IET Image Processing

Research Article

IET Journals
The Institution of Engineering and Technology

Wavelet-based deep learning for skin lesion classification

ISSN 1751-9659
Received on 6th May 2019
Revised 8th October 2019
Accepted on 13th November 2019
E-First on 19th February 2020
doi: 10.1049/iet-ipr.2019.0553
www.ietdl.org

Sertan Serte[1] ✉, *Hasan Demirel*[2]

[1]*Electrical and Electronic Engineering, Near East University, Nicosia, North Cyprus via Mersin 10, Turkey*
[2]*Electrical and Electronic Engineering, Eastern Mediterranean University, Famagusta, North Cyprus via Mersin 10, Turkey*
✉ *E-mail: sertan.serte@gmail.com*

Abstract: Skin lesions can be in malignant or benign forms. Benign skin lesion types are not deadly; however, malignant types of skin lesions can be fatal. Lethal forms are known as skin cancer. These types require urgent clinical treatment. Fast detection and diagnosis of malignant types of skin lesions might prevent life-threatening scenarios. This work presents two methods for the automatic classification of malignant melanoma and seborrhoeic keratosis lesions. The first method builds on modelling skin images together with wavelet coefficients. Approximate, horizontal, and vertical wavelet coefficients are obtained using the wavelet transform, and then deep learning (DL) models are generated for each of the representations and skin images. The second method builds on modelling skin images together with three approximate coefficients. This method utilises a sequential wavelet transformation to produce approximation coefficients. Then DL models are generated for each of the representations and skin images. Transfer learning-based ResNet-18 and ResNet-50 DL models provide model images and wavelet coefficients. Then skin lesion detection is achieved by fusing model output probabilities. Both proposed models outperform the methods only based on image data and other previously proposed methods.

1 Introduction

Skin cancer is a malignant type of skin lesion that can cause death. Melanoma is one of the skin cancer types, and it can lead to more life-threatening situations than other forms of skin cancer. This malignant type also causes more deaths than different skin cancer types [1–3]. Recent surveys [4–7] have presented automated skin lesion classification methods.

Melanoma and naevus are known as melanocytic skin lesions. Basal cell and squamous cell carcinoma are non-melanocytic lesions. Bowen's disease and actinic keratosis are treated similarly. In contrast, dermatofibroma and vascular are benign types of non-melanocytic deformations. Although basal-cell and squamous-cell carcinoma occur more frequently, melanoma is responsible for a greater number of deaths [1, 4].

Automated fast detection of skin lesions can be achieved using deep convolutional neural networks (CNNs). Examples of different CNNs include AlexNet [8], GoogleNet [9, 10], VGG [11], ResNet [12], and DenseNet [13]. Recently, these models have provided the classification of 1000 objects in the ImageNet dataset [14]. The performance results show that these models are close to human-level object-level accuracy. These models also result in high-classification performance in medical image classification. The authors of [15, 16] used CNNs to model automated age-related macular degeneration and diabetic macular oedema detection. Kaymak and Serener [17] also proposed a two-step classification of malignant pigmented skin lesions.

Skin lesion classification methods can mainly be divided into two groups. First, Esteva *et al*. [1] and Han *et al*. [2] used a single CNN model for skin lesion classification tasks. Second, an ensemble of CNNs was proposed to detect skin lesion types. The recent studies have reported that a group of CNNs outperformed a single CNN model for skin lesion classification [18–26]. These ensemble methods are based on creating several sets of skin images and then generating different DL models for each of the collections. The lesion prediction is made by fusing the output probabilities of the models.

The proposed wavelet and approximation coefficient-based methods are different than previously proposed methods. These methods model both skin lesion images as well as wavelet coefficients for more accurate detection of skin lesion type. The proposed wavelet coefficient-based method captures both colour and shape information of the skin image, together with horizontal and vertical details of the skin lesions. More information is extracted from the skin images and modelled by ResNet-18 and ResNet-50 models. Finally, the combining of image data and wavelet coefficients is achieved by using the probabilistic model.

The main contributions of this work are as follows:

i. A novel wavelet coefficient-based DL method is proposed for melanoma and seborrhoeic keratosis detection. This method models information on approximate, horizontal, and vertical coefficients of skin images in conjunction with skin images for accurate classification.
ii. A novel approximation coefficient-based DL method is proposed for melanoma and seborrhoeic keratosis detection. This method models three approximation images in conjunction with skin images for more robust skin lesion classification.

The organisation of this paper is as follows. First, related work regarding skin lesion classification will be given. Second, the proposed wavelet and approximation-based DL methods are described. An associated database of skin lesions is presented. Finally, the performance of the proposed methods is evaluated and discussed.

2 Related work

Skin lesion classification approaches can be divided into two categories. These categories are single DL-based methods and an ensemble of DL methods.

2.1 Single DL-based methods

Esteva *et al*. [1] proposed an approach for the classification of skin lesions using the Inception v3 model. The proposed Inception v3 model builds on three skin lesion datasets. This model detects malignant and benign types of skin lesions. Furthermore, the proposed model provides the detection of keratinocyte carcinomas

and benign seborrhoeic keratoses skin types. The presented approach shows higher performance than dermatologists. Then, skin lesions are recognised using the model.

Similarly, Han *et al*. [2] proposed ResNet-152 for the classification of skin lesions. This model is fine-tuned for the classification of basal cell carcinoma, squamous cell carcinoma, intraepithelial carcinoma, actinic keratosis, seborrhoeic keratosis, malignant melanoma, melanocytic naevus, lentigo, pyogenic granuloma, haemangioma, dermatofibroma, and warts. This study uses several datasets of different ethnic groups. This study shows that ethnicity and image contrast are the factors that decrease the classification accuracy of the skin lesions.

Menegola *et al*. [27] used six publicly available databases for skin lesion recognition. Their approach involved training Deep ResNet-101 and Google Inception-v4 models for malignant melanoma, seborrhoeic keratosis, and naevus skin lesion classification. This study shows that increasing training images by combining databases results in more accurate lesion classification.

Zhang *et al*. [28] proposed attention-based residual learning for melanoma, seborrhoeic keratosis, and naevus skin lesion classification. This method builds on the ResNet DL model, and the attention-based representations model the last layers of the ResNet.

Fujisawa *et al*. [29] proposed four-level skin tumour detection. This method allows the classification of 21 skin lesion types. Skin lesion images are modelled using the GoogleLeNet, and then skin images are classified into four levels. First, malignant and benign types are recognised, and then epithelial, melanocytic categories are organised. Third, actinic keratosis, seborrhoeic keratosis, basal cell carcinoma, and Bowen disease are classified. Finally, these types can be further divided into pigmented and non-pigmented types.

2.2 Ensemble of the DL-based methods

Matsunaga *et al*. [18] proposed a deep neural network ensemble for melanoma, seborrhoeic keratosis, and naevus classification. Their proposed method builds on two binary classifiers. First, a classifier is generated to classify melanoma and other lesions. Second, a classifier is generated to classify seborrhoeic keratosis and the rest of the lesions. Classifiers are created using different sets of skin lesion images and ResNet-50 models. Lesion classification is performed by fusing output the probabilities of the models.

Harangi [19] proposed an ensemble of CNN models for the classification of melanoma, seborrhoeic keratosis, and naevus. The proposed method is based on fusing the output probabilities of the GoogLeNet, AlexNet, ResNet, and VGGNet models. Four probabilistic fusion techniques are proposed for joining the models. These are the sum of the probabilities, product of the probabilities, simple majority voting, and sum of the maximal probabilities. The sum of the maximal probabilities method leads to higher performance than the other fusion techniques. This study also shows that joined models perform better than a single CNN model for skin lesion classification.

Mahbod *et al*. [20] proposed fusing fine-tuned deep features for skin lesion classification. This method builds on training several fine-tuned CNN models and extracting features from the last fully connected layers. Then, these features are fused, and lesion classification is performed using support vector machines.

Nyiri and Kiss [21] propose an ensemble CNN for dermatological image classification. First, VGG16, VGG19, Xception, Inception, ResNet, and DenseNet models are generated using skin images. Then, these models are also generated using segmented skin images. Then skin image and segmented skin image-based CNN models are fused for lesion prediction. The results show that an ensembling of all models resulted in better performance than a single CNN model for lesion prediction. This study also demonstrated that combining CNNs outperforms each of the single CNN models.

Nozdryn-Plotnicki *et al*. [22] generated different DL models on skin lesion images and then modelled probabilistic outputs of the models for lesion classification. The probabilistic fusion method is based on reweighing the prior probabilities and then balance distribution probabilities are used for lesion classification.

Li and Shen [23] proposed a method for the classification of melanoma, seborrhoeic keratosis, and naevus skin lesions. The study is based on using two ResNet models to obtain two fully connected probability maps. Then these maps are fused for lesion type classification.

Gonzalez [24] proposed the integration of several DL models for the classification of melanoma, seborrhoeic keratosis, and naevus. These models are the lesion segmentation model, structure segmentation model, and diagnosis model. The given skin lesion image is segmented and then used as inputs to the structure segmentation model for further feature representation. The diagnosis model uses all features to predict the final skin lesion type.

Bi *et al*. [25] proposed an ensemble method for melanoma, seborrhoeic keratosis, and naevus classification. First, a CNN model is proposed to model three skin lesions. Second, one CNN model performs melanoma detection, and another CNN model is generated to detect the seborrhoeic keratosis class. Finally, the above generated models are fused for the classification of melanoma, seborrhoeic keratosis, and naevus skin lesions.

Recently, Hagerty *et al*. [26] proposed a different approach than the above mentioned studies. Their approach provides six features from skin lesion images. Then ResNet-50 CNN models these feature representations to obtain six feature vectors. The obtained feature vectors are mapped to skin lesion types using logistic regression.

3 Proposed wavelet-based methods

The proposed wavelet and approximation-based methods allow the classification of melanoma, seborrhoeic keratosis, and melanocytic naevus. These methods are based on decomposing skin lesion images into wavelet coefficients and then modelling these coefficients in conjunction with skin images. Modelling coefficients and skin images is achieved using ResNet-18 and ResNet-50 deep learning models. Then output probabilities are fused for lesion classification.

3.1 Preprocessing and implementation

Skin images are converted to grey-scale and then the proposed wavelet transform is used to decompose the images. The Matlab programme converts red–green–blue (RGB) images to grey-scale images. Wavelet transform is also implemented using the Matlab programme.

3.2 Wavelet transform

The wavelet transform technique is proposed to decompose skin lesion images into multiple resolutions [30]. These resolutions show the spatial and frequency characteristics of skin lesions. This method uses a skin lesion image as an input to obtain approximate, horizontal, vertical, and diagonal detail coefficients. An example of a decomposed wavelet coefficient can be seen in Fig. 1.

3.3 Wavelet coefficient-based method

Skin lesion images are transformed into four wavelet coefficients using the wavelet transform. Fig. 1 shows four transformed wavelet coefficients. These coefficients are approximate, horizontal, vertical, and diagonal coefficients. An approximate image is an approximation of the skin lesion image. Horizontal, vertical, and diagonal wavelet coefficients contain horizontal, vertical, and diagonal details of the skin images. After ResNet-18 and ResNet-50 models are generated using skin images, approximate, horizontal, vertical coefficients then output probabilities are fused. A DL-based model generation is described in Section 3.5, and then model fusion for skin lesion prediction is explained in Sections 3.10 and 3.2.

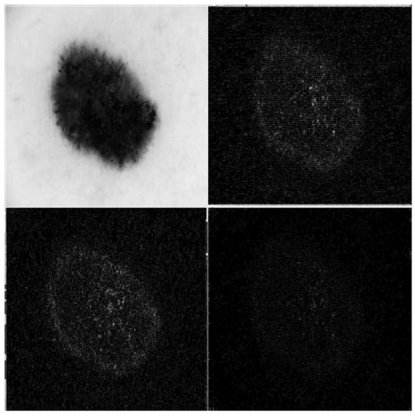

Fig. 1 *Skin image decomposition into approximate, horizontal, vertical, and diagonal images*

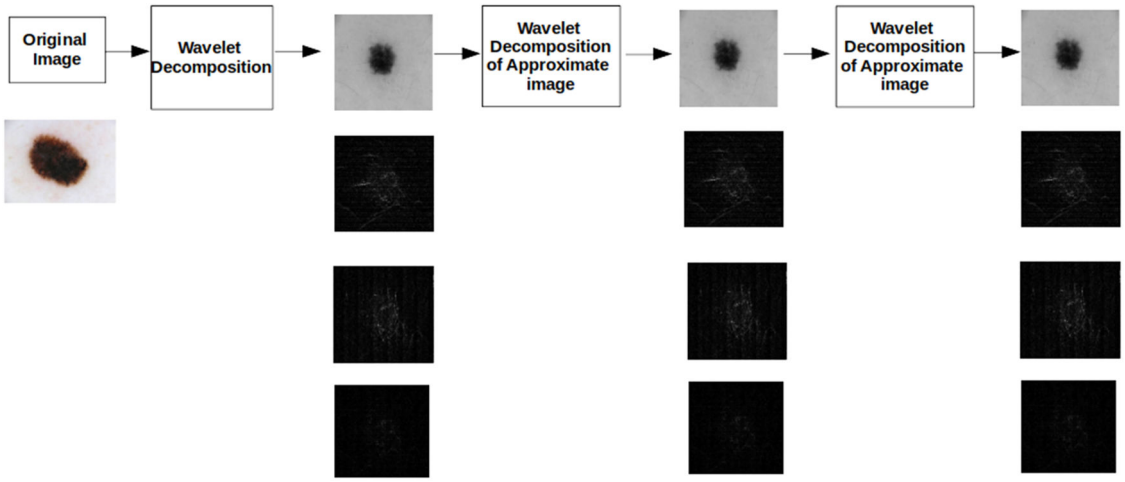

Fig. 2 *Sequential wavelet transformation on skin lesion images*

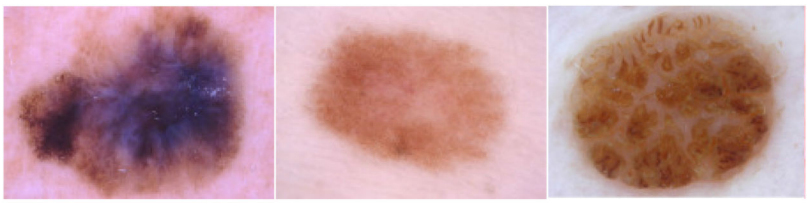

Fig. 3 *Skin lesion types in ISIC 2017 dataset [3]*
(a) Melanoma, ***(b)*** Nevus, ***(c)*** Benign keratoses

Table 1 ISIC 2017 dataset

Skin lesion	No. of images
melanoma	374
seborrhoeic keratosis	254
melanocytic naevus	1372
total	2000

3.4 Approximation coefficient-based method

The approximation coefficient-based DL method is generated in two steps. First, skin lesion images are decomposed into wavelet coefficients using the method described in Section 3.3, and then the first approximation image is obtained. Later the resultant approximation image is decomposed into wavelet coefficients to generate the second approximation image. Finally, the third approximation image is obtained by decomposing the second approximate image into wavelet coefficients. Fig. 2 shows this process. Then fine-tuned ResNet-18 and ResNet-50 DL models are created using skin lesion images in conjunction with three approximation coefficients. Output probabilities of the models are fused for skin lesion prediction. Model generation is described in Section 3.5, and model fusion is explained in Sections 3.10 and 3.2.

3.5 Model generations

The ResNet-18 and ResNet-50 DL models are generated using International Skin Imaging Collaboration (ISIC) 2017 training images. Skin lesion types in ISIC 2017 dataset can be seen in Fig. 3. This dataset consists of 374 melanoma, 254 seborrhoeic keratosis, and 1372 melanocytic naevus images. Table 1 shows the number of skin images of the ISIC 2017 training dataset. The wavelet transform is applied to all pictures of the ISIC 2017 training dataset. The new dataset consists of 374 melanoma skin images and corresponding wavelet coefficients. It also contains 254 seborrhoeic keratosis skin images and corresponding wavelet coefficients. Finally, it contains 1372 naevus images and corresponding wavelet coefficients.

Since the number of skin lesion images and corresponding wavelet coefficients might not be enough for estimating DL model parameters and the number of images are unbalanced, data augmentation is performed on the new dataset. The melanoma and

seborrhoeic keratosis images are rotated 18° while melanocytic naevus are rotated by 45°. The number of pictures in this augmented dataset is given in Table 2.

3.6 Data augmentation

The augmented dataset consists of 7854 approximation, 7854 horizontal, and 7854 vertical wavelet coefficients of melanoma and 7854 melanoma skin images. It also contains 5334 approximation, 5334 horizontal and 5334 vertical wavelet coefficients of seborrhoeic keratosis and 5334 seborrhoeic keratosis skin images. Finally, the dataset includes 12,348 approximation, 12,348 horizontal, and 12,348 vertical wavelet coefficients of melanocytic naevus and 12,348 melanocytic naevus skin images.

Each image of melanoma and seborrhoeic keratosis is rotated 18° between 0 and 360°. Then, 21 rotated images are obtained using a single skin image. Then, $21 \times 373 = 7854$ and $21 \times 254 = 5334$ augmented images are gathered for melanoma and seborrhoeic keratosis, respectively. Similarly, each image of melanocytic naevus is rotated 45° between 0 and 360°. In this case, nine rotated images are obtained using a single image. Then, $9 \times 1372 = 12,348$ augmented images are obtained for melanocytic naevus. Table 2 shows all the number of enlarged images.

We apply 18° image rotation for melanoma and seborrhoeic keratosis classes as augmentation but used 45° for the naevus class. The reason for this is that balanced training data for CNN is vital for training and estimating the CNN parameters correctly. The ISIC 2017 dataset is imbalanced, and the number of melanoma and seborrhoeic keratosis is much less than melanocytic naevus images. Therefore, the data augmentation technique is used for balancing the number of pictures and increasing the dataset. This is achieved by rotating both melanoma and seborrhoeic by 18 and melanocytic naevus by 45° between 0° and 360°.

Table 2 Augmented skin images and corresponding coefficients.

Skin lesion	No. of images
melanoma	7854
seborrhoeic keratosis	5334
melanocytic naevus	12,348
approximation coefficients of melanoma images	7854
approximation coefficients of seborrhoeic keratosis images	5334
approximation coefficients of melanocytic naevus images	12,348
horizontal coefficients of melanoma images	7854
horizontal coefficients of seborrhoeic keratosis images	5334
horizontal coefficients of melanocytic naevus images	12,348
vertical coefficients of melanoma	7854
vertical coefficients of seborrhoeic keratosis	5334
vertical coefficients of melanocytic naevus	12,348
total number of approximation coefficients	25,536
total number of horizontal coefficients	25,536
total number of vertical coefficients	25,536

3.7 Fine-tuned models

The ResNet-18 models are generated for skin images, approximate, horizontal, and vertical wavelet coefficients. The ResNet-50 models are also generated using skin images, approximate, horizontal, and vertical wavelet coefficients. The generated ResNet-18 and ResNet-50 models are based on modification of pre-trained ResNet-18 and ResNet-50 models. These models are trained on the ImageNet dataset, and then they are used to adapt the parameters to skin lesion images and wavelet coefficients. The ImageNet dataset contains approximately one million images, and this large number of images results in well- estimated network parameters. Fine-tuning pre-trained networks enable better estimates of skin lesion types.

3.8 Resizing

All skin images are resized to 256×256 RGB images, and then these images are used as inputs to the CNN models. The $224 \times 224 \times 3$ random crops are extracted, and these crops are modelled using ResNet-18 and ResNet-50 DL models.

3.9 Fine-tuning

ResNet-18 and ResNet-50 pre-trained models are fine-tuned for skin classification. The weights of the convolutional layers are frozen, and only fully connected layers are learned for skin lesion classification. During training, learning parameter, gamma, momentum, and weight decay values of 0.001, 0.1, 0.9, and 0.0001 are used, respectively.

3.10 Lesion prediction

Fig. 4 presents the proposed wavelet coefficient-based methods. A skin image is decomposed into approximate, horizontal, and vertical wavelet coefficients (Section 3.3 and Fig. 1). Then output probabilities are obtained by inputting each of the coefficients to the corresponding CNN models. Finally, these output probabilities are fused using the method described in Section 3.11.

Fig. 5 shows the proposed approximation coefficient-based method. The proposed sequential wavelet transformation on the skin image is performed to obtain three approximate images (Section 3.4 and Fig. 2). Then output probabilities are achieved by inputting each of the approximate photos to the corresponding CNN models. Finally, these output probabilities are fused using the method described in Section 3.11.

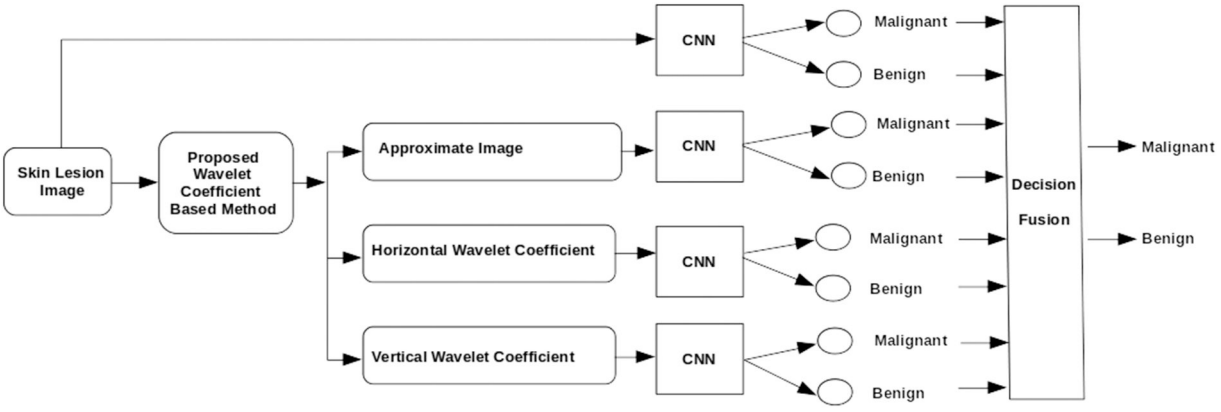

Fig. 4 *Skin lesion prediction using the proposed wavelet-based method*

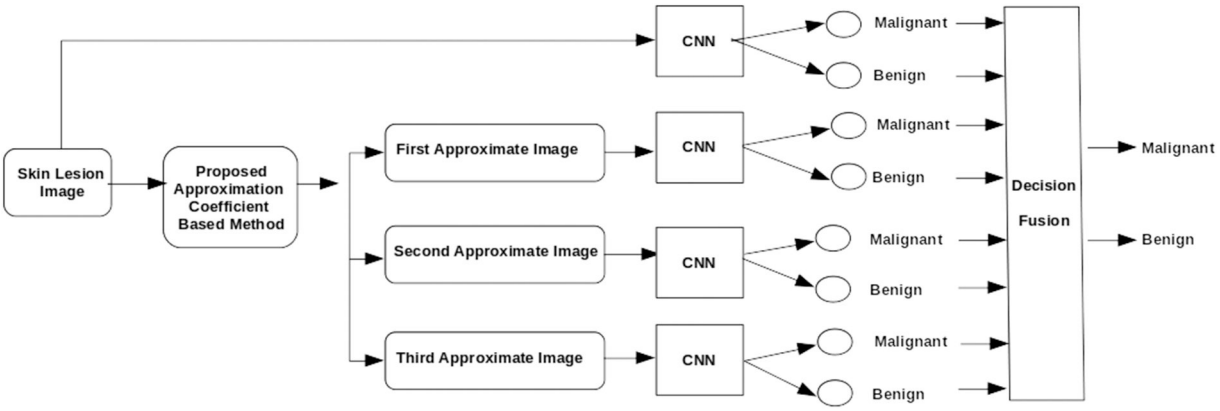

Fig. 5 *Skin lesion prediction using the proposed approximation coefficient-based method*

Table 3 Performance comparison of the proposed methods

Network	Method	AVG-AUC	M-AUC	M-ACC	M-SE	M-SP	SK-AUC	SK-ACC	SK-SE	SK-SP
ResNet-18	I	0.68	0.77	0.80	0.33	0.92	0.59	0.83	0.74	0.86
ResNet-50	I	0.81	0.85	0.87	0.33	1.00	0.76	0.79	0.29	0.98
ResNet-18	A1	0.80	0.81	0.80	0.20	0.95	0.78	0.73	0.31	0.90
ResNet-18	H	0.74	0.75	0.74	0.23	0.87	0.72	0.76	0.33	0.93
ResNet-18	V	0.71	0.79	0.79	0.33	0.90	0.63	0.77	0.40	0.91
ResNet-50	H	0.67	0.62	0.68	0.30	0.77	0.72	0.71	0.43	0.81
ResNet-50	V	0.82	0.83	0.77	0.27	0.89	0.81	0.71	0.26	0.89
ResNet-18	I-A1-H-V	0.89	0.96	0.84	0.23	0.99	0.81	0.79	0.36	0.96
ResNet-50	I-A1-H-V	0.77	0.74	0.83	0.30	0.96	0.80	0.79	0.29	0.99
ResNet-18	A2	0.86	0.88	0.83	0.23	0.98	0.83	0.74	0.21	0.94
ResNet-18	A3	0.86	0.85	0.83	0.27	0.98	0.87	0.73	0.21	0.93
ResNet-50	A1	0.82	0.78	0.77	0.33	0.88	0.85	0.79	0.38	0.94
ResNet-50	A2	0.68	0.73	0.78	0.13	0.94	0.63	0.66	0.55	0.70
ResNet-50	A3	0.82	0.69	0.75	0.23	0.88	0.95	0.71	0.10	0.94
ResNet-18	I-A1-A2-A3	0.82	0.82	0.83	0.27	0.98	0.82	0.81	0.40	0.96
ResNet-50	I-A1-A2-A3	0.92	0.96	0.83	0.13	1.00	0.87	0.79	0.26	0.99
ResNet-50-18	I-A1-A2-A3	0.89	0.91	0.85	0.30	0.99	0.87	0.84	0.50	0.97

3.11 Probabilistic fusion of DL models

The single CNN model provides probability values for melanoma, seborrhoeic keratosis, and naevus skin lesions. These probabilities can be denoted by p_i where $i = 1, \ldots, n$ and n is the number of output probabilities. In this case, the value of n is equal to 3.

An ensemble of CNNs also provides output probabilities. These probabilities are denoted by p_{ij} where $i = 1, \ldots, n$ and $j = 1, \ldots, m$. The value of m represents the number of CNN models. In this study, four CNN models are used, and the value of m is equal to 4.

The fusion of generated probabilities of CNN models is achieved by the sum of probabilities (SMP) [19, 31]. The SMP is denoted by

$$p_i = \frac{\sum_{j=1}^{m} p_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}} \quad i = 1, \ldots, n. \tag{1}$$

where p_{ij} indicates the i class probability of the CNN_j model.

3.12 Implementation

The proposed methods are carried out on a desktop computer. This computer utilises an Intel Core i7-4790 3.6 Hz CPU. It also uses 25 GB of memory. The data augmentation is done using the C++ programming language together with the OpenCV library. The ResNet-18 and ResNet-50 architectures are trained using an NVIDIA GeForce GTX 1080Ti GPU running a Caffe DL framework. The model training requires about 20 min on this system.

4 Performance evaluation

The average area under the receiver operating characteristic (ROC) curve of melanoma and seborrhoeic keratosis is denoted by AVG-AUC. The area under the ROC curve, accuracy, sensitivity, and specificity of the melanoma detection are denoted by M-AUC, M-ACC, M-SE, and M-SP, respectively. Similarly, the average area under the ROC, accuracy, sensitivity, and specificity of the seborrhoeic keratosis detection are denoted by SK-AUC, SK-ACC, SK-SE, and SK-SP, respectively.

The performance evaluation of the proposed methods is evaluated using the publicly available ISIC 2017 dataset. This set contains training, validation, and test image sets. The number of images in each set is 2000, 150, and 600, respectively. This dataset contains skin lesion images of melanoma, seborrhoeic keratosis, and naevus classes.

4.1 Wavelet coefficient-based method

The generated models for skin images and corresponding approximate, horizontal, and vertical coefficients are named I, A1, H, and V, respectively. Similarly, the generated fused models are named I-A1-H-V. All models are created using both ResNet-18 and ResNet-50 DL models. The performances of the proposed methods are reported in Table 3.

These results show that the best performing model is the ResNet-18-based I-A1-H-V fused model. The fusing wavelet coefficients, in conjunction with skin images, improve the performance.

Table 4 Performance comparison of the proposed methods and other methods

Network	Method	AVG-AUC	M-AUC	M-ACC	M-SE	M-SP	SK-AUC	SK-ACC	SK-SE	SK-SP
ResNet-18	I-A1-H-V	**0.89**	**0.96**	0.84	0.23	0.99	**0.81**	0.79	0.36	0.96
ResNet-50	I-A1-H-V	**0.77**	**0.74**	0.83	0.30	0.96	**0.80**	0.79	0.29	0.99
ResNet-18	I-A1-A2-A3	**0.82**	**0.82**	0.83	0.27	0.98	**0.82**	0.81	0.40	0.96
ResNet-50	I-A1-A2-A3	**0.92**	**0.96**	0.83	0.13	1.00	**0.87**	0.79	0.26	0.99
ResNet-50-18	I-A1-A2-A3	**0.89**	**0.91**	0.85	0.30	0.99	**0.87**	0.84	0.50	0.97
Haragi [26]	ensembles of CNNs	0.89	0.85	0.85	0.40	0.71	0.93	0.88	0.71	0.85
Zhang *et al.* [28]	attention based CNN	0.90	0.86	0.84	0.59	0.89	0.95	0.90	0.87	0.93
Li and Shen [23]	lesion indexing	—	0.91	0.86	0.49	0.96	—		—	—
Matsunaga *et al.* [18]	ensembles of CNNs	0.91	0.87	0.83	0.73	0.85	0.95	0.80	0.98	0.77
Diaz [24]	knowledgebased CNN	0.91	0.85	0.82	0.10	0.99	0.96	0.87	0.17	0.99
Menegola *et al.* [27]	ensembles of CNNs	0.91	0.87	0.87	0.54	0.95	0.94	0.89	0.35	0.99
Bi *et al.* [25]	ensembles of CNNs	0.89	0.87	0.85	0.42	0.96	0.92	0.91	0.58	0.97
Mahbod *et al.* [20]	deep features	0.91	0.87	—	—	—	0.95	—	—	—
Hagerty *et al.* [26]	handcrafted (HC) ensemble	—	0.90	—	—	—	—	—	—	—
Hagerty *et al.* [26]	HC and DL ensemble	—	0.94	—	—	—	—	—	—	—

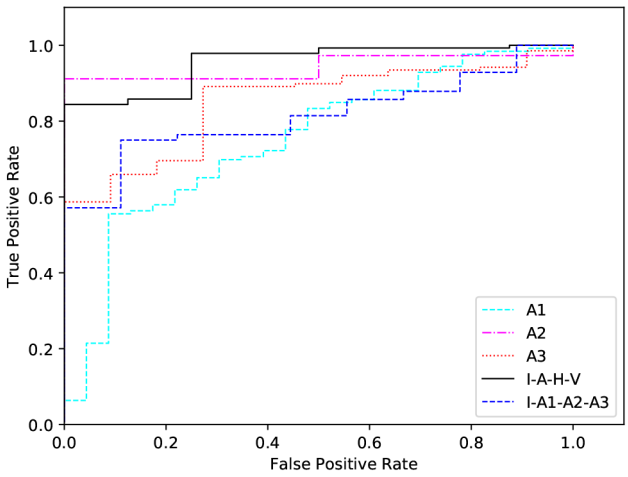

Fig. 6 *Melanoma detection*

4.2 Approximation coefficient-based method

The generated single models for skin images and three corresponding approximate coefficients are named as I, A1, A2, and A3, respectively. Then these individual models are fused and named as I-A1-A2-A3. All models are generated using both ResNet-18 and ResNet-50 DL models. The performances of these models are reported in Table 3.

ResNet-50-based I model and ResNet-18-based three approximate coefficient-based single models are fused. This proposed model is named as a ResNet-50-18 model. The performance of this new model is reported in Table 3. These results show that the best performing model is the ResNet-50-based I-A1-H-V fused model. The fusing approximation coefficients, in conjunction with skin images, improve the performance.

4.3 Comparison of the proposed methods

The comparison of the proposed methods is performed as follows. First, ResNet-18 and ResNet-50-based I models are compared with the proposed single wavelet models. The ResNet-50-based I model outperforms the ResNet-18-based I and wavelet-based A1, H, and V models. However, the proposed ResNet-18-based A1 and A2 models outperform the ResNet-50-based I model. Second, the generated single wavelet coefficient-based models are compared with the fused models. It can be seen from the results that the ResNet-18-based I-A1-H-V and ResNet-50-based I-A1-A2-A3 models outperform all single models. This shows that modelling the skin images in conjunction with wavelet coefficients improves the skin lesion classification performance. Finally, the ResNet-18-based I-A1-H-V and ResNet-50-based I-A1-A2-A3 models outperform the ResNet-50-based I model.

4.4 Comparison of the proposed and other methods

The performance comparison of the proposed methods and other methods is reported in Table 4. First, the performance comparison is performed for melanoma and seborrhoeic keratosis detection. The performances of the ResNet-18-based I-A1-H-V and ResNet-50-based I-A1-A2-A3 methods are close to other methods [18, 20, 23, 24, 26–28]. The ResNet-50-18-based I-A1-A2-A3 model also provides close AVG-AUC values to the other methods. Second, the performance comparison is performed for melanoma detection. The ResNet-18-based I-A1-H-V and the ResNet-50-based I-A1-A2-A3 models result in 96% and 96% M-AUC values, respectively. This shows that these proposed methods outperform other recent methods for melanoma detection.

The ROC curves of the proposed methods are also provided for melanoma detection in Fig. 6. The performance of approximation coefficient-based single A1, A2, A3 models is compared with the fused I-A1-H-V and I-A1-A2-A3. The I-A1-H-V performs better than other proposed methods.

5 Discussion

The only image-based results of the fine-tuned ResNet-18 model are reported as 0.68%, 0.77% and 0.59% for AVG-AUC, M-AUC, and SK-AUC, respectively (Table 3). These values are reported approximately as 0.87, 0.83, and 0.91% in [20]. The same fine-tuned ResNet-18 in [20] performed better than in this study. The reason is that Mahbod *et al.* [20] used external training data (ISIC 2016) in conjunction with ISIC 2017 training data for fine-tuning the ResNet-18 model (Section 2.1 [20]). Mahbod *et al.* [20] also used the grey world colour constancy algorithm for skin image colour standardisation. However, the proposed method is based only on the ISIC 2017 dataset. The proposed methods do not utilise any image colour standardisation. This could explain why the results reported in [20] might be better than the proposed methods. Similarly, the image-based fine-tuned ResNet-50 model performance is higher in [18] than in this study. The reason is that the proposed method in [18] was based on ISIC 2017 and an external dataset was used for fine-tuning the ResNet-50 model. On the other hand, this study utilises only the ISIC 2017 dataset.

The performances of the skin image and wavelet-based methods change according to data and deep learning models. This performance difference is also apparent for the proposed fused methods. This might be due to the learning parameters of the ResNet models. The selection of the best parameters for the best lesion classification performance might not be possible before training the models. Therefore, selecting the best performing models might be possible by generating several models and then selecting the best performing model. This strategy is used to select the best performing fused model of the lesion type classification. In this study, the ResNet-18-based I-A1-H-V and ResNet-50-based I-

A1-A2-A3 models are the best performing combined models for melanoma and seborrhoeic keratosis detection.

The proposed fused methods are compared with skin image-based and wavelet-based methods for melanoma and seborrhoeic keratosis detection. First, the combined methods are compared with only skin image-based methods. The results show that the fused methods outperform the image-based methods for both melanoma and seborrhoeic keratosis detection. Second, the fused methods are compared with only wavelet coefficient-based methods. The results show that the combined methods outperform the only wavelet coefficient-based methods for both melanoma and seborrhoeic keratosis detection. The performance results show that the modelling of skin image data in conjunction with wavelet data improves the accuracy of lesion detection. Furthermore, the higher performance of the fused methods than the other single models shows that the proposed sum of probabilities method efficiently combined the image and coefficient data.

The fusion of skin image and corresponding approximate, horizontal, and vertical coefficients results in higher detection accuracy than other methods [18, 20, 23, 24, 26, 28]. This indicates that horizontal and vertical details of lesions contribute to lesion recognition. The higher detection accuracy also proves that extracting horizontal and vertical features and then modelling these details in conjunction with skin image data improves accuracy.

6 Conclusion

Wavelet and approximate coefficient-based DL models are proposed for malignant melanoma and seborrhoeic keratosis classification. Models are generated for skin images and corresponding coefficients. ResNet-18 and ResNet-50 DL models are used for modelling image and wavelet coefficients. These models are pre-trained on the ImageNet database, and then parameters are adapted to skin and wavelet coefficients. Afterwards, models are fused to obtain the proposed wavelet and approximate coefficient-based methods. The performances of single and fused models are evaluated and compared. The results show that the proposed fused methods outperform the individual image and wavelet-based methods.

A comparison of the proposed methods and other methods has also been performed. Other methods are based on skin images; on the other hand, the proposed methods are based on wavelet coefficients. Obtaining horizontal and vertical details of skin images and then modelling these details in conjunction with skin images leads to higher performance in comparison with image-based methods.

7 References

[1] Esteva, A., Kuprel, B., Novoa, A.R., *et al.*: 'Dermatologist-level classification of skin cancer with deep neural networks', *Nature*, 2017, **547**, (2), pp. 115–118

[2] Han, S.S., Kim, M.S., Lim, W., *et al.*: 'Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm', *J. Invest. Dermatol.*, 2018, **138**, pp. 1529–1538

[3] Tschandl, P., Rosendahl, C., Kittler, H.: 'The HAM10000 dataset: a large collection of multi-source dermatoscopic images of common pigmented skin lesions', ArXiv, 2018

[4] Nehal, K.S., Bichakjian, C.K.: 'Update on keratinocyte carcinomas', *N. Engl. J. Med.*, 2018, **379**, (4), pp. 363–374

[5] Oliveira, R.B., Papa, J.P., Pereira, A.S., *et al.*: 'Computational methods for pigmented skin lesion classification in images: review and future trends', *Neural Comput. Appl.*, 2015, **29**, (3), pp. 613–636

[6] Geller, A.C., Dickerman, B.A., Taber, J.M., *et al.*: 'Skin cancer interventions across the cancer control continuum: a review of experimental evidence (1/1/2000–6/30/2015) and future research directions', *Preventive Med.*, 2018, **111**, pp. 442–450

[7] Li, Y., Shen, L.: 'Dermoscopy image analysis: overview and future directions', *IEEE J. Biomed. Health Inf.*, 2019, **23**, (2), pp. 474–478

[8] Krizhevsky, A., Sutskever, I., Hinton, G.E.: 'ImageNet classification with deep convolutional neural networks'. Proc. 25th Int. Conf. on Neural Information Processing Systems – Vol. 1 (NIPS'12), USA, 2012, pp. 1097–1105

[9] Szegedy, C., Liu, W., Jia, Y.: 'Going deeper with convolutions'. IEEE Conf. on Computer Vision and Pattern, Boston, MA, USA, 2015

[10] Szegedy, C., Vanhoucke, V., Ioffe, S., *et al.*: 'Rethinking the inception architecture for computer vision'. IEEE Conf. on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2016

[11] Simonyan, K., Zisserman, A.: 'Very deep convolutional networks for large-scale image recognition'. Int. Conf. on Learning Representations, San Diego, CA, USA, 2015

[12] Targ, S., Almeida, D., Lyman, K.: 'ResNet in ResNet: generalizing residual architectures'. CoRR, San Juan, Puerto Rico, 2016

[13] Huang, G., Liu, Z., Weinberger, K.Q.: 'Densely connected convolutional networks'. CoRR, Honolulu, HI, USA, 2016

[14] Deng, J., Dong, W., Socher, R., *et al.*: 'ImageNet: a large-scale hierarchical image database'. IEEE Computer Vision and Pattern Recognition, Miami, FL, USA, 2009

[15] Kaymak, S., Serener, A.: 'Automated age-related macular degeneration and diabetic macular edema detection on OCT images using deep learning'. IEEE Int. Conf. on Intelligent Computer, Cluj-Napoca, Romania, 2018

[16] Serte, S., Serener, A.: 'Dry and wet age-related macular degeneration classification using OCT images and deep learning'. Int. Conf. on Electrical–Electronics, Biomedical Engineering, Computer Science, Istanbul, Turkey, 2019

[17] Kaymak, S., Serener, A.: 'Deep learning for two-step classification of malignant pigmented skin lesions'. Symp. on Neural Networks and Applications, Belgrade, Serbia, 2018

[18] Matsunaga, K., Hamada, A., Minagawa, A., *et al.*: 'Image classification of melanoma, nevus and seborrheic keratosis by deep neural network ensemble'. ArXiv, 2017

[19] Harangi, B.: 'Skin lesion classification with ensembles of deep convolutional neural networks', *Biomed. Inf.*, 2018, **86**, pp. 25–32

[20] Mahbod, A., Schaefer, G., Ellinger, I., *et al.*: 'Fusing fine-tuned deep features for skin lesion classification', *Comput. Med. Imag. Graphics*, 2019, **71**, pp. 19–29

[21] Nyiri, T., Kiss, A.: 'Novel ensembling methods for dermatological image classification'. Int. Conf. on Theory and Practice of Natural Computing, Dublin, Ireland, 2018

[22] Nozdryn-Plotnicki, A., Yap, J., Yolland, W.: 'Ensembling convolutional neural networks for skin cancer classification'. ArXiv, 2018

[23] Li, Y., Shen, L.: 'Skin lesion analysis towards melanoma detection using deep learning network', *Sensors*, 2018, **18**, (2), p. 556

[24] Diaz, I.G.: 'Incorporating the knowledge of dermatologists to convolutional neural networks for the diagnosis of skin lesions', 2017

[25] Bi, L., Kim, J., Ahn, E., *et al.*: 'Automatic skin lesion analysis using large-scale dermoscopy images and deep residual networks', 2018

[26] Hagerty, J., Stanley, R., Almubarak, H., *et al.*: 'Deep learning and handcrafted method fusion: higher diagnostic accuracy for melanoma dermoscopy images', 2019

[27] Menegola, A., Tavares, J., Fornaciali, M.: 'RECOD Titans at ISIC Challenge 2017'. Available at https://arxiv.org/pdf/1703.04819

[28] Zhang, J., Xie, Y., Xia, Y.: 'Attention residual learning for skin lesion classification', 2019

[29] Fujisawa, Y., Otomo, Y., Ogata, Y., *et al.*: 'Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board- certified dermatologists in skin tumour diagnosis', *Br. J. Dermatol.*, 2019, **180**, pp. 373–381

[30] Gonzalez, R.C., Woods, R.E.: '*Digital image processing*' (Prentice Hall, Upper Saddle River, NJ, 2008)

[31] Serte, S., Demirel, H.: 'Gabor wavelet-based deep learning for skin lesion classification', *Comput. Biol. Med.*, 2019, **113**, p. 103423