



# DermoNet: A Computer-Aided Diagnosis System for Dermoscopic Disease Recognition

Ibtissam Bakkouri<sup>(✉)</sup> and Karim Afdel

LabSIV, Department of Computer Science, Faculty of Science,  
Ibn Zohr University, Agadir, Morocco  
ibtissam.bakkouri@gmail.com, k.afdel@uiz.ac.ma

**Abstract.** The research of skin lesion diseases is currently one of the hottest topics in the medical research fields, and has gained a lot of attention on the last few years. However, the existing skin lesion methods are mainly relying on conventional Convolutional Neural Network (CNN) and the performance of skin lesion recognition is far from satisfactory. Therefore, to overcome the aforementioned drawbacks of traditional methods, we propose a novel Computer-Aided Diagnosis (CAD) system, named DermoNet, based on Multi-Scale Feature Level (MSFL) blocks and Multi-Level Feature Fusion (MLFF). Further, the DermoNet approach yields a significant enhancement in terms of dealing with the challenge of small training data sizes in the dermoscopic domain and avoiding high similarity between classes and overfitting issue. Extensive experiments are conducted on the public dermoscopic dataset, and the results demonstrate that DermoNet outperforms the state-of-the-art approaches. Hence, DermoNet can achieve an excellent diagnostic efficiency in the auxiliary diagnosis of skin lesions.

**Keywords:** Dermoscopic diseases · Dermoscopic pattern recognition · Multi-layer feature fusion · Multi-scale features · Multi-class classification

## 1 Introduction

Nowadays, skin disease is one of the most difficult diseases to cure [1]. It is well-known that more early diagnosis of skin lesions is a crucial issue for the dermatologists [2]. Over the last few years, skin lesion recognition in dermoscopy images based on deep learning approach has attracted the attention of many researchers. Authors of [4] presented a computer aided diagnosis system to classify abnormal skin lesions in dermoscopic image into melanocytic nevus, seborrheic keratosis, basal cell carcinoma or psoriasis. The proposed method was evaluated on clinical database originated from the dermatology department of Peking Union Medical College Hospital, and the feature extraction and classification were carried out by GoogleNet Inception-V3 based on domain expert knowledge.

Inspired by the layers in the deep networks [5], authors of [6] applied ResNet-50 to extract discriminative features and Support Vector Machine (SVM) with a chi-squared kernel to classify melanoma images. This method was evaluated on the International Skin Imaging Collaboration (ISIC) 2016 Skin lesion challenge dataset, and it reached promising performance. In [7], an error-correcting output coding classifier combined with SVM for the multi-class classification scenario was introduced. The authors investigated the pre-trained AlexNet convolutional neural network model as feature extractor. VGG-16 architecture with fine-tuning technique was suggested to recognize melanoma abnormality in [8]. The proposed method was evaluated on ISIC 2017 challenge dataset, yielding satisfactory results. In this context, the similar approach based on convolutional multi-layer feature fusion network and continuous fine-tuning of CNNs using VGG-16, ResNet-18 and DenseNet-121 architectures were suggested by authors of [9] in order to classify seven skin diseases. The proposed method was evaluated on the testing subset of the HAM10000, and it reached promising results. A convolutional neural-adaptive network for melanoma recognition was proposed by authors [10]. The authors tried to adapt the pre-trained low-level weights on dermoscopic dataset using AlexNet architecture and fine-tuning technique. The images were obtained from ISIC dataset, and the proposed system provided efficient results.

Despite the high achievement of deep CNN for skin lesion diagnosis, it needs some improvement concerning the multi-class classification task, and overcome many problems that confronted most of previous algorithms such as semantic feature ambiguity and overfitting issue. Therefore, to overcome the aforementioned drawbacks of traditional methods, we propose a new mechanism based on CNN for skin lesion recognition, called DermoNet, in order to significantly improve the effectiveness and efficiency of computer-aided diagnosis system. The main contribution, novelty and characteristic of DermoNet are summarized as follows.

- An effective multi-scale feature architecture, called MSFL, is introduced to formulate a robust dermoscopic lesion representation for coping with the complex morphological structures of abnormal skin tissues.
- A new mechanism of feature fusion in CNNs, named MLFF, is proposed and applied for skin lesion classification to overcome the problem of semantic feature ambiguity and high inter-class visual similarities between the classes, and deal with the problem of having small set of dermoscopic data samples.
- Assessment of performance of the proposed methodology on test subset against the backbone as MSFL network without MLFF architecture. five statistical measurements were used for comparison of performances including accuracy, sensitivity, specificity, precision and F1-Score metrics.

The proposed framework has been validated on the public Human Against Machine with 10000 training images (HAM10000) dataset [3], which is designed for skin lesion diagnosis, and has been compared to state-of-the-art skin pattern classification approaches. The performance evaluation analysis shows the

potential clinical value of the proposed framework. To the best of our knowledge, no prior work exists on skin lesion characterization based on convolutional multi-scale feature fusion network that is investigated here.

The remainder of this paper is organized as follows: In Sect. 2, we detail the proposed approach. Then, we present the results of experiments realized on HAM10000 dataset in Sect. 3. Finally, we conclude this paper in Sect. 4.

## 2 Proposed DermoNet Framework

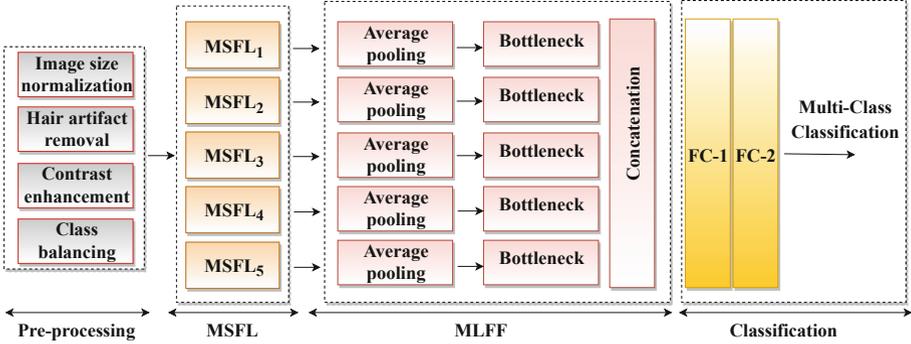
The main contribution of this study is to evaluate a novel CAD approach, called DermoNet, based on multi-scale CNN features combined with multi-level feature fusion for classifying the skin structures as Melanoma (MEL), Melanocytic Nevus (NV), Basal Cell Carcinoma (BCC), Actinic Keratosis (AKIEC), Benign Keratosis (BKL), Dermatofibroma (DF), or Vascular Lesion (VASC). Our proposed methodology based CAD system was developed in several stages. First, the representative square regions with size of  $64 \times 64$  pixels were selected from the digitized dermoscopic images and enhanced by the pre-processing techniques. Then, we built MSFL blocks to extract the most significant features. Inspired by feature fusion, to resolve the challenge of dealing with limited availability of the training dermoscopic samples, improve the performance of our model and avoid overfitting, the multi-level feature fusion was proposed and the extracted features from the last layer of each MSFL block were preserved by applying MLFF architecture to avoid missing useful information during convolution. Finally, the multi-class classification was performed using two Fully Connected layers (FC-1 and FC-2). An overview of the proposed DermoNet approach is given in Fig. 1.

### 2.1 Pre-processing of Dermoscopic Data

As indicated in our previous work [9], pre-processing of the digital dermoscopic data is performed in four stages, including data size normalization technique, hair artifact removal, intensity correction process and class balancing to facilitate application of DermoNet algorithm. The determinant of the Hessian combined with pyramidal REDUCE decomposition was carried out to normalize data size. To reduce the effects of hair artifact and enhance dermoscopic image brightness, we employed a computerized method as described in [11] and local adaptive bi-histogram equalization [12], respectively. As the dermoscopic image volumes of different categories vary widely, data augmentation is needed to balance the image volumes of different classes. In this paper, we balanced the dataset by augmenting the data for the minority classes, applying geometric transformation techniques as indicated in [9]. Finally, after data preparation process, the number of dermoscopic images is expanded at least by a factor of 20.

### 2.2 Multi-Scale Feature Level (MSFL) Extraction

The suggested multi-scale CNN is quite different from previous conventional CNN [13–15]. Rather than using a stack of consecutive convolutional layers in



**Fig. 1.** Flow chart of the proposed DermoNet method. The multi-scale feature architecture consists of five MSFL blocks. The proposed MLFF architecture is composed by three types of layer: average pooling, bottleneck and concatenation layers.

each block, we used a multi-scale convolutional layer per block with a relatively strong depth throughout the whole net, introducing a merge layer to maintain the information flow and ensure that no feature is lost along the network. Multi-scale layers allow for preserving details of different sizes from the input images. Moreover, we used Leaky Rectified Linear Unit (LReLU) activation function to avoid the zero gradients in negative part and increase the bias shift. Batch Normalization (BN) was also used to fight the internal covariate shift problem. The proposed multi-scale CNN method was evaluated in five blocks denoted by  $MSFL_n$ , where  $n$  is the number of blocks. As shown in Fig. 2, in each MSFL block, four scales of convolutional layers with kernel shape of  $[3 \times 3] \times 128$ ,  $[5 \times 5] \times 128$ ,  $[7 \times 7] \times 128$  and  $[9 \times 9] \times 128$ , LReLU activation function, BN and element-wise multiplication layers are involved. The multi-scale feature blocks were randomly initialized using the Xavier weight initialization algorithm and trained with the dermoscopic dataset.

### 2.3 Multi-Level Feature Fusion (MLFF)

Recent research in feature fusion in CNN has a great effect on pattern recognition which can cope with the problem of semantic feature ambiguity and high inter-class visual similarities between the classes and deal with the problem of having small set of dermoscopic data samples with very short calculation times and low computational costs. The proposed MLFF stores hierarchical multi-scale features where each level of MLFF preserves the output of the last layer of the MSFL block after applying average pooling operation with stride  $S = 2$  and kernel size  $2 \times 2$  and bottleneck layer with ratio  $\alpha = 0.125$ . The fusion layer is used to ensure maximum information flow between layers in the network. The MLFF method is first applied to skin lesion recognition problem and the model provides a solution to the most complex structures, like melanoma and basal cell carcinoma, perfectly. It can not only solve the overfitting issue like others

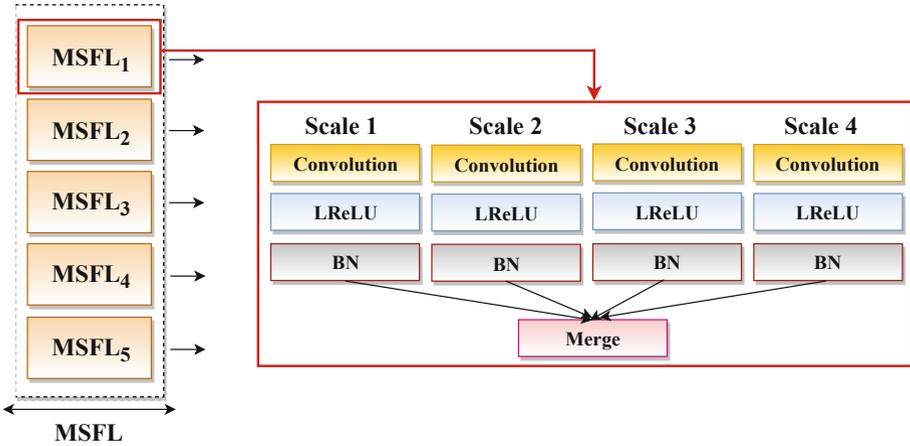


Fig. 2. Structure of a basic multi-scale feature block.

conventional methods, but also has a very simple structure which makes our model much easier to train.

### 3 Experimental Results

#### 3.1 Data Acquisition

In our experiments, we evaluated the performance of our proposed system on HAM10000 dataset which is the official ISIC 2018 challenge dataset hosted by the annual MICCAI conference in Granada, Spain, but is also available to research groups who do not participate in the challenge [3]. This dataset was collected over a period of 20 years from two different sites, the department of dermatology at the medical university of Vienna, Austria and the skin cancer practice of Cliff Rosendahl in Queensland, Australia. It includes seven generic classes: MEL, NV, BCC, AKIEC, BKL, DF and VASC. The dataset consists of 10015 dermoscopic images with size of  $600 \times 450$  showing skin lesions which have been diagnosed based on expert consensus, serial imaging and histopathology. The distribution is completely imbalanced, and then we applied the geometric transformation techniques as described in [9] to overcome class prevalence. After balancing process, we obtained 2000 images per class. The typical data set proportions are 60% for training, 20% for validation and 20% for testing [9].

#### 3.2 Experiment Settings and Implementation

The proposed system was trained using Adam algorithm with learning rate  $\alpha = 10^{-3}$ , first moment-decay  $\beta_1 = 0.9$  and second moment-decay  $\beta_2 = 0.999$  with Xavier algorithm for weight initialization and LReLU activation function

inserted after each convolutional layer. The softmax activation function was combined with the cross-entropy loss for training the network. The optimization ran for 30 epochs with mini batch size of 32 samples. To help our models generalize better in these circumstances, prevent over-fitting, we introduce dropout layer after the first fully connected layer with factor  $\rho = 0.5$ . The proposed method has been developed using the Keras library with python wrapper, and Compute Unified Device Architecture (CUDA) enabled parallel computing platform to access the computational resources of Graphics Processing Unit (GPU). The available hardware, used for training, is a PC with a Core i7 CPU, 8 GB RAM and a single NVIDIA GeForce GTX 1060 with 6 GB memory. Training dermoscopic images using the convolutional multi-scale feature fusion algorithm took about 6 h , 48 min and 43.19s. For testing, on average, it takes 3.56s in GPU and 5.37s in CPU per image.

### 3.3 Classification Results

The evaluation of the performance of our proposed system was fulfilled with the statistical quality metrics such as accuracy (Acc), sensitivity (Sen), specificity (Spe), precision (Pre) and F1-Score (FSc). As shown in Table 1, we compare the results of the proposed method (MSFL+MLFF) with regard to MSFL without multi-level fusion based on the number of MSFL blocks ( $MSFL_n$ ) using all five quality metrics. As indicated in Table 1, considering  $MSFL_6$  without MLFF as a base model,  $MSFL_6$  achieved the accuracy of 96.72%, sensitivity of 92.14%, specificity of 95.97%, precision of 91.85% and F1-Score of 91.99%. Compared to these results,  $MSFL_5+MLFF$  made improvements by 1.61%, 2.07%, 2.97%, 3.32% and 2.69% in terms of accuracy, sensitivity, specificity, precision and F1-Score, respectively. Noticeably,  $MSFL_5+MLFF$  shows the highest improvement for skin lesion classification, outperforming the two competing methods: MSFL at all levels and four instances of MSFL+MLFF including  $MSFL_2+MLFF$ ,  $MSFL_3+MLFF$ ,  $MSFL_4+MLFF$  and  $MSFL_6+MLFF$ .

### 3.4 Comparison with State of the Art

Since research on HAM10000 dermoscopic dataset is still relatively scarce, we tried to implement some existing algorithms used for skin lesion recognition and test them on our dermoscopic dataset. Then, the [7, 9, 10, 16, 17] were implemented and tested on HAM10000 dataset. The visual comparison given in Table 2 demonstrates that our best performing network outperforms the pre-trained AlexNet with an error-correcting output coding classifier combined with support vector machine [7], convolutional neural-adaptive networks [10], pre-trained GoogleNet inception-V3 [16], pre-trained very deep residual networks with 152 layers [17] and convolutional multi-layer feature fusion network with continuous fine-tuning of CNNs using VGG-16, ResNet-18 and DenseNet-121 architectures [9].

**Table 1.** Quantitative result of different MSFL blocks with and without MLFF on testing subset in terms of accuracy, sensitivity, specificity, precision and F1-Score metrics.

Methods	MSFL <sub>n</sub>	Acc. (%)	Sen. (%)	Spe. (%)	Pre. (%)	FSc. (%)
MSFL	MSFL <sub>2</sub>	86.37	81.49	85.19	80.66	81.07
	MSFL <sub>3</sub>	89.11	83.70	82.56	80.01	81.81
	MSFL <sub>4</sub>	92.89	89.36	92.84	90.25	89.80
	MSFL <sub>5</sub>	96.34	91.68	95.76	92.12	91.89
	MSFL <sub>6</sub>	<b>96.72</b>	<b>92.14</b>	<b>95.97</b>	<b>91.85</b>	<b>91.99</b>
MSFL + MLFF	MSFL <sub>2</sub>	90.71	87.22	90.89	91.04	89.08
	MSFL <sub>3</sub>	92.95	89.56	92.18	90.65	90.10
	MSFL <sub>4</sub>	96.48	91.80	95.92	91.39	91.59
	MSFL <sub>5</sub>	<b>98.33</b>	<b>94.21</b>	<b>98.94</b>	<b>95.17</b>	<b>94.68</b>
	MSFL <sub>6</sub>	98.00	93.54	98.71	94.63	94.08

**Table 2.** Comparison between the proposed method and other algorithms using the HAM10000 dataset.

Methods	Acc. (%)	Sen. (%)	Spe. (%)	Pre. (%)	FSc. (%)
Method [16]	91.56	87.26	92.48	92.52	89.81
Method [7]	83.91	81.47	84.20	81.76	81.61
Method [10]	84.28	79.34	84.79	84.01	81.60
Method [17]	86.77	84.91	86.67	85.37	85.13
Method [9]	98.09	93.35	98.88	93.36	93.35
<b>Ours</b>	<b>98.33</b>	<b>94.21</b>	<b>98.94</b>	<b>95.17</b>	<b>94.68</b>

## 4 Conclusion

In this work, we have presented a novel CAD system, DermoNet, for practical skin lesion recognition. It is developed using the combination between MSFL and MLFF architectures. The aim of MSFL is to formulate a robust dermoscopic lesion representation for coping with the complex morphological structures of abnormal skin tissues, while the MLFF was applied to overcome the problem of semantic feature ambiguity, deal with the small set of dermoscopic data samples and avoid overfitting issue. We conducted extensive experiments on HAM10000 dataset, and the comparison of DermoNet model with existing methods shows that this adduced work outperforms the most representative state-of-the-art strategies. So, we can conclude that DermoNet may assist the dermatologists to make the final decision for the further treatment. In the future, we will continue our work in integrating useful features from an explainable CNN for other task recognition in medical imaging field.

## References

1. Hollestein, L., Nijsten, T.: An insight into the global burden of skin diseases. *J. Invest. Dermatol.* **134**, 1499–1501 (2014)
2. Chen, M., Zhou, P., Wu, D., Hu, L., Hassan, M., Alamri, A.: AI-Skin: skin disease recognition based on self-learning and wide data collection through a closed-loop framework. *Inf. Fusion* **54**, 1–9 (2019)
3. Tschandl, P., Rosendahl, C., Kittler, H.: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci. Data* **5** (2018)
4. Zhang, X., Wang, S., Liu, J., Tao, C.: Towards improving diagnosis of skin diseases by combining deep neural network and human knowledge. *BMC Med. Inform. Decis. Making* **18** (2018)
5. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)
6. Yu, Z., et al.: Melanoma recognition in dermoscopy images via aggregated deep convolutional features. *IEEE Trans. Biomed. Eng.* **66**, 1006–1016 (2019)
7. Dorj, U.-O., Lee, K.-K., Choi, J.-Y., Lee, M.: The skin cancer classification using deep convolutional neural network. *Multimed. Tools Appl.* **77**(8), 9909–9924 (2018). <https://doi.org/10.1007/s11042-018-5714-1>
8. Thao, L., Quang, N.: Automatic skin lesion analysis towards melanoma detection. In: 2017 21st Asia Pacific Symposium on Intelligent and Evolutionary Systems (IES) (2017)
9. Bakkouri, I., Afdel, K.: Computer-aided diagnosis (CAD) system based on multi-layer feature fusion network for skin lesion recognition in dermoscopy images. *Multimed. Tools Appl.* (2019). <https://doi.org/10.1007/s11042-019-07988-1>
10. Bakkouri, I., Afdel, K.: Convolutional neural-adaptive networks for melanoma recognition. In: Mansouri, A., El Moataz, A., Nouboud, F., Mammass, D. (eds.) ICISP 2018. LNCS, vol. 10884, pp. 453–460. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-94211-7\\_49](https://doi.org/10.1007/978-3-319-94211-7_49)
11. Lee, T., Ng, V., Gallagher, R., Coldman, A., McLean, D.: Dullrazor®: a software approach to hair removal from images. *Comput. Biol. Med.* **27**, 533–543 (1997)
12. Tang, J., Mat Isa, N.: Adaptive image enhancement based on bi-histogram equalization with a clipping limit. *Comput. Electr. Eng.* **40**, 86–103 (2014)
13. Yong, L., Bo, Z.: An intrusion detection model based on multi-scale CNN. In: 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC) (2019)
14. Li, J., Zhang, R., Li, Y.: Multiscale convolutional neural network for the detection of built-up areas in high-resolution SAR images. In: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (2016)
15. Alhichri, H., Alajlan, N., Bazi, Y., Rabczuk, T.: Multi-scale convolutional neural network for remote sensing scene classification. In: 2018 IEEE International Conference on Electro/Information Technology (EIT) (2018)
16. Esteva, A., et al.: Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–118 (2017)
17. Han, S., Kim, M., Lim, W., Park, G., Park, I., Chang, S.: Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. *J. Invest. Dermatol.* **138**, 1529–1538 (2018)