augmentations have been widely employed to overcome the issue of data scarcity in deep learning-based skin lesion detection. A thorough evaluation of the pros and disadvantages of different methods of improvement is still required.

REFERENCE

[1] P. Tschandl, C. Rosendahl, H. Kittler “The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions, (eng)” Çataloluk, H., & Kesler, M. (2012).

[2] A diagnostic software tool for skin diseases with basic and weighted K-NN. In 2012 international symposium on innovations in intelligent systems and applications (pp. 1-4): IEEE.

[3] Dreiseitl, S., Binder, M., Hable, K., & Kittler, H. (2009). Computer versus human diagnosis of melanoma: evaluation of the feasibility of an automated diagnostic system in a prospective clinical trial. Melanoma Research, 19, 180-184.

[4] J. Kawahara, S. Daneshvar, G. Argenziano, G. Hamarneh” 7-Point Checklist and Skin Lesion Classification using Multi-Task Multi-Modal Neural Nets, (eng)” IEEE J. Biomed. Health inf. (2018)

[5] Haenssle, H. A., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., Kalloo, A., Hassen, A. B. H., Thomas, L., & Enk, A. J. A. o. O. (2018).

[6] Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. 29, 1836-1842.
[7] S. Dreiseitl, L. Ohno-Machado, H. Kittler, S. Vinterbo, H. Billhardt, and M. Binder, “A comparison of machine learning methods for the diagnosis of pigmented skin lesions,” Journal of biomedical informatics, vol. 34, no. 1, pp. 28–36, 2001.
[8] C. Barata, M. Ruela, M. Francisco, T. Mendonc¸a, and J. S. Marques, “Two systems for the detection of melanomas in dermoscopy images using texture and color features,” IEEE Systems Journal, vol. 8, no. 3, pp. 965–979, 2014.
[9] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
[10] G. Argenziano and H. P. Soyer, “Dermoscopy of pigmented skin lesions–a valuable tool for early diagnosis of melanoma,” The Lancet Oncology, vol. 2, no. 7, pp. 443–449, 2001.