

DTLNet: Deep Transfer Learning-based Hybrid Model for Skin Lesion Detection and Classification

Doma Murali Krishna¹, Sanjay Kumar Sahu², G R L V N Srinivasa Raju³

Abstract - In humans, MELs are a kind of skin cancer that develops in the pigment-producing cells (melanocytes) that are responsible for producing pigment in the skin. Skin cancer, such as melanoma (MEL) must be detected and diagnosed as early as possible in order to limit the number of people who die from the disease. Deep transfer learning (DTL)-based skin lesion detection and classification (SLDC) system, referred to as DTLNet, is the subject of this article, which focuses on its implementation in a computer network. Initially, the skin lesions are preprocessed by using the hybrid gaussian-wiener filter (HGWF), which removes the noises from skin images. Secondly, skin lesion segmentation operation is performed by using transfer learning based AlexNet model. Then, the hybrid features were extracted by using deep learning convolutional neural network (DLCNN) model. Finally, multi class classification operation is performed by using SoftMax Classifier of DLCNN, which classifies the eight different types of skin cancers. Further simulations performed on ISIC-2019 dataset showed that the proposed framework resulted in superior performance as compared to the state of art approaches.

Keywords: Skin cancer detection and classification, deep learning, transfer learning, convolutional neural network, ISIC-2019 dataset.

1. INTRODUCTION

MEL and non-MEL skin cancers are frequently the two most well-known forms of skin cancer [1]. But the patients are suffering with the many forms of skin cancers, they include MEL, melanocytic nevus (NV), basal cell carcinoma (BCC), actinic keratosis (AKIES), benign keratosis (BKL), dermatofibroma (DF), vascular lesion (VASC), and squamous cell carcinoma (SCC) (SCC). Figure 1 illustrates the all these eight forms of skin cancers and they are considered from ISIC-2019 dataset.

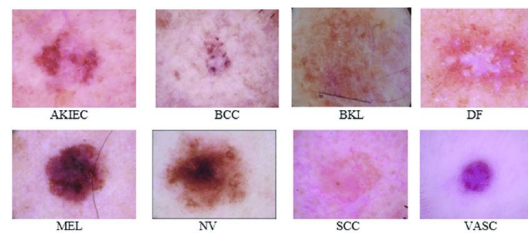


Figure 1. Sample images of ISIC-2019 dataset.

MEL is a treatable condition if detected and treated at an early stage. Biopsy, pathology report, and medical imaging analysis are all methods of detecting MEL in its early stages. Dermoscopy is a non-invasive imaging tool that is often used to identify MEL early in order to increase the likelihood of a positive outcome. When dermatologists examine dermoscopy photos, it is costly and needs a high degree of competence in order to accurately diagnose the condition. In light of this difficulty, it has been determined that effective computer-aided diagnostic procedures are required to help in the early detection of MEL from skin lesions [1]. The presence of MEL may be associated with

¹Research Scholar, School of Electronics and Electrical Engineering Lovely Professional University, Phagwara, Punjab, India. mkrishna557@gmail.com

²School of Electronics and Electrical Engineering Lovely Professional University, Phagwara, Punjab, India. sanjay.23393@lpu.co.in

³Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh, India. grlvnsr@svecw.edu.in

a higher visual similarity between cancerous and non-cancerous cells, making it difficult to distinguish between MEL and non-MEL skin cancer. Numerous computer-added SLDC methods were proposed for skin cancer identification. The skin cancer detection methods were developed by using basic image processing approaches, machine learning algorithms, deep learning models [2], transfer learning [3] and ensemble learning prototypes. Among those various models, deep learning resulted in high performance but suffering with computation complexity problems and transfer learning models are resulted in better performance with low complexity. Hence, this article is focused on utilization of both deep learning transfer learning models and developed a hybrid network called as DTLNet for SLDC.

The organization of article is, section 2 deals with literature survey with their drawbacks. Section 3 deals with the detailed analysis of proposed DTLNet. Section 4 deals with summary of the article with possible future scope.

2. LITERATURE SURVEY

This section gives the detailed analysis of diverse related works on SLDC system using segmentation model, ensemble models, and transfer learning models.

Deep learning approaches are resulting in superior performance in various applications including SLDC models. In [4], the author reviewed automated skin cancer diagnosis as well as the use of image processing and machine learning in skin cancer detection and prevention. Another article provided an overview of the most recent research efforts in skin lesion detection and classification using CNN, transfer learning, and ensemble techniques, as well as their results. Deep learning architecture was employed by the authors in [5], they concentrated mostly on lesion attribute recognition, lesion border segmentation, and lesion diagnosis, and they achieved the highest accuracy of 92.74 percent on the ResNet neural network. A deep learning architecture for segmentation was created in [6] by the authors, who named it the pyramid scene parsing network (PSPNet). But the system segmentation accuracy is measured for lesser epochs. In U-Net [7] they implemented skin lesion segmentation with tversky index for loss optimization. Further, the complexity of this work is increased as number of layers are increased. Finally, multi-layer residual convolutional neural network (MLRNet) [8] is developed for skin lesion segmentation, which is also utilized modified gaussian and guided image filters for noise removal. Compared to all other approaches, MLRNet resulted in superior performance. With transfer learning-based approaches, great accuracy was attained while reducing the demand for huge datasets for a variety of classification tasks. The authors employed pixel-based fusion and multilayer feature reduction to run two tests on the ISBI-2016 and ISIC-2017 datasets for segmentation and classification, and they were able to reach an accuracy of 95 percent for MEL classification. Recent research has concentrated on creating an ensemble of multiple models in order to obtain high accuracy while utilizing dermoscopic lesions. For SLDC, an ensemble of Deep Neural Networks models, such as AlexNet, VGGNet, and GoogleNet, were used. In order to classify skin lesions, deep learning-based techniques such as artificial neural networks (ANN), backpropagated-ANN, DenseNet 201[9], CNN with data augmentation (CNN-DG) [10], DLCNN [11], and Hybrid CNN (HCNN) [12] have been used. The complexity of this work is increased due to synchronization failure between various models.

3. PROPOSED METHOD

The evaluated proposed DTLNet were applied on ISIC-2019 dataset for multi class classification. This dataset was collected under different conditions and had different characteristics.

Table 1. Proposed DTLNet algorithm.

Input: ISIC-2019 training dataset, test skin lesion
Output: Predicated Classes, Quantitative evaluation.
Training process
Step 1: Perform the HGWF pre-processing operation on ISIC-2019 dataset, which eliminates the different types of artifacts from the dataset.
Step 2: Apply the AlexNet based transfer learning model for skin lesion segmentation on pre-processed outcomes.
Step 3: Apply the DLCNN architecture for extracting the multiple disease dependent features with high correlation and create the feature database.
Testing process
Step 4: Consider the test skin lesion and perform the steps 1 to 3, which extracts the test skin lesion features.
Step 5: Perform the SLDC by comparing the test features with trained features using SoftMax classifier, which classifies the multiple classes of skin lesion.
Step 6: Perform the quantitative evaluation and calculate the various performance metrics.

In Figure 2, the mechanism the proposed DTLNet for diagnosing skin diseases is presented and Table 1 presents the proposed DTLNet algorithm. Images were enhanced using HGWF. The lesion segmentation was performed using the AlexNet algorithm. Feature extraction was conducted using DLCNN, where the deep feature maps were contained inter disease dependent and disease specific features. These features were classified using SoftMax classifier for multi class classification including SCC, VASC, DF, BKL, AKIES, BCC, NV and MEL.

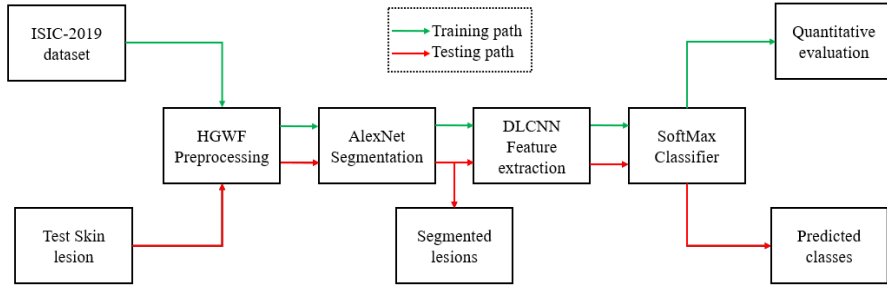


Figure 2. Proposed DTLNet framework.

A. HGWF-Preprocessing.

Skin lesions are suffering with the different types of noises such as salt-pepper, gaussian, random, jitter, poison, etc. Further, the skin lesions are also suffering with the hair artifacts, which also degrades the performance of segmentation and classification.

Thus, this article newly introduces the HGWF based preprocessing for skin lesion enhancement.

B. Segmentation using AlexNet

Deep learning models are used to segment the skin lesions. But the standard deep learning models are suffering with the vanishing gradient problems for the huge datasets like ISIC-2019. Even though, the deep learning models are resulted in better performance by analyzing each pixel, but they are suffering with the high computational complexity. Thus, this work is focused on adaption of transfer learning model for skin lesion segmentation.

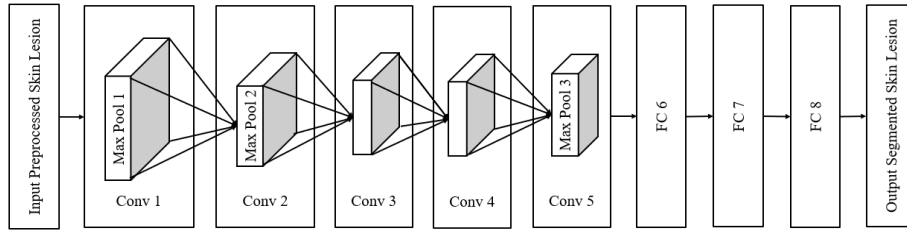


Figure 3. AlexNet architecture for skin lesion segmentation.

Table 2. Layer details of AlexNet model

Layer	Number	Filter	Stride	Padding	Feature	Activation
Conv 1	96	11 x	4	-	55 x 55x 96	ReLU
Max	-	3x3	2	-	27x27x96	-
Conv 2	256	5x5	1	2	27x27x256	ReLU
Max	-	3x3	2	-	13 x 13x	-
Conv 3	384	3x3	1	1	13 x 13x	ReLU
Conv 4	384	3x3	1	1	13 x 13x	ReLU
Conv 5	256	3x3	1	1	13 x 13x	ReLU
Max	-	3x3	2	-	6 x 6 x 256	-
FC 6	-	-	-	--	1 x 4096	ReLU
FC 7	-	-	-	--	1 x 4096	ReLU
FC 8	-	-	-	--	1 x 1000	SoftMax

Transfer learning models are resulting the superior performance in image analysis applications including image recognition, image segmentation, background extraction and edge analysis. AlexNet is one such transfer learning model, which contains the low computational complexity than other transfer learning models like ResNet, GoogleNet, and MobileNet. Figure 3 shows the segmentation process of skin lesion using AlexNet architectures, which is used to highlight the cancer region by classifying the pixels. Table 2 shows the detailed information of each layer presented in AlexNet model. Further, the AlexNet performs the skin lesion segmentation operation by analyzing the Asymmetry, Border, Colour, Diameter, Edge (ABCDE) properties.

The preprocessed skin lesion input image is applied as an input to the convolution layer 1 of the image processing algorithm (Conv1). The kernel or filter-based feature

detectors are located in the convolution layer, and they are responsible for extracting the features by executing the convolution operation between the input and the kernel matrix. Edges, horizontal lines, borders, bends, and vertical lines, among other things, are extracted using feature detectors in this case for efficient segmentation purpose. Further, rectified linear unit (ReLU) is also used to select segmented region.

C. Feature extraction and classification

After the segmentation, feature extraction plays the major role in SLDC operation. Features are the statistical parameters, which holds the different attributes of skin lesions based on their individual classes. The conventional image processing-based feature extractors are failed to extract the detailed ABCDE from the huge datasets and they are supported to low level feature extraction on small datasets.

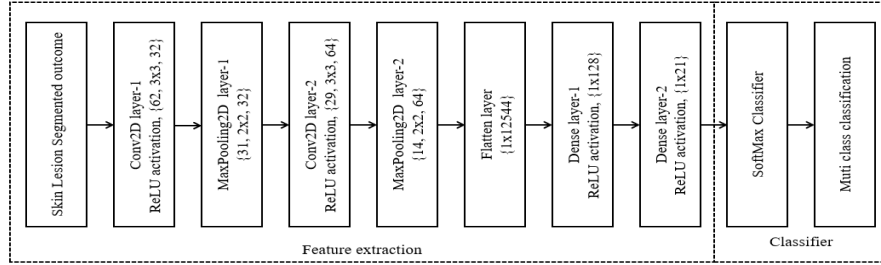


Figure 4. DLCNN feature extraction and classifier.

The DLCNN models are capable of extracting the detailed spatial, spectral, texture, and color features from the segmented images with high correlation. The DLCNN models trained with these features and performs the SLDC operation. Table 3 gives a full analysis of each layer, including the layer's dimension, filter size or kernel size, number of filters, and parameters, as well as the number of filters and parameters.

Table 3: Layer wise analysis of DTLNet.

Layer name	Layer dimension	Filter size	filter s	Paramete rs
Conv2D-1	62x62	3x3	32	896
MaxPooling2D-1	31x31	2x2	32	0
Conv2D-2	29x29	3x3	64	18496
MaxPooling2D-2	14x14	2x2	64	0
Flatten	1x12544	-	-	0
Dense-1	1x128	-	-	1605760
Dense-2	1x21	-	-	2709
SoftMax	1x8	-	-	0

Convolutional Layer: It is the most important operational block in DLCNN since it is responsible for performing the convolution operation between the skin lesion and the weight matrix and for generating local features. The mathematical operation of convolution layer is given as follows:

$$F_c(i, j) = \sum_m^M \sum_n^N f_s(i + m, j + n) K(m, n) \quad (1)$$

In this case, the input image or matrix is designated by f_s , the 2D filter is denoted by k with $m \times n$ as the filter size, and the 2D feature map output is marked by F_c . In this case, the 2D feature map output is denoted by F_c . After the convolution layer is completed, the resulting output is applied to the ReLU-based activation function, which adds the non-linearity connection between distinct features. The following relation illustrates the mathematical analysis of the ReLU activation function (f_{relu}).

$$f_{relu}(x) = \max(0, x) \quad (2)$$

MaxPooling Layer: In the DLCNN environment, the MaxPooling layer is used for down sampling purposes. It is used to reduce the input spatial size and also reduces the network parameters by a factor of two

SoftMax Classifier: All of the layers given in the proposed DTLNet architecture are layered together to form a DLCNN model that can recognize multiple classes. Among the features of the proposed DLCNN model are a SoftMax classifier, which helps to minimize the complexity by reducing the training time, and which also distinguishes between skin cancer and skin lesion multiclass categorization. The following is the definition of the mathematical connection between these properties:

$$Output = ReLU(dot(input, kernel) + bias) \quad (3)$$

Here, the output holds the different classified classes of skin cancer.

4. Simulation Results

This section gives detailed simulation analysis of proposed DTLNet with respect to subjective and objective analysis. Further, the performance of the proposed method is compared with state of art approaches using same ISIC-2019 dataset.

A. Dataset

For the purpose of DTLNet training and testing, the publicly accessible and real-time ISIC-2019 challenge dataset is taken into consideration. This dataset comprises the BCN 20000 and HAM10000 images, which are both available on the internet.

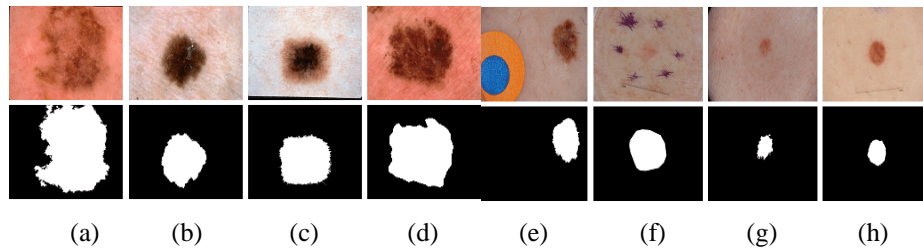


Figure 6. Segmented outcomes using AlexNet. (a) SCC, (b) VASC, (c) DF, (d) BKL, (e) AKIES, (f) BCC, (g) NV and (h) MEL

The BCN 20000 is a collection of 19424 images, each having a resolution of 1024 by 1024 pixels. The ISIC dataset comprises a total of 25,331 images, which are distributed in the following ways: SCC has 628 images, VASC has 253, DF has 239, BKL has 2,624 images, AKIEC has 867 images, BCC has 3,323 images, NV has 12,875 images, and MEL images have 4,522 images. Essentially, the whole dataset is divided into three

categories. They are as follows: 10% of the dataset is considered for testing, 10% for validation, and 80% of the dataset is considered for training. Figure 6 depicts the segmented outputs obtained via the use of the proposed AlexNet model.

B. Objective evaluation

The performance of any system cannot be estimated just on the basis of visual or subjective evaluations. In this way, objective assessment is quite beneficial for estimating the performance of different algorithms.

Segmentation performance analysis: An examination of the six objective factors for determining the segmentation effectiveness of AlexNet is presented in this article. They are segmentation accuracy (SACC), sensitivity (SEEN), specificity (SSPE), precision (SPR), recall (SRE), F1-Score (SF1).

Table 4 compares the segmentation performance of proposed AlexNet model with the existing methods such as PSPNet [6], U-Net [7], and MLR-Net [8]. The proposed method resulted in superior performance and also minimizes the vanishing gradient problems presented in the existing approaches.

Table 4. Performance comparison various segmentation approaches

Method	SACC (%)	SPR (%)	SRE (%)	SF1 (%)	SEEN (%)	SSPE (%)
PSPNet [6]	85.34	90.88	86.73	78.83	95.23	96.29
U-Net [7]	93.86	92.40	93.23	86.35	96.45	97.45
MLR-Net [8]	92.07	90.18	98.19	93.19	98.18	81.81
Proposed AlexNet	96.42	98.23	97.82	97.93	100	100

Classification Performance Analysis: This article considers classification accuracy, sensitivity, specificity, precision, recall, F1-Score for analyzing the classification performance of eight diseases. Table 5 compares the classification performance of proposed DTLNet with the various existing methods such as DLCNN [11], DenseNet-201[9], CNN-DG [10] and HCNN [12]. These existing approaches were implemented with the segmentation and classification models only, whereas the proposed DTLNet model contains preprocessing, segmentation, feature extraction, and classification steps. Finally, proposed DTLNet is outperformed due to its novel design

Table 5. Classification performance comparison of various SLDC methods

Method	Accuracy (%)	Precision (%)	F1-Score (%)	Sensitivity (%)	Specificity (%)
DLCNN [11]	81	74	77.45	83.4	84.84
DenseNet-201[9]	85.8	82.4	84.67	89.3	84.63
CNN-DG [10]	86.2	87.2	78.14	91.5	90.39
HCNN [12]	95.39	93.24	92.28	90.2	93.48
Proposed DTLNet	96.42	98.23	97.93	100	100

5. Conclusion

This research proposed a DTLNet model, which is deep transfer learning-based method and it is used to perform the preprocessing, segmentation and classification operations. Initially, HGWF is developed for removal of noises and also enhances the skin lesions. Then, transfer learning based AlexNet is used for skin lesion segmentation, which is accurately localize the area of disease effected regions. In addition, DLCNN model is developed for the extracting the deep features and SoftMax classifier is utilized to classify the eight different classes of skin lesion. The proposed DTLNet model is capable of classifying the multiple classes of skin lesion including SCC, VASC, DF, BKL, AKIES, BCC, NV and MEL. Simulation results shows that the performance of proposed DTLNet resulted in superior as compared to the conventional methods. This work can be extended to implement with bio-optimization approaches for enhancing the classification accuracy.

References

- [1]. Ganster, H.; Pinz, P.; Rohrer, R.; Wildling, E.; Binder, M.; Kittler, H. Automated melanoma recognition. *IEEE Trans. Med. Imaging* 2001, 20, 233–239.
- [2]. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* 2014, arXiv:1409.1556.
- [3]. Chollet, F. Deep learning with separable convolutions. *arXiv* 2016, arXiv:1610.02357.
- [4]. Razmjoooy, N.; Ashourian, M.; Karimifard, M.; Estrela, V.V.; Loschi, H.J.; Do Nascimento, D.; França, R.P.; Vishnevski, M. Computer-aided diagnosis of skin cancer: A review. *Curr. Med. Imaging* 2020, 16, 781–793.
- [5]. Li, Y.; Shen, L. Skin lesion analysis towards melanoma detection using deep learning network. *Sensors* 2018, 18, 556.
- [6]. Zou, J., Ma, X., Zhong, C., Zhang, Y.: Dermoscopic image analysis for ISIC challenge (2018). *arXivPrepr arXiv:180708948*
- [7]. Abraham, Nabila, and NaimulMefraz Khan. "A novel focal tversky loss function with improved attention u-net for lesion segmentation." *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, 2019.
- [8]. D. M. Krishna, S. K. Sahu and G. R. L. V. N. Srinivasa Raju, "MLRNet: Skin Lesion Segmentation using Hybrid Gaussian Guided Filter with CNN," 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2021, pp. 1337-1343, doi: 10.1109/ICECA52323.2021.9676020.
- [9]. Guissous, AllaEddine. "Skin lesion classification using deep neural network." *arXiv preprint arXiv:1911.07817* (2019).
- [10]. Sun, Qilin, et al. "Skin Lesion Classification Using Additional Patient Information." *BioMed Research International* 2021 (2021).
- [11]. Kassem, Mohamed A., Khalid M. Hosny, and Mohamed M. Fouad. "Skin lesions classification into eight classes for ISIC 2019 using deep convolutional neural network and transfer learning." *IEEE Access* 8 (2020): 114822-114832.
- [12]. A. Mahbod, G. Schaefer, C. Wang, G. Dorffner, R. Ecker, and I. Ellinger, "Transfer learning using a multi-scale and multinet ensemble for skin lesion classification," *Computer Methods and Programs in Biomedicine*, vol. 193, 2020.