JOURNAL OF TECHNIQUES

Journal homepage: http://journal.mtu.edu.iq

RESEARCH ARTICLE - ENGINEERING

Comparison Between Convolutional Neural Network CNN and SVM in Skin Cancer Images Recognition

Zaid G. Hadi[1*], Ahmed R. Ajel[1], AyadQ. Hussain[1]

[1] Electrical Engineering Technical College, Middle Technical University, Baghdad, Iraq.

[*] Corresponding author E-mail: zaidghazi2020@gmail.com

Article Info.

Article history:

Received
06 September 2021

Accepted
13 December 2021

Publishing
31 December 2021

Abstract

As compared with benign, the common human malignancy of skin cancer can be diagnosed visually starting from clinical screening and ending with histopathological examination. Accurate automatic classification of skin lesion images is a great challenge as the image features are very close in these images. In this paper we used two methods for image recognition. The first method was carried out with Convolution neural networks (CNN) that promise to provide a potential classifier for skin lesions. This work presents a dermatologist-level classification of skin cancer by using residual network (ResNet-50) as a deep learning convolutional neural network (DLCNN) that maps images to class labels. It presents a classifier with a single CNN to automatically recognize benign and malignant skin images. As for the second method, we used the Support Vector Machine. Which is a supervised learning algorithm and it is used for classification of data for the different classes based on a separating hyperplane. The network inputs are only disease labels and image pixels. About 320 clinical images of the different diseases have been used to train the CNN. The model performance has been tested with untrained images from the two labels. This model identifies the most common skin cancers and can be updated with a new unlimited number of images. The DLCNN was trained by the ResNet-50 model and it showed good classification of the benign and malignant skin categories. The ResNet-50 as a DLCNN has achieved a significant recognition rate of more than 97% on the testing images, which proves that the benign and malignant lesion skin images are properly classified. Support vector machine (SVM) classifier for the classification of skin cancer, for the feature extraction step achieved 86.9% accuracy. This means in CNN; we had more accuracy with 11%. These results were attained using MATLAB.

Keywords: CNN; Skin Cancer; SVM; Lesion Classification; Melanoma Classification; DLNN; residual network (ResNet-50)

1. Introduction

1.1. Background

It is reported that there is 5.4 million new cases of skin cancer in the US alone and despite melanoma patients are less than 5%, it accounts for approximately 75% of all types of skin cancer in the US that related cause over 10,000 deaths per year [1,2]. Melanoma is approximately 10 times more likely to spread and disseminate than all the other types of skin cancer the other thing is it's also very capricious and unpredictable most of the other types of skin cancers. It usually expects to see them in areas that are sun-exposed dominant or occupationally some exposed dominant melanoma on the other hand very unpredictable. It's seen around people's feet even in people's eyeballs are very unpredictable the other problem with it is it's like an octopus it tends to borrow and it has tentacles that tend to spread out quickly. Early detection of these cases is critical, where the 5-years survival rate drops from over 99% to 14% if detected in its latest stages [3].

According to [4-6], it is possible to classify human skin disease taxonomy, where red refers to malignant, orange denotes Non-neoplastic, black for melanoma, and green refers to benign, as shown in Fig. 1.

This classification figure has been prearranged visually and clinically with the help of medical specialists. Deep learning convolutional neural networks (DLCNN) skin image classifier enables to embed a dermatologist level classifier into our phone or PC by just snap a picture of user skin lesion and instantly getting such diagnosis [3].

Nomenclature			
CNN	Convolutional Neural Network	RGB	Red Green Blue Images
DLCNN	Deep learning convolutional neural networks	CPU	Central Processing Unit
SVM	Support Vector Machine	**Symbols**	
ResNet-50	pre-trained architecture in MATLAB	X	Input Vector
DLNN	Deep Learning Neural Network	W	Modifiable Weight Vector
GPU	Graphical Processing Unit	b	Bias

A popular way of classifying images is that most images take as input and by using a series of compositional filters and other miscellaneous things the deep learning communities come up with for better vision applications, and then match the images into class labels such as scattered off. The use of deep learning with an automated dermatology classifier will provide a new disease-partitioning algorithm for skin disease taxonomy that mapping each disease according to training classes. The idea in this work is to detect skin cancer in clinical images by using a skin cancer classifier based on a convolution residual neural network. The considered data set in the paper consists of 320 clinical images including 160 benign and malignant classes. The skin lesion images have been resized to match the input of the DLCNN classification network. Residual Network, which is called ResNet-50, as a DLCNN, is considered in this paper. ResNet-50 is a pre-trained architecture that is already trained on a database subset of the Image-Net for large-scale visual image recognition. A million images have been trained on ResNet-50, which has 177 layers corresponding to 50-layer residual networks. It can classify 1000 images into object classes.

A comparison was made between the CNN and SVM methods for the skin cancer disease images recognition and to find which one of these methods is best performance and more speed and more accuracy.

Several deep learning techniques have been introduced to improve structure performance in solving different image processing and engineering problems. The learning techniques of DLCNN have undergone a large growth, which sharply improves its performance in various applications such as medical image classifiers [7,8].

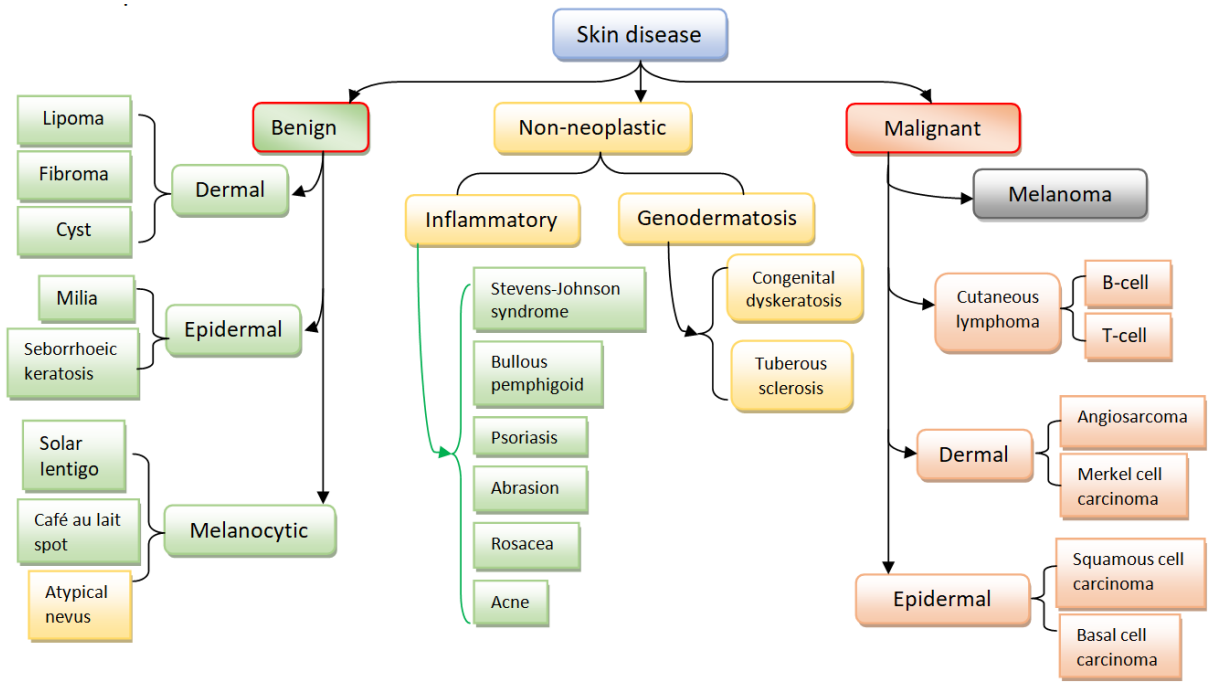

Fig. 1 A diagram illustration of human skin disease taxonomy, where it is organized according to the clinical and visual similarity of the skin diseases [4-6]

1.2. Literature review

Several studies related to dermatological classifier [9–11] but they are have lacked to generalize the capability of disease-partitioning with unsatisfactory information and a spotlight on standard tasks like histological [12] and dermoscopy image classifier [13]. Histological images are obtained by invasive microscopy and biopsy, while dermoscopy images are obtained by a specific device, and both modalities create consistent images. Smartphone images as photographic images show inconsistency in features such as lighting, zoom, and angle, which makes a classifier significantly further challenging [14,15]. Other previous related studies require extraction of visual features, lesion segmentation, and extensive processing, and pre-training before classification. In this work, a data-driven approach [16,17] is considered to overcome these challenges. Hundreds of images with photographic variability are trained to ensure robust classification. In addition, no hand-crafted features are required in our system. The presented model trains the image raw pixels and labels by a single structure for dermoscopy and photographic images. This study presents a generalized classification approach with a flexible number of input clinical datasets that reach hundreds of thousands to generalize the application over extensive dermoscopy and photography images.

Advances in DLCNN with large-scale datasets [18] that recently create different visual tasks systems of human performances, for example; playing Atari game [19,20], object recognition [21,22], and strategic games such as Go [23]. Hagerty et al. [24] used a utilized residual network algorithm called ResNet-50 as a DLCNN for image processing to fuse the features from individual methods by hypothesizing different error

profiles and using three hand-crafted biological modules for predicting the probability of melanoma categorization. However, the adopted handcraft features-based method is not clear. In this work, an image-based classification using the CNN system is developed to match dermatologists satisfactory in some significant key diagnostic aspects such as melanoma/ nonmelanoma classification by means of carcinoma and dermoscopy classifications. An open-source dataset has been pre-trained over approximately 1.3 million clinical images with large-scale object categories [25], and the dataset is trained using Renet50 learning architecture [26].

1.3. Manuscript Outlines

The manuscript is arranged in five sections. An introduction includes systematic literature review by unfolding research questions, literature review scope, information source In section one, in section two introduces the research method, which includes the image preprocessing to analyse of skin cancer including feature learning and the classification by using two types of methods CNN with ResNet50 and SVM, in section three the results have been discussed. Moreover, deep learning-based classifier, the performance and dataset and compare between them were discussed in this section, section four provides the conclusions of this study and Acknowledgement in section five.

2. Methodology

2.1. Diagnosis

The Convolutional Neural Networks (CNN) is trained with skin disease classes. Our dataset is organized with dermatologist-labeled clinical images that are coming from an online open-access clinical source. We divide the dataset into 70% for training images and validations, while 30% for testing images. An algorithm, that portioning skin diseases such as melanoma and non-melanoma, is developed as a fine training class. Initially, skin lesion images are separated from adjoining artifacts and normal skin tissue like hairs, air bubbles, and veils, with mean thresholding, color space transformation, and removal of the region of interest. Then, the region of interest image employing the binary mask image a set of features based on texture, shape, and color, is obtained. The features of DLNN are acquired appropriate CNN structure that is trained in a Matlab-function called ResNet-50 classification algorithm that includes CNN architecture to extract image features [27]. The obtained classification is trained on a dataset consist of both benign and malignant skin-lesion images.

2.2. ResNet-50 architecture

The deep learning (DL) part of our approach depends on a transfer learning described in [27], and a deep residual network called (ResNet), which is a convolution-based architecture [28]. Its input matrix can accept 224x224 RGB images. DL ResNet architecture is effective in a large-scale image visual recognition to classify 1000 objects [29–31]. ResNet can achieve a correct recognition error of 3.6%. ResNet is a convolution-based architecture. Bicubic interpolation is usually used to reduce the resolution. A sequential series of residual blocks (see Fig. 2) is the main property of ResNet, which results in 2048 feature vectors.

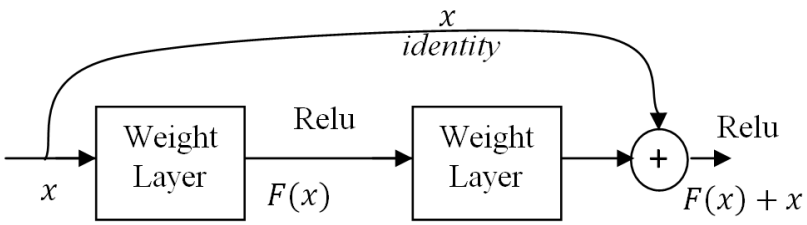

Fig. 2 Residual block of ResNet-50 [28]

Since this work considers the Residual network as a DLCNN to solve accurately the clinical skin lesion image recognition problem, the underlying architecture is illustrated layer by layer, starting from an RGB image to an output production, including convolutional layers, batch normalization layer, and max-pooling layers. There are too many convolution layers attached one by one, and there are many skip connections called residual neural networks. A main conceptual block diagram of the proposed approach is demonstrated in Fig. 3, 4, and 5. To make the structure simpler to understand, the complete architecture is broken down into 19 small blocks. The first block is shown in Fig. 3.

It is known that transfer learned models have a predefined shape, which is in the start described by 1x 3 x 224 x 224. The first one shows the batch size, then the number of RGB channels, then 224 x 224 are width and height of an input image. ResNet-50 is applying 64 convolutions of 7x 7 kernel size and a padding of 3 and strides of 2 that represent the tensor operation. The size of the image after padding is the same as width but when applying the kernel impact the size will be reduced to 230 after padding. The output from this first convolutional layer is 64 and the output size will be 1x 112x 112x 64. The batch normalization applies to a simple layer and there are 64 mean variance beta and lambda which is the scaling factor that doesn't make any changes with the tensor other than normalizing.

In the second block, the max-pooling is applied, which has strides 2, padding 1, and the kernel size is 3x3, as shown in Fig. 4.

Therefore, the size is reduced and the complete shape is divided by 2 and the output shape will be 1x56x56x64. In this block, we have multiple convolutions on one side, then pass normalization, and the input dimension will remain the same. Another one convolution, which has the same kernel 1x1padding, and strides are also 1x1 that makes no changes with the shape. There is a parallel connection with the simple shape of 256x64x1x1 that provides output 1x56x56x256. Therefore, only the 64 got converted to 256.

Going forward in the CNN, these blocks are similar in nature actually, they are just tacked up one by one. The last end of this chain is having 256 filters in one convolution. So it will again change the shape to 256 and this skip connection, which has been taken from this value (denoted by Relu), which is directly connecting to the sum. Getting understood of block 2 and bock 3 all rest of the blocks are similar in nature. The last block of the ResNet-50 architecture is shown in Fig. 4.

The last layer or block 19 includes applying an average pooling of 7 x 7. The average pooling is a little bit different though converting the 7 x 7 to 1 x 1 within one channel of 7 x 7. The output will be 1 x 2048. Then a general matrix multiplication that is inclusive of the SoftMax here. The output of the SoftMax is 1000 classes and the input to the SoftMax is 2048. Therefore, using this complete ResNet-50 CNN, this last SoftMax layer can be replaced by a defined number of classes and keeping all previous layers not trainable.

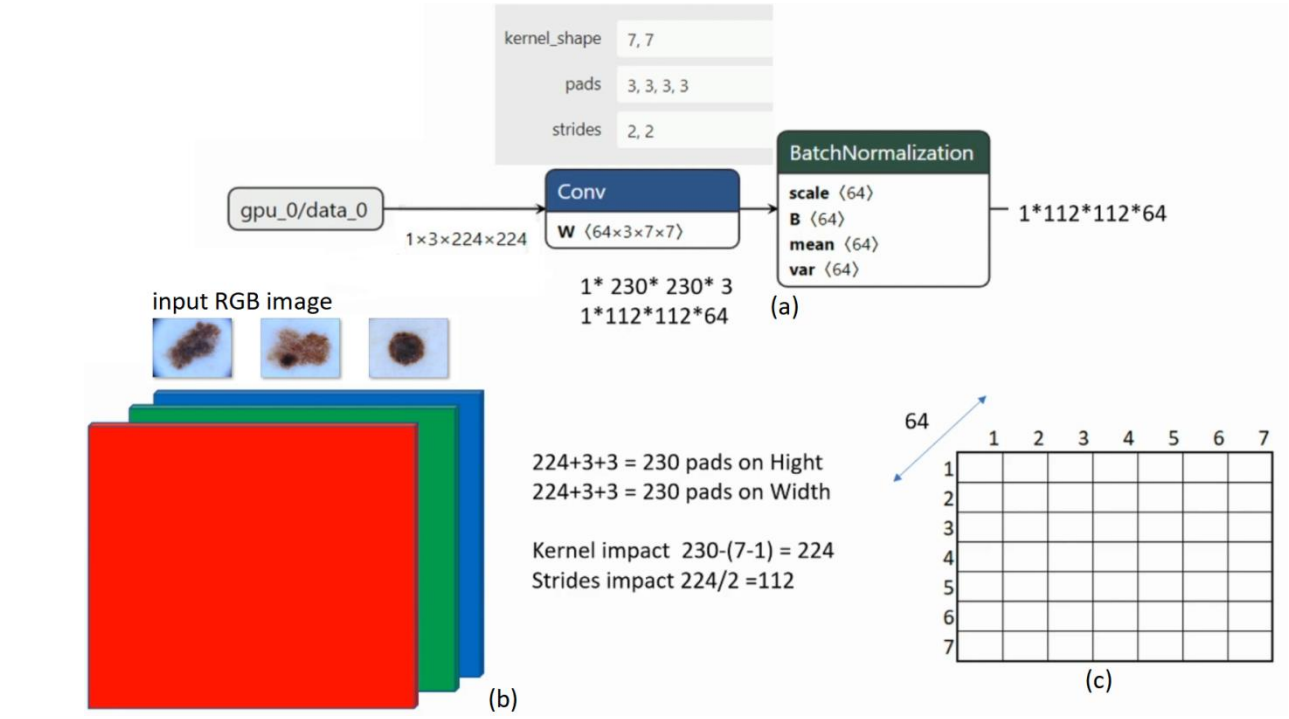

Fig. 3 Block 1 of 19 ResNet-50 architecture; where (a) represents the 1st convolution operation, (b) is the RGB input image, and (c) is 7x7

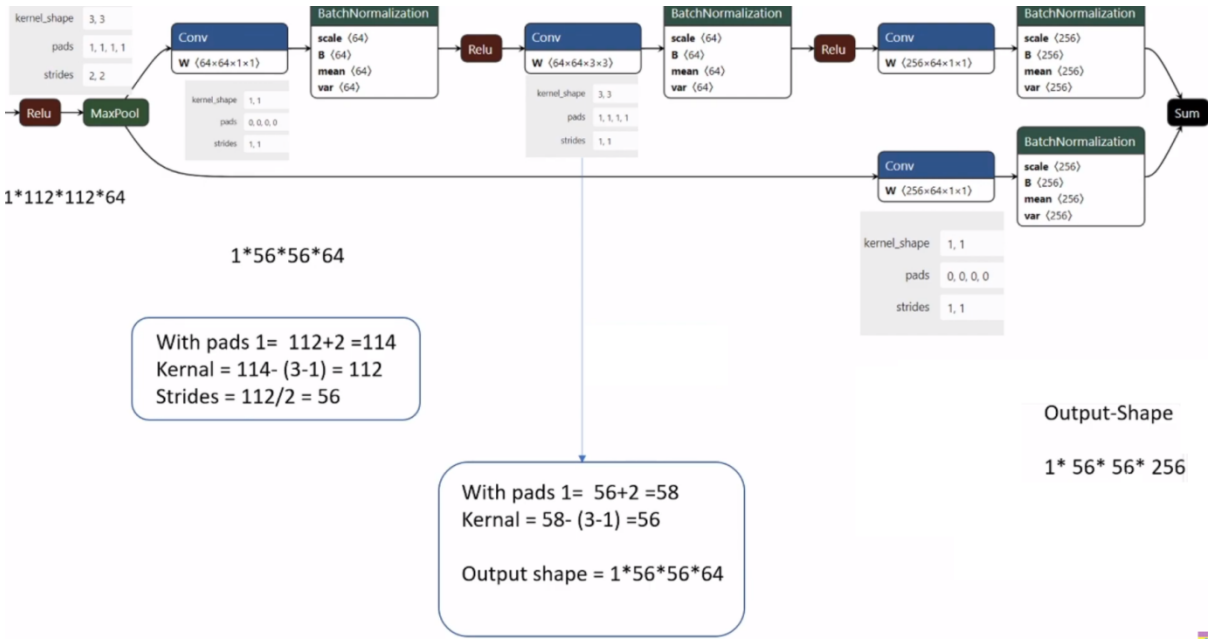

Fig. 4 Block 2 of 19 of the ResNet-50 architectures

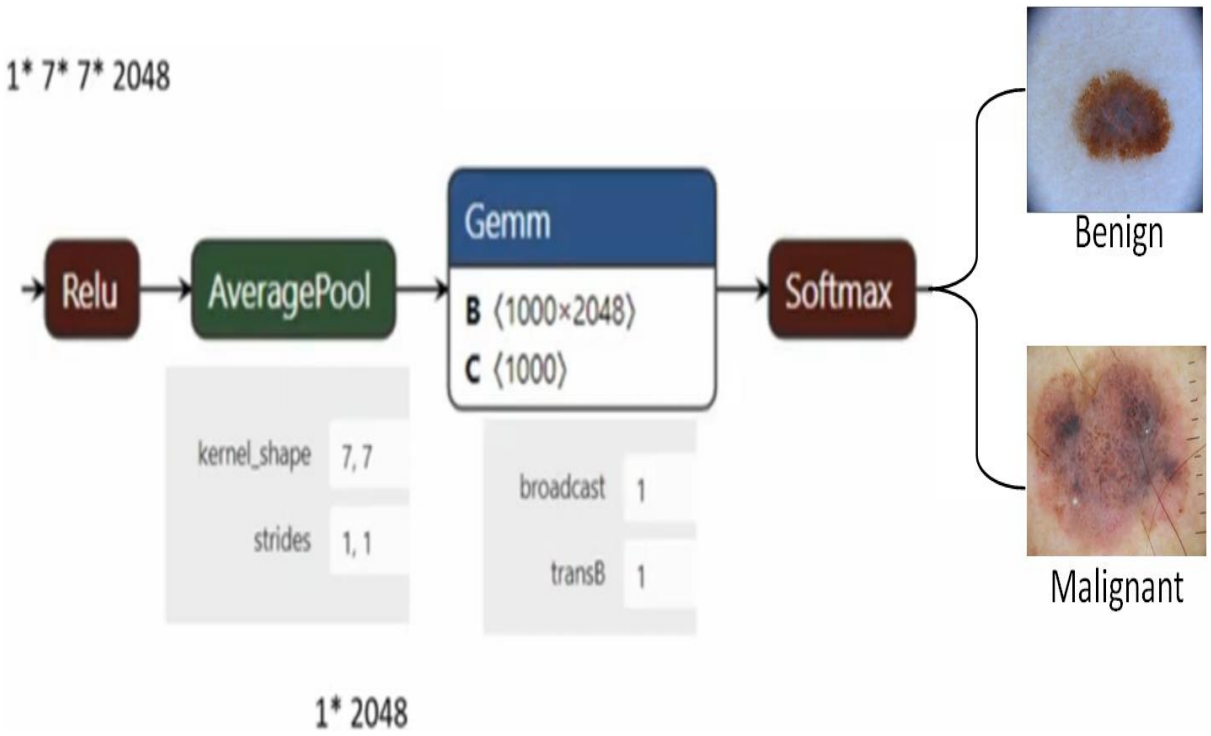

Fig. 5 The 19th (final block) of 19 of the ResNet-50 architecture

2.3. Support vector machine (SVM)

Support Vector Machine is a supervised learning method that uses a hyperplane that separates the data to classify them into distinct classes. The SVM has two types of classifications; it is non-linear and the linear classification. The training dataset is mapped into higher dimensional space and infinite in support vector machine. Mapped the vectors of the inputs to the feature space by using the functions of Kernels, that calculates the dot product in the feature space immediately. It tries to keep the gap between him and the training samples as small as possible as shown in fig. 6.

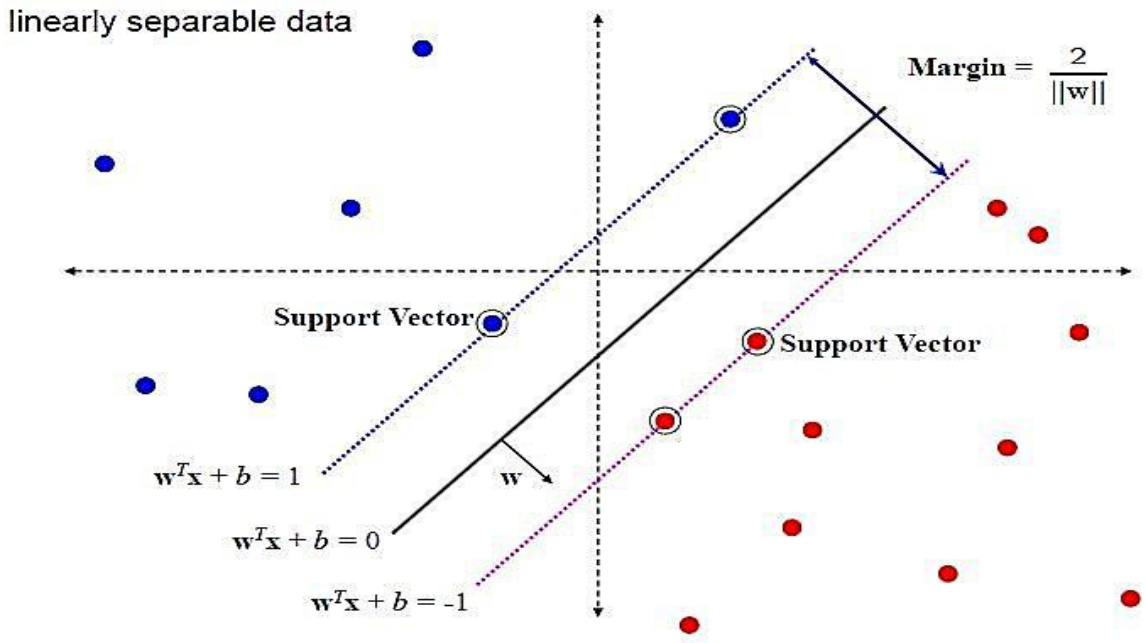

Fig. 6 Support Vector Machine [32]

The separation that is does in the hyperplane works by the equation:
WT X + b = 0
X: input vector.
W: modifiable weight vector.
b: bias.
By using the equation of the kernel which converts the incoming input to the space of higher dimensions and the second phase in the new space searches for a linear hyperplane to separate the samples [30], as shown in fig. 7.

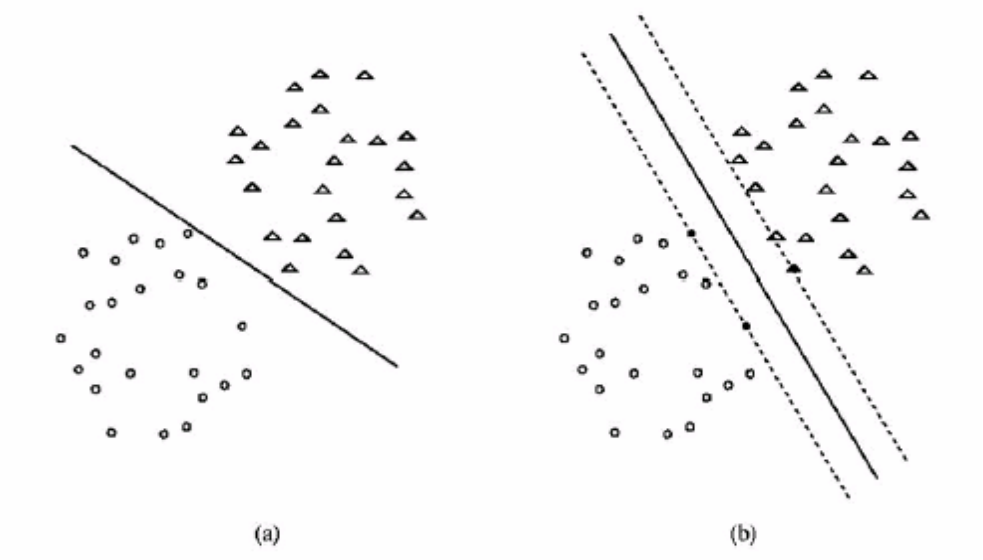

Fig. 7 Linear and Nonlinear Support Vector Machine [33]

3. Result and Comparison

3.1. Result

The feasibility of the developed DLCNN to categorize the skin lesion images is verified by training 70% of the dataset and testing the remaining images that are not identified by the network. All patches of the residual CNN (ResNet-50) were (n = 1687), trained, and augmented in the training cohort in ResNet-50 to enlarge the robustness of the architecture. The upper graph of Fig. 8 demonstrates the accuracy during the learning, which was found to be 0.9779, while the cross-entropy loss was close to 0.132 after 8th epochs training (248 iterations), which can be seen at the bottom of the figure. The elapsed time here was 1 min 0 seconds as the hardware resource is a single CPU and not GPU. The details are also shown in Fig. 8.

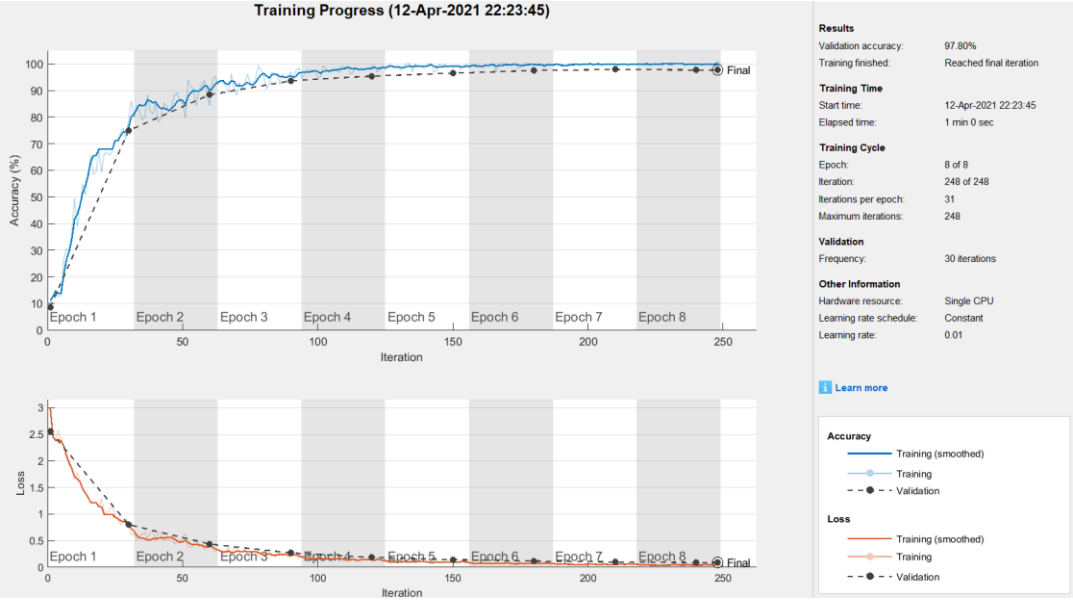

Fig. 8 The process of training and validation of the DLCNN using ResNet-50 model when training skin lesion images

The accuracy is in the upper while the cross-entropy loss is in the bottom. Two categories have been classified over 248 maximum iterations. The input image is responded to or activated by each layer of the DLCNN. Some of these layers are performing the extraction of image features. The basic image features such as edges and blobs are captured by DLCNN at the beginning layers of the network. To take visual insight for the network filter weights, Fig. 9. shows weights samples at the 1st convolution layer.

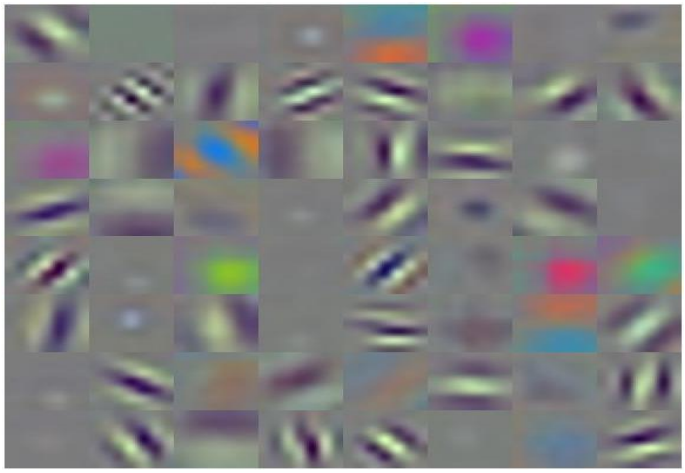

Fig. 9 First convolutional layer weights

The DLCNN trained by the ResNet-50 model showed good classification of the benign and malignant skin categories. The ResNet-50 in DLCNN verified a significant recognition rate of more than 97% on the testing images, which proves that the benign and malignant lesion skin images are properly classified. Now, the predictive accuracy of DLCNN architecture can be obtained in each patch by a confusion matrix, which is displayed in Fig. 10.

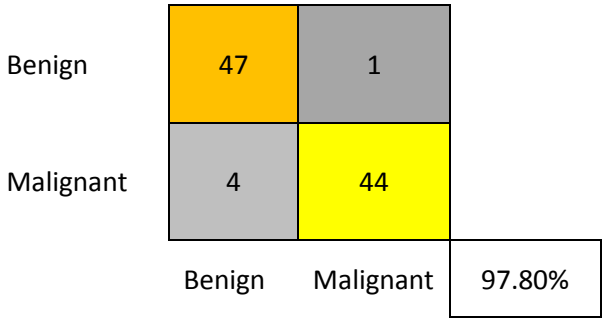

Fig. 10 predictive accuracy of DLCNN architecture

3.2. Comparison

Simulation was performed on the system using both the CNN and SVM methods simulated by [34]. The system was tested with 320 images in CNN so it has 97% accuracy in convolutional neural network and timing 1 minute and 0 second more by 11% then SVM that has accuracy 86% [34] so the result of our comparison that CNN is a best way for skin cancer image recognition then SVM.

4. Conclusion

This paper presents a developed residual network as a deep learning convolutional neural network and support vector machine for image recognition purposes to classify the lesion skin images into benign and malignant images. About 320 clinical images of the different diseases have been used to train the CNN and SVM.

The aim is to make a comparison between SVM which performed by [34] results and CNN. To find the better performance in image recognition in accuracy, so we found that using SVM recognition having accuracy of 86% and in CNN method by using Deep ResNet-50 network with skip connections was employed as an approach to perform the partitioning images task, which results in higher classification performance then SVM in 11% to have an accuracy in 97% in CNN.

The results of training and the verification by the testing images show that ResNet-50 DLCNN is efficiently able to accurately generalize the grade of benign and malignant images with a very small margin of error. The significance and robustness of this model in grading skin lesion images are because of its control of learning high and low levels to extract image features utilizing its deep residual networks, pooling layers, skip connections, layer depth, and convolutions. These features help to extract unimaginable features that contribute to reaching high recognition rates throughout the training of the network. Such recognition architecture must be assigned as the mainly efficient, accurate, errorless, reliable, and flexible models for image partitioning to classify clinical lesion skin image applications.

Acknowledgement

I extend my heartfelt thanks to my wonderful college, the Electrical Engineering Technical College and its professors, and to the Middle Technical University, who provided me with all the possibilities to complete my research paper.

References

[1] K. G. Paulson et al., "Age-Specific Incidence of Melanoma in the United States," JAMA Dermatology, 2020.
[2] ACS, "Key Statistics for Basal and Squamous Cell Skin Cancers," Am. Cancer Soc., 2019.
[3] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, 2017.
[4] M. N. Bajwa et al., "Computer-aided diagnosis of skin diseases using deep neural networks," Appl. Sci., 2020.
[5] R. Zell et al., "ICTV virus taxonomy profile: Picornaviridae," J. Gen. Virol., 2017.
[6] S. W. Kashem, M. Haniffa, and D. H. Kaplan, "Antigen-presenting cells in the skin," Annual Review of Immunology. 2017.
[7] R. H. Abiyev and A. Helwan, "Fuzzy Neural Networks for Identification of Breast Cancer Using Images' Shape and Texture Features," J. Med. Imaging Heal. Informatics, 2018.
[8] Z. Xian, X. Wang, S. Yan, D. Yang, J. Chen, and C. Peng, "Main Coronary Vessel Segmentation Using Deep Learning in Smart Medical," Math. Probl. Eng., 2020.
[9] B. N. Walker et al., "Dermoscopy diagnosis of cancerous lesions utilizing dual deep learning algorithms via visual and audio (sonification) outputs: Laboratory and prospective observational studies," EBioMedicine, 2019.
[10] A. Hekler et al., "Effects of Label Noise on Deep Learning-Based Skin Cancer Classification," Front. Med., 2020.
[11] W. B.N. et al., "Dermoscopy diagnosis of cancerous lesions utilizing dual deep learning algorithms via visual and audio (sonification) outputs: Laboratory and prospective observational studies," EBioMedicine, 2019.
[12] M. C. M. van Zon, J. D. van der Waa, M. Veta, and G. A. M. Krekels, "Whole-slide margin control through deep learning in Mohs micrographic surgery for basal cell carcinoma," Exp. Dermatol., 2021.
[13] A. Rezvantalab, H. Safigholi, and S. Karimijeshni, "Dermatologist level dermoscopy skin cancer classification using different deep learning convolutional neural networks algorithms," arXiv. 2018.
[14] O. T. Jones, C. K. I. Ranmuthu, P. N. Hall, G. Funston, and F. M. Walter, "Recognising Skin Cancer in Primary Care," Adv. Ther., 2020.
[15] S. S. Chaturvedi, J. V. Tembhurne, and T. Diwan, "A multi-class skin Cancer classification using deep convolutional neural networks," Multimed. Tools Appl., 2020.
[16] H. H. Wang, Y. H. Wang, C. W. Liang, and Y. C. Li, "Assessment of Deep Learning Using Nonimaging Information and Sequential Medical Records to Develop a Prediction Model for Nonmelanoma Skin Cancer," JAMA Dermatology, 2019.
[17] D. E. Webster et al., "The Mole Mapper Study, mobile phone skin imaging and melanoma risk data collected using ResearchKit," Sci. Data, 2017.
[18] L.-J. Li et al., "ImageNet: a Large-Scale Hierarchical Image Database Shrimp Project View project hybrid intrusion detction systems View project ImageNet: A Large-Scale Hierarchical Image Database," 2009 IEEE Conf. Comput. Vis. Pattern Recognit., 2009.
[19] M. Jaderberg, "Deep Learning for Text Spotting University of Oxford Supervisors : Professor Andrew Zisserman Doctor Andrea Vedaldi," Thesis, D Phil, 2014.
[20] N. A. Tu, T. Huynh-The, K. S. Wong, D. M. Bui, and Y. K. Lee, "Distributed Feature Extraction on Apache Spark for Human Action Recognition," in Proceedings of the 2020 14th International Conference on Ubiquitous Information Management and Communication, IMCOM 2020, 2020.
[21] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," Int. J. Comput. Vis., 2015, doi: 10.1007/s11263-015-0816-y.
[22] K. H. Mahmud, Adiwijaya, and S. Al Faraby, "Klasifikasi Citra Multi-Kelas Menggunakan Convolutional Neural Network," e-Proceeding Eng., 2019.
[23] D. Silver et al., "Mastering the game of Go with deep neural networks and tree search," Nature, 2016.
[24] J. R. Hagerty et al., "Deep Learning and Handcrafted Method Fusion: Higher Diagnostic Accuracy for Melanoma Dermoscopy Images," IEEE J. Biomed. Heal. Informatics, 2019, doi: 10.1109/JBHI.2019.2891049.
[25] "Robust Melanoma Screening - Datasets." https://sites.google.com/site/robustmelanomascreening/dataset (accessed Apr. 10, 2021).
[26] "ResNet-50 convolutional neural network - MATLAB resnet50." https://www.mathworks.com/help/deeplearning/ref/resnet50.html;jsessionid=bdfe63b8d35bfa4b58f44aca0512 (accessed Apr. 10, 2021).
[27] K. Weiss, T. M. Khoshgoftaar, and D. D. Wang, "08," J. Big Data, 2016.
[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016.
[29] A. Vilardi, "Very Deep Convolutional Neural Networks for face identification," IEEE/ACM Trans. Audio, Speech, Lang. Process., 2016.
[30] M. T. Hagos and S. Kant, "Transfer learning based detection of Diabetic Retinopathy from small dataset," arXiv. 2019.
[31] H. Zhu, H. Wei, B. Li, X. Yuan, and N. Kehtarnavaz, "A review of video object detection: Datasets, metrics and methods," Applied Sciences (Switzerland). 2020.
[32] A. Zisserman, Lecture 2: "The SVM Classifier. C19 Machine Learning "(Hilary Term 2015). Available Online At: Http: //Www. Robots. Ox. Ac. Uk/~ Az/Lectures/Ml/Lect2. Pdf (Accessed March 23, 2019), 2015
[33] C.W. Hsu, And C.J. Lin, "A Comparison of Methods for Multiclass Support Vector Machines. IEEE Transactions on Neural Networks", 13(2), pp.415-425, 2002.
[34] X. Yuan, Z. Yang, G. Zouridakis, and N. Mullani, , "SVM-Based Texture Classification And Application To Early Melanoma Detection", In International Conference Of The IEEE Engineering In Medicine and Biology Society, pp. 4775-4778, IEEE, 2006.