INTERNATIONAL JOURNAL OF FOUNDATIONS OF COMPUTER SCIENCE

Published by World Scientific Publishing Co. Pte Ltd
Online ISSN: 01290541
Volume 32, Issue 05 (2021)

ENSEMBLE OF IMPROVED EFFICIENTNET MODEL FOR MELANOMA DETECTION IN DERMOSCOPIC IMAGES

Hemant Kumar[1,*]
Email: hemantime@gmail.com, hemantkumar@csjmu.ac.in
ORCID ID: 0000-0003-0603-4394
Department of Information Technology,
Chhatrapati Shahu Ji Maharaj University, Kanpur, India.

Shivneet Tripathi[2]
Email: shivneet@csjmu.ac.in
Department of Computer Applications,
Chhatrapati Shahu Ji Maharaj University, Kanpur, India.

Amit Virmani[3]
Email: amitvirmani@csjmu.ac.in
Department of Computer Applications,
Chhatrapati Shahu Ji Maharaj University, Kanpur, India.

Rashi Agrawal[4]
Email: rashi@csjmu.ac.in
Department of Information Technology,
Chhatrapati Shahu Ji Maharaj University, Kanpur, India.

Indrajit Patra[5]
Email: ipmagnetron0@gmail.com
ORCID ID: 0000-0001-6704-2676
National Institute of Technology, Durgapur, India.

Ashish Gupta[6]
Email: roy.gupta@utoronto.ca
School of Continuing Studies,
University of Toronto, Canada.

Sunil Kumar[7]
Email: sunilymca24@gmail.com
Department of Information Technology,
Chhatrapati Shahu Ji Maharaj University, Kanpur, India.

**Corresponding Author: Hemant Kumar (hemantime@gmail.com)*

ABSTRACT

Melanoma has increased in prevalence over the last three decades, and early detection is critical for lowering the mortality rate linked to its type of skin cancer. Thus, access to an automated, dependable system capable of detecting the existence of melanoma using a dermatoscopic image of lesions can be an essential instrument in the field of clinical diagnostics. Among the state-of-the-art technologies for automated or computerized clinical diagnostics, deep learning based on Convolutional Neural Networks should given attention, which have been utilized to create classification and detection systems for various disorders. The method suggested in this article utilizes an imaging stage to generate the features of the dermatoscopic picture using the EfficientNet-based Convolutional Neural Network, followed by an attention mechanism to reweight the features. Following that, we utilize the dense layer to process metadata and build a fully connected layer that categorizes lesions as "benign" or "malignant". We have trained, validated and tested the suggested model using the database related to the 2020 International Symposium on Biomedical Imaging challenge. The ensemble model's ROC-AUC is 0.9628, while its recall, precision, and F1-score are 0.8639, 0.8746, and 0.8692, respectively. Thus, it shows that the model can discriminate accurately between benign and malignant lesions without bias toward any class.

KEYWORDS

Skin Lesions, Melanoma Detection, EfficientNet, Attention Mechanism.

1. INTRODUCTION

The WHO reports more than 3.1 million individuals worldwide being diagnosed with melanoma in 2015, and there were 60,000 died. Melanoma is the third most prevalent malignancy among men and women in their twenties [1]. The most frequent cancer for men and women in the USA is melanoma. Melanoma is most prevalent in Australia and New Zealand [2][3][4]. According to research, this illness is often aggravated by UV radiation throughout the day, sunbed tanning, and skin tone, among other factors. Early detection, physicians believe, is the safest method of diagnosing any kind of malignant skin disease. When the illness is diagnosed early, the five-year survival rate increases to about 95% [5].

Dermoscopy, also known as Epiluminescent Microscopy (ELM), is a surgical procedure used to evaluate whether a malignant tumor is cancerous or benign. This method uses dermatoscopy, a device that consists of a light source and a magnifying lens. It is used to enhance the perception of clinical patterns like pigmented veils, networks, globs, ramifications, and shades. This method is beneficial in the diagnosis of melanoma [4].

When classifying skin lesions as malignant or benign, the initial step is picture segmentation. Image segmentation occurs when filters are utilized to highlight lesions within images (the area of interest). Next, images are transmitted to dermatologists who examine several lesions from the same patient and look for anomalies representing melanoma [6]. Because the present methodology is not digital and primarily depends on dermatologists, this research aims to utilize the present data set to construct a model with maximum accuracy to predict whether a lesion is malignant or non-malignant [1].

Computer-aided detection (CAD) systems and methods have been advanced since the development of image processing techniques in the classification of Skin Lesions without startling or painful surgical treatment in the early stage of the disease. Currently, convolutional neural networks (CNN) perform very well in a broad range of computer vision, including picture classification, segmentation, and target recognition[7].

Many studies have used CNN to categorize skin lesions, but they continue to face difficulties:

- the high level of intra-class variation and similarity between classes;
- several artifacts, such as hair, color lighting, and so on;
- CNN's limited capacity to generalize;
- imbalanced classes of lesion images.

Challenges of Skin Cancer Detection

Due to the variety of images and sources, there are numerous difficulties with diagnosing skin cancer. First, the type of human color complicates and complicates the diagnosis of skin cancer. The visual characteristics of skin lesion pictures include the following:

- The main challenges in skin cancer are the different sizes and shapes of the images, making accurate identification difficult. According to this point of view, pre-processing is critical for proper analysis in this instance.
- A few wasted signals that were not initially part of a picture but may disrupt to get a desirable outcome must be sacrificed. As a result, all noise and artifacts should be eliminated during the pre-processing processes.
- Low contrast from nearby tissues may also contribute to the difficulty of analyzing skin cancer correctly.
- Due to light beams, color texture, and reflections, color lighting poses extra difficulties.
- Some moles in the human body may never become cancer cells, but they identify skin cancer correctly from malignant images.

- Another problem in skin cancer diagnosis is the present bias, altering the models' performance to reach a better outcome.

The remaining paper is organized as follows: Section 2 shows Review summarises the current state of knowledge about deep learning classification methods for melanoma diagnosis; Section 3 presents proposed methods and the fundamental ideas, an overview of the dataset, and the suggested technique in depth; Section 4, Training and Results illustrates the specifics of the training or training loop, as well as the results of experiments; Section 6, provides context for the conclusion and future scope.

2. RELATED WORKS

Recently, numerous techniques have been proposed for the detection of melanoma from skin images. The summary of the literature review is illustrated in table 1.

Table 1. Related Works of Melanoma Detection.

References	Contribution	Training Techniques	Dataset	Results
[8]	Skin Cancer classification using the pre-trained model on a large dataset.	Transfer Learning	PH2	Accuracy: 92.4%
[9]	Transfer learning was applied on two datasets for the detection of melanoma.	Transfer Learning	ISIC	Accuracy: 91.6%
[10]	Classification of skin lesions utilizing pre-trained transfer learning models.	Transfer Learning	ISIC	Accuracy: 97.5%
[11]	Melanoma classification using Bag-of-Features and CNN.	Bag-Of-Features	unpublished dataset	Accuracy: 95%
[12]	Skin cancer classification using the Gradient Descent Algorithm for CNN.	Gradient Descent Algorithm	ISIC	Accuracy: 88%
[13]	Melanoma recognition using Residual Networks and FCN.	Deep Residual Network	ISIC	Accuracy: 95.7%
[14]	Detection of skin cancer by noise reduction and picture segmentation.	Gradient and feature adaptive contour	PH2	Accuracy: 89.7%
[15]	Skin lesions classification using an ensemble learning model.	Ensemble DNN	ISIC	Accuracy: 84.7%
[16]	Use of inter and intra-network fusion for skin cancer classification.	Fusion + Ensemble DNN	ISIC	Accuracy: 95.1%
[17]	Melanoma diagnosis using VGG16 and RseNet50.	Transfer Learning + crowdsourcing	ISIC	Accuracy: 90.2%
[18]	"You Only Look Once (YOLO)" is based on deep learning CNN	YOLO Algorithm + CNN	ISIC	Accuracy: 89.6%
[19]	Ensemble two CNN architecture and metadata for multiclass classification.	EfficientNet + SENet154	ISIC	Accuracy: 94.8%

3. PROPOSED METHOD

Our model consists of two phases, the first phase is an image phase (including augmentation and feature extraction and attention mechanism) and a metadata phase (including preprocessing, two dense layers). In the image phase, first, we reduce the size of the image according to different EfficientNet models and apply the Progressive Sprinkles augmentation techniques.EfficientNet begins by extracting image features through its convolutional layers. After that, the attention mechanism is used to reweight features, enhancing the activation of essential regions. Next, we passed metadata in two dense layers with swish activation function, dropout, and batch normalization layer and obtained the output (using sigmoid activation function). Subsequently, the classification of pictures is performed using the fused features. The proposed methodology is shown in figure 1. This process performs for all the models (i.e., B0, B1,..., B7) of EfficientNet families and ensembles all the results and obtains the results.

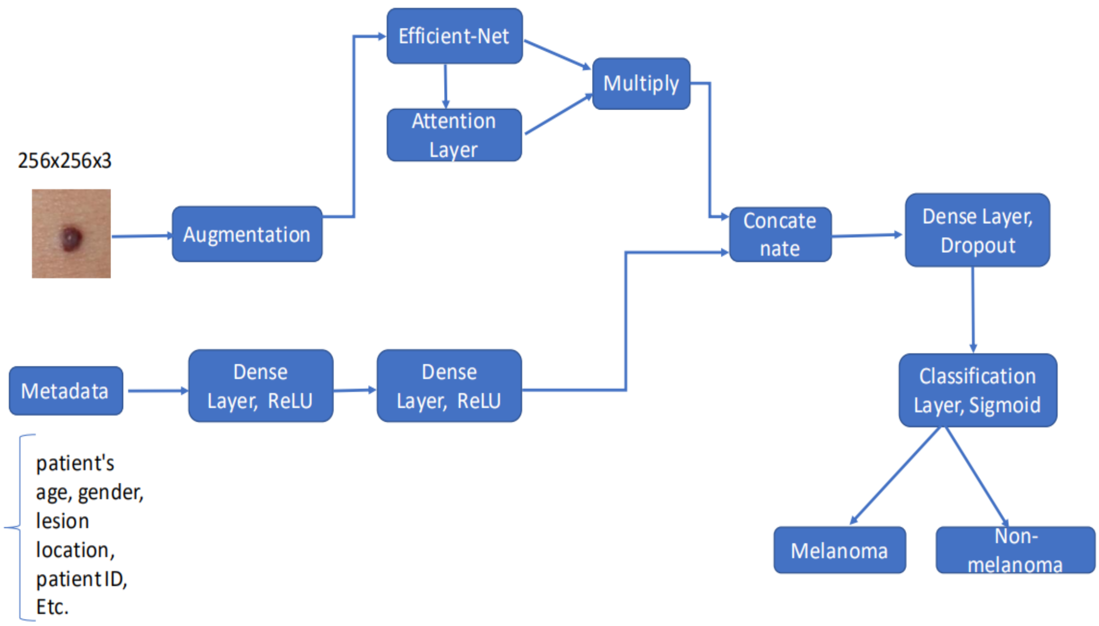

Figure 1. In the image stage, use the EfficientNet, including the Attention Mechanism. In the Metadata stage, use two dense layers, and concatenate them with the output generated from the image stage.

3.1 Dataset

The International Skin Imaging Collaboration (ISIC) collected data as part of a worldwide strategy to improve melanoma diagnostic efficiency, the dataset shown in figure 2. The data comprised information such as the patient's age, gender, lesion location, patient ID, and whether the lesion was malignant or non-malignant; each picture had a row within the data [20]. This dataset includes meta-information such as the patient's age group (by five-year increments), anatomical location (eight possibilities), and sex (male/female). In addition, there are missing values for specific photos in the metadata. Some Images of melanoma and Non-melanoma are shown in figure 3.

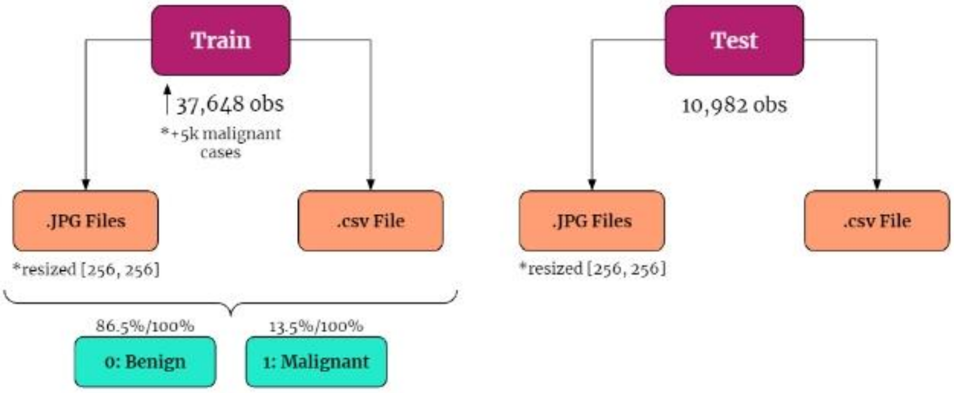

Figure 2. Dataset Overview

Figure 3. Melanoma and Non-melanoma images of Skin Lesions

3.2 Data Augmentation and Normalization

Data augmentation allows for a rebalancing of the dataset's classes, reducing the load on imbalanced classes. Data Augmentation is a successful method for increasing the quantity of training data by randomly altering many training data properties. We employ random brightness and contrast adjustments, random rotation, random flipping, random scaling (or, if the image needs it, proper padding/cropping), and random shearing.

Image normalization is a method that is used to distribute the pixel values in an image uniformly [21]. Furthermore, We utilize the data augmentation variant of Cutout, Progressive Sprinkles, which offers superior learning and transfer due to full pictures with random masking. Images of the Progressive Sprinkles augmentation method are shown in figure 4.

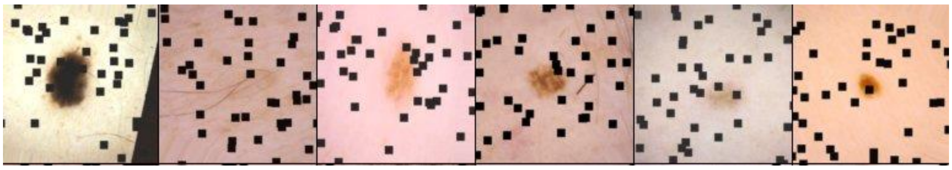

Figure 4. Augmented image with Progressive Sprinkles methods

3.3 EfficientNet

Tan et al. [22] propose a technique for determining the ideal balance of depth, breadth, and resolution in a CNN shown in figure 5. Additionally, the same study described an EfficientNet design for CNNs. The basic model has inverted residual blocks like in MobileNetV2. This structure can be enlarged while maintaining the set of compound coefficient formulas.

$$depth = \alpha^{\theta} \tag{1}$$

$$width = \beta^{\theta} \tag{2}$$

$$resolution = \gamma^{\theta} \tag{3}$$

$$\alpha * \beta^2 * \gamma^2 = 2 \tag{4}$$

While $\alpha, \beta, \gamma \geq 1$

θ is a user-defined coefficient that specifies the number of accessible resources, while α, β, and γ are parameters that define the depth, width, and resolution respectively.

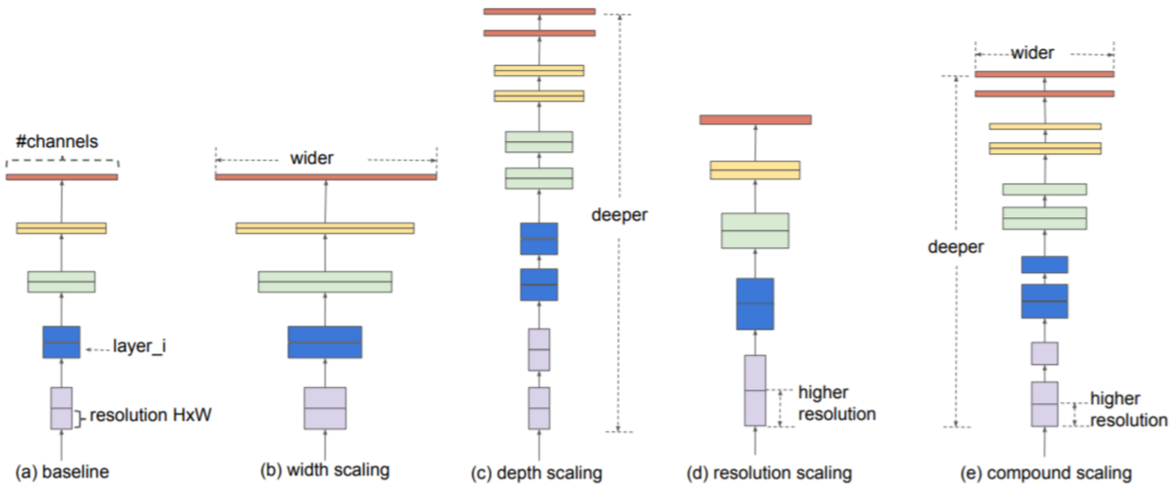

Figure 5. EfficientNet Architecture for determining the ideal balance of depth, breadth, and resolution

3.4 Attention Mechanism

The main idea is that a Global Average Pooling is simple since certain areas are more important than others. So, before pooling, we design an attention mechanism that turns pixels in the GAP on/off and then rescales the results depending on the number of pixels [26]. Thus, the model might be thought of as a "global weighted average" pooling method.

Steps to generate Attention Layer

1. **Compatibility Score**

   The "compatibility score" is calculated by integrating the local features l with the global feature vector g (i.e., dot product).

$$c_i^s = g.\, l_i^s \tag{5}$$

2. **Attention Weights**

$$a_i^s = \frac{exp(c_i^s)}{\sum_j^n \quad exp(c_i^s)} \tag{6}$$

We're going to compress the compatibility scores c into the range (0,1) using a softmax/Sigmoid and name the result a.

3. **Attention Mechanism Outputs for Each Layer**

$$g_a^s = \sum_{i=1}^n \quad a_i^s . l_i^s \tag{7}$$

4. **Final Attention Output Predicts Classification**

We want to categorise an item using the attention outputs g_a from layer l.

3.5 Evaluation Metric

The number of samples that are true-positive (TP), true-negative (TN), false-negative (FN), and false-positive (FP) is a widely used statistic for binary classification assessment (FN). The evaluation metric is precision, recall, and F1-score.

$$Sensitivity/Recall = \frac{TP}{TP + FN} \tag{8}$$

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$F1 - score \ = \ 2 * \frac{(Precision \ * \ Recall)}{(Precision \ + \ Recall)} \tag{10}$$

5. TRAINING AND RESULTS

We trained all the efficient-net models on Kaggle's GPU. For training, some parameters are illustrated in table 2.

We proposed the EfficientNet structure (i.e., B0 to B7), and it is fused with additional patient information(or metadata). After generation output, we ensemble all the EfficientNet structures (i.e., B0, B1..., B7) and achieved a ROC-AUC of 0.9628, as shown in figure 6. Moreover, we compared our proposed method to other methods.

Table 2. Hyper-parameters used for training the EfficientNet.

Parameter Name	**Value**
Input Image	256x256x3
Epochs	30
Optimizer	Adam
Loss Function	Cross Entropy Loss
Batch Size	128
Learning Rate	1e-3

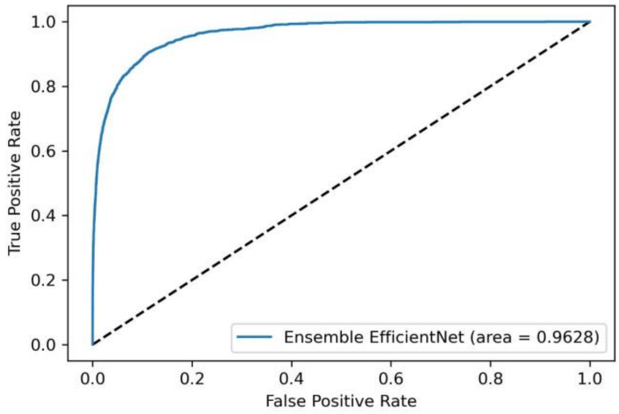

Figure 6. ROC-AUC Curve for Ensemble EfficientNet

According to F1-score, our EfficientNet B6 model outperformed its family illustrated in Table 3.

Table 3. Our Experimental Results on Different EfficientNet Models.

EfficientNet Model	Precision	Recall	F1-Score
B0	0.8534	0.8441	0.8487
B1	0.8683	0.856	0.8621
B2	0.8824	0.8653	0.8738
B3	0.8764	0.8651	0.8698
B4	0.873	0.8868	0.8798
B5	0.8871	0.8676	0.8772
B6	**0.8909**	**0.8795**	**0.8852**
B7	0.8674	0.8466	0.8569

According to F1-score, our ensemble EfficientNet model outperformed from recent studies of ensemble methods illustrated in Table 4.

Table 4. Comparison of Our Ensemble EfficientNet Method to Recent studies.

References	Precision	Recall	F1-Score
[23]	0.8800	0.7500	0.8100
[24]	--	0.8100	0.8128
[25]	**0.8900**	0.8300	0.8300
[26]	0.8599	0.8284	0.8351
[27]	--	0.7570	0.8610
[28]	0.8800	0.7600	0.8200
Our Proposed Method: Ensemble EfficientNet	0.8746	**0.8639**	**0.8629**

6. Conclusion and Future Scope

We demonstrated that an ensemble of EfficientNet-based algorithms might achieve competitive performance in classifying dermoscopic pictures to identify skin lesions/melanoma. This may evolve into a highly automated and accurate technique to categorize dermoscopic pictures in collaboration with experienced dermatologists. However, with a significant difference in the total number of pictures per class, the unbalanced dataset makes it harder to generalize visual features of the lesions. Additionally, the bigger dataset with more feature variation will improve the overall performance of the classification by learning a more accurate representation, reducing the risk of overfitting, and successfully generalizing. The proposed approach achieves a 0.9628 greatest AUC value for the EfficientNet model ensemble. According to the results of various experiments provided in these works and compared to the required evaluations, increasing the number of original photos does not ensure a better classification result in terms of classification metrics. Consequently, we believe that to enhance classifier performance. We must add more images of the same type; i.e., two databases not subjected to the same preprocessing are not always complimentary. Common picture modifications or adjustments, such as rotation, cropping, zooming, contrast/brightness alteration, etc., may influence the performance of the image classifier. This data increase may serve as a regularizer in neural networks, preventing or lowering overfitting and boosting performance while handling unbalanced data sets.

REFERENCES

[1] A. Gálvez and A. Iglesias, "Memetic improved cuckoo search algorithm for automatic B-spline border approximation of cutaneous melanoma from macroscopic medical images," *Advanced Engineering Informatics*, vol. 43, p. 101005, Jan. 2020, doi: 10.1016/j.aei.2019.101005.

[2] "Radiation: Ultraviolet (UV) radiation and skin cancer." https://www.who.int/news-room/q-a-detail/radiation-ultraviolet-(uv)-radiation-and-skin-cancer (accessed Jul. 24, 2021).

[3] "Skin cancer," *WCRF International*. https://www.wcrf.org/dietandcancer/skin-cancer/ (accessed Jul. 24, 2021).

[4] J.-A. Almaraz-Damian, V. Ponomaryov, S. Sadovnychiy, and H. Castillejos-Fernandez, "Melanoma and Nevus Skin Lesion Classification Using Handcraft and Deep Learning Feature Fusion via Mutual Information Measures," *Entropy*, vol. 22, no. 4, p. 484, Apr. 2020, doi: 10.3390/e22040484.

[5] P. Tang, Q. Liang, X. Yan, S. Xiang, and D. Zhang, "GP-CNN-DTEL: Global-Part CNN Model With Data-Transformed Ensemble Learning for Skin Lesion Classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 10, pp. 2870–2882, Oct. 2020, doi: 10.1109/jbhi.2020.2977013.

[6] A. C. Green, N. Pandeya, S. Morton, J. Simonidis, and D. C. Whiteman, "Early detection of melanoma in specialised primary care practice in Australia," *Cancer Epidemiology*, vol. 70, p. 101872, Feb. 2021, doi: 10.1016/j.canep.2020.101872.

[7] A. Naeem, M. S. Farooq, A. Khelifi, and A. Abid, "Malignant Melanoma Classification Using Deep Learning: Datasets, Performance Measurements, Challenges and Opportunities," *IEEE Access*, vol. 8, pp. 110575–110597, 2020, doi: 10.1109/access.2020.3001507.

[8] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Jan. 2017, doi: 10.1038/nature21056.

[9] V. Pomponiu, H. Nejati, and N.-M. Cheung, "Deepmole: Deep neural networks for skin mole lesion classification," Sep. 2016, Accessed: Jul. 24, 2021. [Online]. Available: http://dx.doi.org/10.1109/icip.2016.7532834.

[10] A. Mahbod, G. Schaefer, C. Wang, R. Ecker, and I. Ellinge, "Skin Lesion Classification Using Hybrid Deep Neural Networks," May 2019, Accessed: Jul. 24, 2021. [Online]. Available: http://dx.doi.org/10.1109/icassp.2019.8683352.

[11] S. Sabbaghi, M. Aldeen, and R. Garnavi, "A deep bag-of-features model for the classification of melanomas in dermoscopy images," Aug. 2016, Accessed: Jul. 24, 2021. [Online]. Available: http://dx.doi.org/10.1109/embc.2016.7590962.

[12] S. Demyanov, R. Chakravorty, M. Abedini, A. Halpern, and R. Garnavi, "Classification of dermoscopy patterns using deep convolutional neural networks," Apr. 2016, Accessed: Jul. 24, 2021. [Online]. Available: http://dx.doi.org/10.1109/isbi.2016.7493284.

[13] M. Prathiba, D. Jose, R. Saranya, and Nandhinidevi, "Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks," *IOP Conference Series: Materials Science and Engineering*, vol. 561, p. 012107, Nov. 2019, doi: 10.1088/1757-899x/561/1/012107.

[14] T. Sreelatha, M. V. Subramanyam, and M. N. G. Prasad, "Early Detection of Skin Cancer Using Melanoma Segmentation technique," *Journal of Medical Systems*, vol. 43, no. 7, May 2019, doi: 10.1007/s10916-019-1334-1.

[15] M. A. A. Milton, "Automated Skin Lesion Classification Using Ensemble of Deep Neural Networks in ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection Challenge," *arXiv.org*, Jan. 30, 2019. https://arxiv.org/abs/1901.10802 (accessed Jul. 24, 2021).

[16] A. Mahbod, G. Schaefer, I. Ellinger, R. Ecker, A. Pitiot, and C. Wang, "Fusing fine-tuned deep features for skin lesion classification," *Computerized Medical Imaging and Graphics*, vol. 71, pp. 19–29, Jan. 2019, doi: 10.1016/j.compmedimag.2018.10.007.

[17] A. Soudani and W. Barhoumi, "An image-based segmentation recommender using crowdsourcing and transfer learning for skin lesion extraction," *Expert Systems with Applications*, vol. 118, pp. 400–410, Mar. 2019, doi: 10.1016/j.eswa.2018.10.029.

[18] S. Banerjee, S. K. Singh, A. Chakraborty, A. Das, and R. Bag, "Melanoma Diagnosis Using Deep Learning and Fuzzy Logic," *Diagnostics*, vol. 10, no. 8, p. 577, Aug. 2020, doi: 10.3390/diagnostics10080577.

[19] N. Gessert, M. Nielsen, M. Shaikh, R. Werner, and A. Schlaefer, "Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data," *MethodsX*, vol. 7, p. 100864, 2020, doi: 10.1016/j.mex.2020.100864.

[20] V. Rotemberg *et al.*, "A patient-centric dataset of images and metadata for identifying melanomas using clinical context," *Scientific Data*, vol. 8, no. 1, Jan. 2021, doi: 10.1038/s41597-021-00815-z.

[21] S. G. K. Patro and K. K. sahu, "Normalization: A Preprocessing Stage," *IARJSET*, pp. 20–22, Mar. 2015, doi: 10.17148/iarjset.2015.2305.

[22] M. T. Le Quoc V., "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Accessed: Jul. 24, 2021. [Online].

[23] Md. A. R. Ratul, M. H. Mozaffari, W.-S. Lee, and E. Parimbelli, "Skin Lesions Classification Using Deep Learning Based on Dilated Convolution," Cold Spring Harbor Laboratory, Nov. 2019. Accessed: Jul. 24, 2021. [Online]. Available: http://dx.doi.org/10.1101/860700.

[24] M. A. Al-masni, D.-H. Kim, and T.-S. Kim, "Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification," *Computer Methods and Programs in Biomedicine*, vol. 190, p. 105351, Jul. 2020, doi: 10.1016/j.cmpb.2020.105351.

[25] S. S. Chaturvedi, K. Gupta, and P. S. Prasad, "Skin Lesion Analyser: An Efficient Seven-Way Multi-class Skin Cancer Classification Using MobileNet," in *Advances in Intelligent Systems and Computing*, Singapore: Springer Singapore, 2020, pp. 165–176.

[26] J. Zhang, Y. Xie, Y. Xia, and C. Shen, "Attention Residual Learning for Skin Lesion Classification," *IEEE Transactions on Medical Imaging*, vol. 38, no. 9, pp. 2092–2103, Sep. 2019, doi: 10.1109/tmi.2019.2893944.

[27] N. Gessert *et al.*, "Skin Lesion Classification Using CNNs With Patch-Based Attention and Diagnosis-Guided Loss Weighting," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 2, pp. 495–503, Feb. 2020, doi: 10.1109/tbme.2019.2915839.

[28] P. Carcagnì *et al.*, "Classification of Skin Lesions by Combining Multilevel Learnings in a DenseNet Architecture," in *Lecture Notes in Computer Science*, Cham: Springer International Publishing, 2019, pp. 335–344.

[29] Hemant Kumar, Amit Virmani, Shivneet Tripathi, Rashi Agrawal, Sunil Kumar, "Transfer Learning and Supervised Machine Learning Approach for Detection of Skin Cancer: Performance Analysis and Comparison ", *dcth*, vol. 10, no. 1, pp. 1845–1860, Sep. 2021.