Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 197 (2022) 198–207

Procedia Computer Science

www.elsevier.com/locate/procedia

Sixth Information Systems International Conference (ISICO 2021)

Melanoma image classification based on MobileNetV2 network

Rarasmaya Indraswari[a,*], Rika Rokhana[b], Wiwiet Herulambang[c]

[a]*Department of Information Systems, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia*
[b]*Department of Electrical Engineering, Politeknik Elektronika Negeri Surabaya, Indonesia*
[c]*Department of Informatics, Universitas Bhayangkara, Surabaya, Indonesia*

Abstract

Melanoma is one of the most common types of cancer that can lead to high mortality rates if not detected early. Recent studies about deep learning methods show promising results in the development of computer-aided diagnosis for accurate disease detection. Therefore, in this research, we propose a method for classifying melanoma images into benign and malignant classes by using deep learning model and transfer learning. MobileNetV2 network is used as the base model because it has lightweight network architecture. Therefore, the proposed system is promising to be implemented further on mobile devices. Moreover, experimental results on several melanoma datasets show that the proposed method can give high accuracy, up to 85%, compared with other networks. Furthermore, the proposed architecture of the head model, which uses a global average pooling layer followed by two fully-connected layers, gives high accuracy while maintaining the network's efficiency.

Keywords: Computer vision; deep learning; transfer learning; melanoma; MobileNetV2

1. Introduction

Skin cancer is one of the most common types of cancer in the world, especially in white populations [1]. Melanoma, which is due to the rapid growth of melanocytes, is one of the most common types of skin cancer [2]. Melanocytes are cells that produce melanin, which gives colour to skin, hair, and eyes [3]. Most of the common moles are benign melanoma [4]. However, some moles turn out to be malignant melanoma. Surgical excision is the primary treatment

* Corresponding author.
E-mail address: raras@its.ac.id

1877-0509
10.1016/j.procs.2021.12.132

for malignant melanoma [5]. It is effective if the lesion is detected early before it spread to other organs. However, if the lesion already spreads to other organs, it will be more difficult to treat, leading to high mortality rates [6]. Therefore, early detection of malignant melanoma lesions is important to increase the chance of recovery. The primary method for dermatologists to evaluate melanoma's prognosis and severity is by measuring its depth using biopsy [7].

Many computer vision methods have been developed to assist the medical field in detecting diseases easily, quickly, cheaply, and accurately. Computer vision aims to make computers be able to see like humans, including human ability to distinguish objects based on their images or videos. Applications based on computer vision are often used in the medical field because they can be used for non-invasive diagnostics. There were several recent studies show that computer vision methods can be used to detect malignant melanoma automatically using dermoscopic images (digital images obtained by epiluminescence microscopy) [2], [3], [7]–[11]. The methods vary from the conventional machine learning and computer vision approach [7]–[9], which rely on the image segmentation process and manual feature determination, to deep learning with transfer learning approach [2], [3], [10], [11].

Sonia (2016) uses energy features extracted using Non-Sub sampled Contourlet Transform (NSCT) to classify 120 dermoscopic images into benign and malignant classes [9]. This method uses Bayes classification and gives 96.7% accuracy, 97.5% sensitivity, and 96.3% specificity. Pérez-Ortiz, et al. (2016) use 86 features representing the shape, colour, pigment network, and texture of melanoma lesion to classify it into 5 classes: 1 benign class and 4 different stages of malignant melanoma lesion [7]. This study proposes an ordinal cascade classification model for multi-class segmentation using Support Vector Machine (SVM), combined with the oversampling method for the pre-processing to solve the imbalance dataset problem. This study uses 556 dermoscopy images with 313 images in the benign class and gives 58.36% classification accuracy. Indraswari, et al. (2017) use the information of the size, shape, and colour of the lesion to classify it into benign and malignant classes using SVM [8]. Experiment on 60 dermoscopy images gives 83.3% accuracy, 80% sensitivity, and 86.7% specificity.

The most recent studies usually utilize deep learning methods for classifying melanoma lesions because they can provide accurate results with enough training data. Deep learning does not rely on the image segmentation process and manual feature determination, reducing the possibility of stacked errors from the previous processes [10]. Moreover, research from Brinker, et al. (2019) shows that deep learning method, such as Convolutional Neural Network (CNN), is promising to be used as computer-aided diagnosis tools because it can outperform 136 of 157 dermatologists in classifying melanoma images [12]. Lopez, et al. (2017) modify the fully-connected layer of VGG16 network so that it can be used for binary classification of melanoma images in the International Symposium on Biomedical Imaging (ISBI) 2016 dataset [11][13]. Using transfer learning with the VGG16 architecture as the base model, the proposed method achieved 68.67% accuracy, 33.11% sensitivity, and 49.58% precision using 379 test images. Kassani & Kassani (2019) make a comparative study on several deep learning architectures for the base model of the transfer learning process in classifying melanoma images in the International Skin Imaging Collaboration (ISIC) 2018 dataset [3]. ResNet50 achieved the best performance as the base model with 93.73% precision, 92.53% recall, 92.74% F-score, and 92.08% accuracy. Rokhana, et al. (2020) propose a CNN architecture for classifying melanoma images in ISIC-Archive dataset [10][14]. The proposed network architecture consists of 6 convolutional layers and 3 max-pooling layers, which gives 84.67% accuracy, 91.97% sensitivity, and 78.71% specificity.

Currently, deep learning methods continue to develop. One of the newest deep learning methods that is widely used today is MobileNetV2 because it has one of the most lightweight network architectures [15]. MobileNetV2 has a novel layer module, the inverted residual with linear bottleneck, that significantly reduces the memory needed for processing [16]. Therefore, to provide an accurate and efficient classification of melanoma lesions, in this research we propose a melanoma image classification based on MobileNetV2 network. The proposed method uses MobileNetV2 as the base model for the transfer learning process and adds a global pooling layer followed by two fully-connected layers as the head model. By using a lightweight network architecture, the developed model can be implemented further on mobile devices, making it easier to detect melanoma. Therefore, it is hoped that this disease can be detected as early as possible to increase people's life expectancy.

2. Methodology

2.1. Dataset

In this research, we use 4 datasets containing melanoma images, which are divided into 2 classes: benign and malignant. The first dataset is from ISIC-Archive repository [14], the second dataset is from ISBI 2016 [13], the third dataset is from MED-Node [17], and the fourth dataset is from PH^2 Database [18]. The number of images contained in each dataset is shown in Table 1. All of the dataset contains dermoscopic images, except for MED-Node dataset which contains non-dermoscopic images. MED-Node dataset uses simple digital images taken with a Nikon D3 or Nikon D1x body and a Nikkor 2.8/105 mm micro lens, which are much easier to obtain [17]. Moreover, the class in the MED-Node dataset is melanoma (which equals to malignant melanoma) and naevus (which equals to benign melanoma). The PH^2 dataset actually consists of 200 images. However, there are only 180 images with label, which is melanoma (which equals to malignant melanoma) and atypical nevus (which equals to benign melanoma). The example of benign and melanoma lesion is shown in Fig. 1 and Fig. 2, respectively.

ISIC-Archive and ISBI 2016 datasets were already separated into training and testing set. For the experimental results, we divide the images in MED-Node and PH^2 dataset stratified-randomly based on its label with the ratio of training data : testing data = 70 : 30. From Table 1, it can be seen that ISBI 2016 dataset has a high imbalance between benign and malignant classes. PH^2 dataset also has a fairly noticeable data imbalance. The class of the data was binarized into 0 (negative class) for the benign class and 1 (positive class) for the malignant class.

Table 1. Dataset.

Dataset	Total data (images)		Training set (images)		Testing set (images)	
	Benign	Malignant	Benign	Malignant	Benign	Malignant
ISIC-Archive [14]	1,800	1,497	1,440	1,197	360	300
ISBI 2016 [13]	1,031	248	727	173	304	75
MED-Node [17]	100	70	70	49	30	21
PH^2 Database [18]	80	40	56	28	24	12

(a) (b) (c) (d)

Fig. 1. Example of benign lesion from (a) ISIC-Archive [14]; (b) ISBI 2016 [13]; (c) MED-Node [17]; and (d) PH^2 Database [18].

(a) (b) (c) (d)

Fig. 2. Example of malignant lesion from (a) ISIC-Archive [14]; (b) ISBI 2016 [13]; (c) MED-Node [17]; and (d) PH^2 Database [18].

2.2. Transfer learning

Transfer learning is a machine learning method that learns to finish a new task by transferring knowledge from a model used to finish other tasks. This method uses a pre-trained model on a large dataset as the base model that will be used to complete other tasks that use other datasets. Transfer learning is widely used because this method allows to perform classification accurately with a small dataset. This is because a deep learning model that was trained from scratch using a small dataset is very difficult to obtain high accuracy because the model does not get enough information regarding the data variations. Therefore, it cannot learn the important features in the data. In transfer learning, the pre-trained model obtains the knowledge from the large dataset, which is represented as the weights in the network. Those weights are then used in another network for a different task with a different dataset. Therefore, instead of training the second network from scratch, usually with a small dataset, we "transfer" the learned features of the first network to the second network.

In computer vision, the first layers of a deep convolutional network are usually used for learning the general features in images, such as lines or shapes. The network learns about the specific features for a particular task, for example image classification, in the last layers of the network. Therefore, in transfer learning, the weights of the base network or model are usually not changed (frozen). Then, the learning process is done to the last layer of the network, which is generally a fully-connected layer, so that the network can learn the specific features to classify the new dataset. Several studies also propose adding several layers on top of the base model, which we called the head model, to further improve the network's performance.

2.3. MobileNetV2

The proposed method uses MobileNetV2 [16] as the base model. MobileNetV2 was developed from MobileNetV1 [19] by adding inverted residual with linear bottleneck modules. MobileNet architecture was based on depthwise separable convolution. The standard 2D convolution processes all of the input channels directly to produce one output channel by convolving in the depth dimension (channel) as well. The depthwise convolution separates the input image and the filter into different channels and then convolves each input channel with the corresponding filter channel. After the filtered output channel have been produced, those output channels are then stacked back. In separable depthwise convolution, the stacked output channels are then filtered using a 1×1 convolution, called as pointwise convolution, to combine the stacked output channels into one channel. The depthwise separable convolution produces the same output as the standard convolution, but it is more efficient because it reduces the number of parameters involved in the process [19]. By counting the depthwise and pointwise convolution as separated layers, MobileNetV1 has 28 convolutional layers that produce output with size $7\times7\times1280$ pixels.

Both MobileNetV1 and MobileNetV2 take input images with the size of $224\times224\times3$ pixels. Therefore, the input images on the dataset are resized and cropped into 224×224 pixels. MobileNetV2 insert 19 inverted residual bottleneck layers after the first convolution layer with 32 filters and then ended with a pointwise convolution that produces output with size $7\times7\times1280$ pixels [16]. Residual block connects the beginning and end of a convolutional block with a skip-connection with the purpose of conveying information to the deeper layer of the network. In the standard residual block, the beginning and end of the convolutional block usually have more channels than the layers between. However, in the inverted residual block as used in MobileNetV2, the connected layers have fewer channels than the layers between, which leads to far fewer parameters than the standard residual block. Therefore, the difference between MobileNetV1 and MobileNetV2 layers can be seen in Fig. 3. There are several convolutional blocks with a skip-connection on MobileNetV2, while the skip-connection does not exist in MobileNetV1.

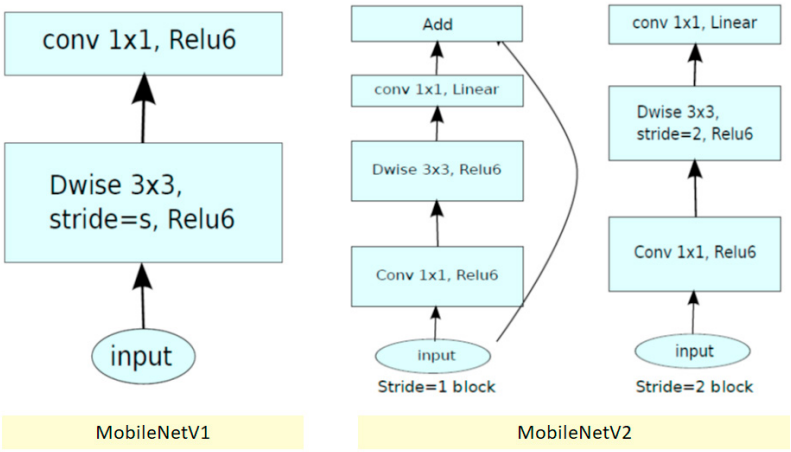


Fig. 3. The difference between MobileNetV1 and MobileNetV2 layers [16].

2.4. Proposed method

The methodology of the proposed method is shown in Fig. 4. In this research, the base model, which is MobileNetV2 network, was pre-trained using ImageNet dataset that consisted of about 1,000 object classes, 1,281,167 training images, 50,000 validation images, and 100,000 test images [20]. We add several layers, which was called as the head model, on top of the convolutional layers of the MobileNetV2. The first layer of the head model is a global pooling layer that takes the 7×7×1280 pixels feature map, which is the output of the base model, as the input. The global pooling layer is a pooling operation that generates a one-dimensional feature vector, reducing the dimension of the data significantly. The proposed method uses an average pooling operation with kernel size 7×7 in the global pooling layer that produces an output feature map with the size of 1×1×1280 pixels. The global pooling layer is then followed by two fully-connected or dense layers. The fully-connected layers consisted of 128 and 64 nodes, respectively, with ReLU (Rectified Linear Unit) activation function [21]. Next is the output layer that consists of 2 nodes because we use one-hot encoding, and there are 2 classes in the dataset, which are benign and malignant. The output layer uses a softmax activation function.

The base model uses the pre-trained weights, and the layers are being frozen. Therefore, in the training process, only the weights of the head model were trained. The optimization algorithm used in the training process is Adam's optimizer [22] because it can reach convergence quickly [23]. Therefore, we use a relatively small number of epochs, which is 20. The number of the image in every batch is 32 images. Binary cross-entropy was utilized as the cost function that the network needs to minimalize during the training process.

To improve the generalization ability of the model, we do data augmentation for the training set. Data augmentation transforms the given data (in this case is the images on the training set) based on one or more of the operations that have been determined. In this study, the operations used for the data augmentation process are including rotation up to 30°, flip, zoom, and shift. The augmented data is then used as the input of the network in each iteration of the training process so that the network does not see the same data in every iteration. Therefore, the amount of data that becomes the network's input in each iteration of the training process is always the same. However, the data used in each iteration process will vary, and the resulting model will have better generalization ability. By having good generalization ability, the model will be able to deal with testing data that has never been learned before.

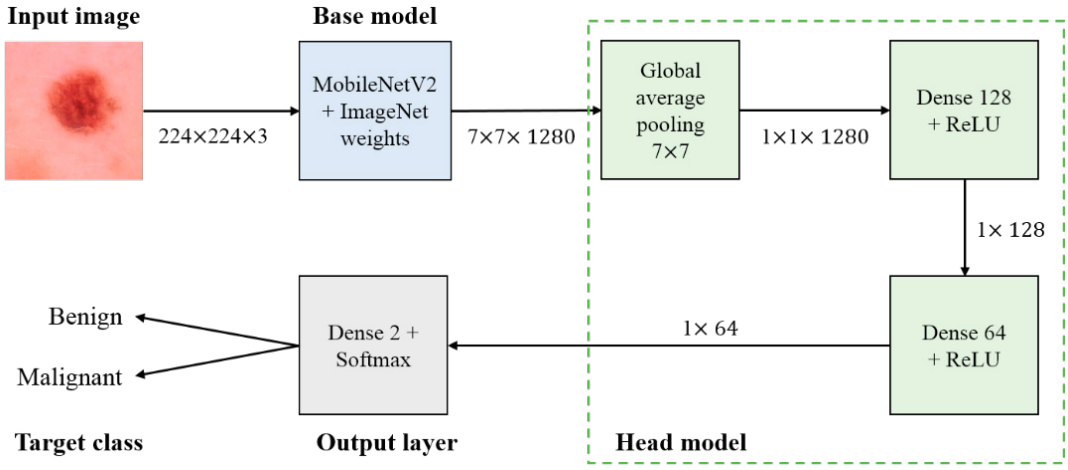


Fig. 4. Proposed method.

3. Results and discussion

This research was implemented by using Python programming language and Keras library. The experiments were run using GPU NVIDIA GeForce GTX 860M, RAM 2 × 8GB 2400 MHz DDR4. The performance of the proposed method was evaluated by using accuracy, sensitivity, specificity, and precision evaluation metric. Accuracy measures the number of data in the testing set that was correctly classified. Sensitivity measures the number of malignant lesions that were correctly classified from all of the malignant data in the testing set. Specificity measures the number of benign lesions that were correctly classified from all of the benign data in the testing set. Precision measures the number of the correctly classified malignant lesion from all of the data that were classified as malignant lesions by the system. Moreover, we also record the running time of the models so that we compare the efficiency of each model.

3.1. Performance of the proposed method

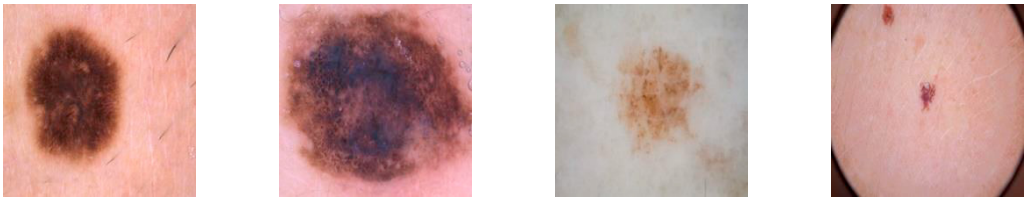
The proposed method was tested using 4 datasets, as shown in Table 2. The proposed method can achieve up to 85% accuracy on ISIC-Archive dataset, which contains the largest number of training images compared with other datasets. However, the proposed method gives low sensitivity to the ISBI 2016 and PH2 datasets, showing that the proposed method failed to classify the malignant class correctly. This is because those two datasets have a high imbalance, in which the number of malignant data is far smaller than the number of benign data. Therefore, further processing, such as using an over-sampling algorithm or using a weighted cost function [24], needs to be done to address the imbalance dataset problem.

Table 2. Performance of the proposed method.

Dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Running time (s)
ISIC-Archive [14]	85	85	85	83	2,394
ISBI 2016 [13]	83	36	95	64	724
MED-Node [17]	75	76	73	67	97
PH^2 Database [18]	72	33	92	67	68

Fig. 5 presents several examples of the classification result on the testing set of ISIC-Archive dataset. We display images that are potentially misclassified by the proposed method because their characteristics are similar to those of other classes. However, the proposed method has succeeded in classifying the image in Fig. 5(a) and Fig. 5(c). The proposed method failed to classify the benign data in Fig. 5(b), which is categorized as a false positive case. The

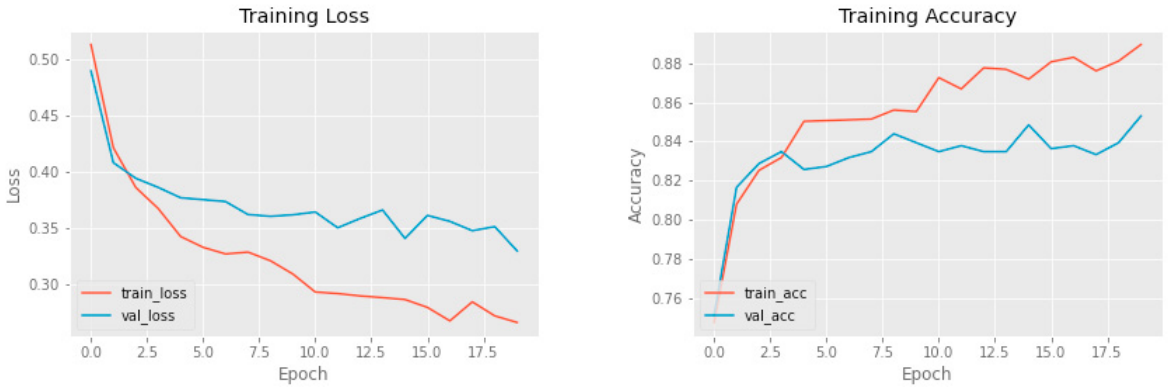
proposed method failed to classify the malignant data in Fig. 5(d), which is categorized as a false negative case. The true positive and false negative cases affect the sensitivity value, while the true negative and false positive cases affect the specificity value.



	(a)	(b)	(c)	(d)
Test image				
Class	Benign	Benign	Malignant	Malignant
Output	Benign	Malignant	Malignant	Benign
Status	True negative	False positive	True positive	False negative

Fig. 5. Example of the classification results on the ISIC-Archive testing set.

The proposed method did not experience an overfitting problem. Overfitting is a condition in which the resulting model fits too well with the training data and does not have good generalization ability, thus provides low accuracy when used on new data. Overfitting checking can be done by conducting a validation process at each iteration of the training process of the deep learning model. In the validation process, the network with the newly updated weights is tested on validation data that is not contained in the training data. Loss and accuracy values for training data and validation data in each iteration of the training process can be graphed as shown in Fig. 6. In this study, the validation data that we use are data from the testing set. Fig. 6 presents the graph of the loss and accuracy values during the training process of the proposed method on the ISIC-Archive dataset. Overfitting can be detected by looking at this graph. The characteristic of an overfitting model is that the resulting model has good performance in classifying the training data but has poor performance in classifying new data, such as validation data. Therefore, an overfit model usually experiences a decrease in the value of the training loss but experiences an increase in the value of the validation loss as the training process iteration goes on. It can be seen in Fig. 6(a) that both of the training loss and validation loss continue to decrease over iteration of the training process. Therefore, it can be concluded that the proposed model does not overfit and have good generalization ability. This can happen because the proposed model uses a relatively small epoch value and performs data augmentation to increase the variability of the training data.



(a) (b)

Fig. 6. Performance of the training process of the proposed method on ISIC-Archive dataset: (a) training loss; (b) training accuracy.

3.2. Comparison of the base model

In this experiment, we change the base model of the proposed method into ResNet50V2 [25], InceptionV3 [26], and InceptionResNetV2 [27], and then measure their performance on ISIC-Archive dataset. The networks were also pre-trained using ImageNet dataset and added with the same head model architecture. We compared the proposed method with ResNet50V2 because in [3], it was said that ResNet50 gives the best performance in classifying melanoma images compared with AlexNet, VGGNet16, VGGNet19, and Xception network. As shown in Table 3, the base model of the proposed method, which is MobileNetV2 network, gives the highest accuracy compared with other networks. Moreover, the proposed method also gives the fastest running time, which confirms that MobileNetV2 is one of the most lightweight network architectures [15]. ResNet50V2 gives the highest sensitivity, which means that using this architecture as the base model can result in accurate detection of malignant class. However, ResNet50V2 model gives the lowest specificity value, which means that this architecture is not as good as other architectures in detecting the malignant class. The result of ResNet50V2 is in contrast to the result of InceptionResNetV2, which gives the highest specificity value but has the lowest sensitivity value.

Table 3. Comparison of the base model.

Base model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Running time (s)
MobileNetV2 [16] **(Proposed)**	**85**	85	85	83	**2,394**
ResNet50V2 [25]	84	**87**	81	79	6,105
InceptionV3 [26]	81	79	82	78	3,844
InceptionResNetV2 [27]	78	63	**91**	**85**	7,899

3.3. Experiment on the head model

We experiment on different architectures of the head model using ISIC-Archive dataset, as shown in Table 4. First, we experiment on using max-pooling layer instead of the proposed average pooling layer. However, the proposed average operation gives better performance than the maximum operation. This was supported by the many uses of the average operation for the global pooling layer, as in MobileNetV2 original network [16]. We then utilized the head model proposed in [3], which directly uses 2 fully-convolutional layers after the base model. However, our proposed method outperformed this approach. Last, we hypothesized that adding a trainable convolution layer on top of the base model can increase the method's performance. This was proven because this architecture achieved 86% accuracy. However, this approach also gives the longest running time, which indicates that we need to trade off the accuracy with the network's efficiency. Because the accuracy of the proposed method is not too different from the accuracy of using an additional convolution layer, we decided that the proposed method has better performance because it can provide high accuracy while maintaining the efficiency of the network.

Table 4. Experiment on the head model.

Architecture of the head model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Running time (s)
Average pooling layer 7×7 – Dense 128 – Dense 64 **(Proposed)**	85	**85**	85	83	2,394
Max-pooling layer 7×7 – Dense 128 – Dense 64	82	78	85	81	**2,125**
Dense 128 – Dense 64 [1]	83	79	**87**	**84**	2,469
Convolution 3×3×320 – Average pooling layer 7×7 – Dense 128 – Dense 64	**86**	84	**87**	**84**	2,915

3.4. Comparison with the previous study

There were other studies that use the same dataset used in this research and also use the deep learning method, as shown in Table 5. The proposed method gives higher accuracy than [10] that trains their convolutional neural network from scratch. However, the difference is not too prominent. It may happen because the number of training sets in the ISIC-Archive dataset is quite large, so that it can provide good accuracy results when the training process is carried out from scratch. Lopez, et al. [11] uses 3 training approaches for processing ISBI 2016 dataset, which are: 1) training from scratch; 2) transfer learning using VGG16; and 3) transfer learning with fine-tuning which is using pre-trained weights from ImageNet for the first 4 convolutional layers and using pre-trained weight from the training from scratch process for the remaining convolutional layers. Overall, our proposed method can give highest accuracy. However, since ISBI 2016 is an imbalanced dataset, our proposed method did not give accurate results in classifying the malignant class. The training from scratch and fine-tuning approach can give better results in detecting the malignant class. However, they must have a lower specificity value since their accuracy is lower than the proposed method.

Table 5. Comparison with the previous study.

Dataset	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
ISIC-Archive [14]	Proposed Method	**85.00**	85.00	**85.00**	83.00
	Rokhana, et al. [10]	84.76	**91.97**	78.71	-
ISBI 2016 [13]	Proposed Method	**83.00**	36.00	95.00	64.00
	Lopez, et al. [11] – 1) Training from scratch	66.00	57.99	-	67.77
	Lopez, et al. [11] – 2) Transfer learning	68.67	33.11	-	49.58
	Lopez, et al. [11] – 3) Transfer learning & fine-tuning	81.33	**78.66**	-	**79.74**

4. Conclusion

Melanoma is one of the most common types of cancer in the world that has high mortality rates if not detected early. In this research, we propose a method for classifying melanoma images into benign and malignant classes by using a deep learning model and transfer learning. The proposed method consists of a pre-trained frozen base model and a trainable head model. MobileNetV2 network is used as the base model because it has lightweight network architecture. Comparison of base model architecture shows that the MobileNetV2 gives the highest accuracy and efficiency compared with ResNet50V2, InceptionV3, and InceptionResNetV2. The proposed architecture of the head model, which uses a global average pooling layer followed by two fully-connected layers, gives high accuracy while maintaining the network's efficiency. Overall, experimental results on several melanoma datasets show that the proposed method can give high accuracy, which is up to 85% on ISIC-Archive dataset. However, further improvement is needed so that the proposed method can address the imbalanced dataset problem. Moreover, the proposed system is promising to be implemented further on mobile devices for assisting early detection of melanoma.

References

[1] Narayanan, D. L., R. N. Saladi, and J. L. Fox. (2010) "Ultraviolet radiation and skin cancer." *International Journal of Dermatology* 49 (9): 978–986, doi: 10.1111/j.1365-4632.2010.04474.x.

[2] Sultana, N. N., and N. B. Puhan. (2018) "Recent deep learning methods for melanoma detection: A review." *Communications in Computer and Information Science* 834 (April): 118–132, doi: 10.1007/978-981-13-0023-3_12.

[3] Kassani, S. Hosseinzadeh, and P. Hosseinzadeh Kassani. (2019) "A comparative study of deep learning architectures on melanoma detection." *Tissue Cell* 58 (March): 76–83, doi: 10.1016/j.tice.2019.04.009.

[4] Ghadially, F. N. (1996) "Trauma and melanoma production." *Nature* 211 (5054): 1199-1199.

[5] Reilly, F. O. (2010) "Melanoma in solid organ transplant recipients." *American Journal of Transplantation* 10 (5): 1297–1304, doi: 10.1111/j.1600-6143.2010.03078.x.

[6] Sandru, A., S. Voinea, E. Panaitescu, and A. Blidaru. (2014) "Survival rates of patients with metastatic malignant melanoma." *Journal of medicine and life* 7 (4): 572–576.

[7] Pérez-Ortiz, M., A. Sáez, J. Sánchez-Monedero, P. A. Gutiérrez, and C. Hervás-Martínez. (2016) "Tackling the ordinal and imbalance nature of a melanoma image classification problem." *Proceedings of the International Joint Conference on Neural Networks* 2016 (October): 2156–2163, doi: 10.1109/IJCNN.2016.7727466.

[8] Indraswari, R., W. Herulambang, R. Rokhana, and U. B. Surabaya. (2017) "Melanoma classification using automatic region growing for image segmentation." in *ICTA 2017 UBHARA Surabaya*: 165–172.

[9] Sonia, R. (2016) "Melanoma image classification system by NSCT features and Bayes classification." *International Journal of Advances in Signal and Image Sciences* 2 (2): 27–33.

[10] Rokhana, R., W. Herulambang, and R. Indraswari. (2020) "Deep convolutional neural network for melanoma image classification." in *2020 International Electronics Symposium (IES)*: 481–486, doi: 10.1109/IES50839.2020.9231676.

[11] Lopez, A. R., X. Giro-i-nieto, J. Burdick, and O. Marques. (2017) "Skin lesion classification from dermoscopic images using deep learning techniques." in 13th IASTED international conference on biomedical engineering (BioMed): 49–54.

[12] Brinker, T. J., Achim Hekler, Alexander H. Enk, Joachim Klode, Axel Hauschild, Carola Berking, Bastian Schilling, S. Haferkamp, D. Schadendorf, T. Holland-Letz, J. S. Utikal. (2019) "Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task." *European Journal of Cancer* 113: 47–54, doi: 10.1016/j.ejca.2019.04.001.

[13] Gutman, D., Noel CF Codella, Emre Celebi, Brian Helba, Michael Marchetti, Nabin Mishra, and Allan Halpern. (2016) "Skin lesion analysis toward melanoma detection : A challenge at the International Symposium on Biomedical Imaging (ISBI) 2016 , hosted by the International Skin Imaging Collaboration (ISIC)." *arXiv Prepr.* arXiv1605.01397.

[14] Fanconi, C. (2019) "Skin cancer: Malignant vs. benign - processed skin cancer pictures of the ISIC archive." [Online]. Available: https://www.kaggle.com/fanconic/skin-cancer-malignant-vs-benign.

[15] Eyiokur, F. I., H. K. Ekenel, and A. Waibel. (2021) "A computer vision system to help prevent the transmission of COVID-19." *arXiv Prepr.* arXiv2103.08773 [Online]. Available: http://arxiv.org/abs/2103.08773.

[16] Sandler, M., A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen. (2018) "MobileNetV2: Inverted residuals and linear bottlenecks." *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*: 4510–4520, doi: 10.1109/CVPR.2018.00474.

[17] Giotis, I., N. Molders, S. Land, M. Biehl, M. F. Jonkman, and N. Petkov. (2015) "MED-NODE : A computer-assisted melanoma diagnosis system using non-dermoscopic images." *Expert Systems With Applications* 42 (19): 6578–6585, doi: 10.1016/j.eswa.2015.04.034.

[18] Mendoncÿa, T., P. M. Ferreira, J. Marques, A. R. S. Marcÿal, and J. Rozeira. (2013) "A dermoscopic image database for research and benchmarking." [Online]. Available: https://www.fc.up.pt/addi/ph2 database.html.

[19] Howard, A. G., Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. (2017) "MobileNets: Efficient convolutional neural networks for mobile vision applications." *arXiv Prepr.* arXiv1704.04861 [Online]. Available: http://arxiv.org/abs/1704.04861.

[20] Deng, J., W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei. (2009) "Imagenet: A large-scale hierarchical image database." in *2009 IEEE conference on computer vision and pattern recognition*: 248–255.

[21] Agarap, A. F. (2018) "Deep learning using Rectified Linear Units (ReLU)." *arXiv Prepr.* arXiv1803.08375 [Online]. Available: http://arxiv.org/abs/1803.08375.

[22] Kingma, D. P., and J. L. Ba. (2015) "Adam: A method for stochastic optimization." *3rd International Conference on Learning Representations,* ICLR 2015 - Conference Track Proceedings: 1–15.

[23] Ruder, S. (2016) "An overview of gradient descent optimization algorithms." *arXiv Prepr.* arXiv1609.04747 [Online]. Available: http://arxiv.org/abs/1609.04747.

[24] Indraswari, R., T. Kurita, A. Z. Arifin, N. Suciati, and E. R. Astuti. (2019) "Multi-projection deep learning network for segmentation of 3D medical images." *Pattern Recognition Letters* 125: 791–797, doi: 10.1016/j.patrec.2019.08.003.

[25] He, K., X. Zhang, S. Ren, and J. Sun. (2016) "Identity mappings in deep residual networks." *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9908 LNCS: 630–645, doi: 10.1007/978-3-319-46493-0_38.

[26] Szegedy, C., V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. (2016) "Rethinking the Inception architecture for computer vision." *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2016 (December): 2818–2826, doi: 10.1109/CVPR.2016.308.

[27] Szegedy, C., S. Ioffe, V. Vanhoucke, and A. A. Alemi. (2017) "Inception-v4, Inception-ResNet and the impact of residual connections on learning." *31st AAAI Conference on Artificial Intelligence*, AAAI 2017: 4278–4284.