Çanakkale Onsekiz Mart University

Journal of Advanced Research in Natural and Applied Sciences

dergipark.org.tr/tr/pub/jarnas

Open Access

e-ISSN: 2757-5195 | 2024, Vol. 10, Issue 4, Pages: 996-1006 | doi.org/10.28979/jarnas.1505804

Classification of Melanoma Cancer Using Deep Convolutional Neural Networks

Ali Güneş[1], Emrah Dönmez[2]

[1]Orzax Inc, 901 Mittel Dr Wood Dale, Wood Dale, Chicago, United States

[2]Department of Software Engineering, Faculty of Engineering and Natural Science, Bandırma Onyedi Eylül University, Balıkesir, Türkiye

Article Info

Received: 27 Jun 2024

Accepted: 16 Oct 2024

Published: 31 Dec 2024

Research Article

Abstract— Accurate detection of skin diseases is crucial in healthcare, with early diagnosis being particularly vital for effective treatment. Melanoma, a form of skin cancer with a high potential for metastasis, requires early detection to significantly improve treatment success and prevent further spread across the skin. This study investigates the application of machine learning techniques to diagnose skin lesions, focusing on differentiating between benign moles and malignant melanoma. A Convolutional Neural Network (CNN) model was developed to explore machine learning's efficacy in this context. The initial model featured a primary architecture, progressively refined by adding additional layers and filters to increase its complexity. This iterative enhancement aimed to improve the model's capability to extract and analyze features from skin images. Each model configuration was meticulously evaluated through a series of experiments to determine its diagnostic performance. The results revealed that the proposed CNN model achieved a high accuracy rate of 91%. This significant finding demonstrates the effectiveness of machine learning approaches in the early diagnosis and management of melanoma. The study confirms that advanced CNN architectures can enhance diagnostic precision, thereby contributing to improved patient outcomes in detecting and treating skin diseases.

Keywords — *CNN, artificial learning, melanoma, mole (nevus), skin cancer*

1. Introduction

Melanoma is a form of skin cancer that arises when melanocytes, the cells responsible for giving skin its tan or brown hue, start to grow uncontrollably. Cancer begins when cells in the body proliferate without regulation. Virtually any cell in the body has the potential to become cancerous and subsequently spread to other regions. Melanoma is also known as malignant melanoma or cutaneous melanoma. Typically, melanoma cells generate melanin, causing the tumors to appear brown or black. Nonetheless, certain melanomas lack melanin production and may present as pink, tan, or even white. Figure 1 illustrates melanocyte cells. The image credit in Figure 1 is attributed to Designua/Shutterstock.

[1]alisenug@gmail.com (Corresponding Author); [2]emrahdonmez@bandirma.edu.tr

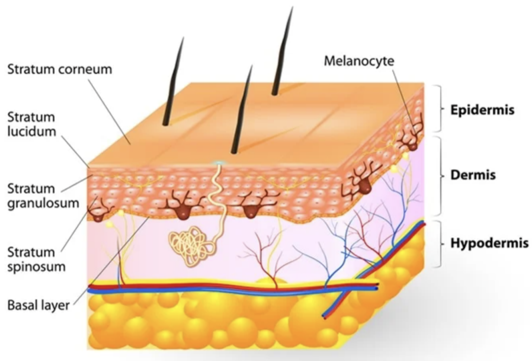

Figure 1. Melanocytes located between the upper and lower skin

In 2018, skin cancer emerged as a widespread and perilous condition on a global scale, with 300,000 new diagnoses and over 1 million fatalities occurring each month worldwide [1]. Melanoma is increasingly common worldwide, ranking as the 19th most prevalent disease and exhibiting one of the highest mortality rates [1]. According to the International Agency for Research on Cancer (IARC), approximately 19.3 million new cancer cases were diagnosed in 2020, and about 10 million individuals succumbed to the disease. Furthermore, in the United States, 100,350 new cases were diagnosed in 2020, and about 6,850 people died [2, 3].

Controlling cancer-related mortality is a challenging process; however, recent advances in image processing and artificial intelligence approaches can help in the early detection of melanoma, as early detection and prognosis can improve survival rates. Extensive research solutions and computer vision algorithms have been proposed in the literature to diagnose skin lesions at the earliest stage and overcome the complexities of traditional approaches [4]. Classification methods, such as Decision Trees (DT) [5], Support Vector Machines (SVM) [6], and Artificial Neural Networks (ANN) [7], have been presented in various approaches. Many machine learning techniques face data processing challenges that necessitate high-contrast, noise-free, and clean images, which is often not the case with skin cancer data. Skin classification relies on features like color, texture, and structure. Due to the high inter-class homogeneity and intra-class heterogeneity of skin lesions, using low-quality feature sets can lead to inaccurate classification results [8]. A “low-quality feature” refers to an attribute or variable in the dataset that does not significantly contribute to the model’s predictive power or performance. Such features can impair model performance and result in inaccurate or sub-optimal outcomes. Traditional methods are parametric and often require normally distributed training data; however, skin cancer data is typically unpredictable and may not conform to these assumptions. Each lesion presents a unique pattern, making these methods insufficient. Therefore, deep learning (DL) techniques are more effective for skin classification, enabling dermatologists to diagnose lesions more accurately.

Detecting skin diseases is crucial in healthcare due to its significant impact on patient outcomes. Among various skin conditions, melanoma, a type of skin cancer, is particularly critical because early detection is known to substantially enhance the success of treatment. Identifying melanoma in its initial stages is vital as it prevents the cancer from spreading further across the skin, which can otherwise lead to more severe health issues. This study focuses on diagnosing skin lesions, specifically moles (benign tumors) and melanoma (a malignant skin cancer), through machine learning techniques.

To achieve this, the research involved developing a Convolutional Neural Network (CNN) model, which was initially simple. Gradually, more layers and filters were incorporated into the model, increasing its complexity. The performance of these enhanced CNN models was carefully evaluated and analyzed. The experimental results revealed that the proposed method achieved an impressive success rate of 91%. These results highlight the effectiveness and potential of machine learning approaches in the early diagnosis and treatment of melanoma, demonstrating their significant role in improving healthcare outcomes.

The remainder of this paper is structured as follows: In the second part, studies on melanoma cancer are presented. The main motivation of the study, the methods, and the dataset are presented in the third and fourth sections, respectively. The results of the experimental studies are presented in the fifth section, while the discussion and conclusions on the results obtained are included in the last section.

2. Related Works

DL has achieved great success in medical image classification and many other fields and is gradually replacing other machine-learning methods. Medical images play an essential role in diagnosis and treatment processes. By learning complex relationships in these images, DL models can be effective in many areas, such as disease diagnosis, lesion detection, and drug discovery. Moreover, DL has advantages over other machine learning methods in tasks such as image classification and feature extraction with high accuracy rates.

Promising results have been obtained in studies conducted in the literature. Litjens et al. [9] have shown that DL models can be successfully used in prostate and breast cancer staging. This research reveals that DL methods have significant potential in cancer diagnosis and staging using medical images. Their study shows that DL models can be more frequently used in the early diagnosis and treatment planning of essential diseases, such as prostate and breast cancer. Liu et al. [10] have developed a new deep learning model to detect prodromal and mild cognitive impairment stages of Alzheimer's disease. Their study demonstrates the potential of DL methods for early diagnosis and progression monitoring of Alzheimer's disease. Their research involves the creation of a customized model using deep neural networks on large datasets. The new model has successfully identified markers specific to Alzheimer's disease. A new CNN model proposed by Wang et al. [11] presents an approach that requires less user interaction and can perform medical image segmentation faster than existing methods. Automatic image segmentation refers to the process of identifying and distinguishing specific structures, e.g., organs or lesions, in medical images. Electroencephalography (EEG) signals are data recording electrical activity used to measure brain activity. Analysis of these signals plays a vital role in the diagnosis of epilepsy, sleep disorders, and assessment of neurological and psychiatric conditions such as attention deficit hyperactivity disorder (ADHD). Acharya et al. [12] have used a 13-layer CNN to analyze EEG signals and achieved high accuracy classification.

Esteva et al. [13] have studied skin cancer classification using deep neural networks. They have investigated the feasibility of using deep neural networks to achieve dermatologist-level classification performance. Besides, they have used a large skin cancer dataset containing 129,450 images. The researchers have achieved the accuracy rate $72.1\% \pm 0.9\%$ when using Google's Inception v3 architecture by scaling the images to 299×299. Harangi [14] has evaluated the effectiveness of deep CNN ensembles in classifying skin lesions. Ensemble methods are a model created by combining several different CNNs.

DL techniques' primary advantage lies in their ability to directly apply to classification tasks without requiring pre-processing steps. In a notable study, Yap et al. [15] have introduced a method that integrates various image modalities and patient metadata. They have employed the Residual Network 50 (ResNet50) [16] network distinctly on both dermoscopic and macroscopic images, subsequently combining their features for the final classification task. This multi-model classifier demonstrated superior performance compared to the baseline model, which utilized only macroscopic images, achieving an Area Under Curve (AUC) value of 0.866. Similarly, Gessert et al. [17] have developed an ensemble model comprising EfficientNets [18], Squeeze-and-Excitation Network (SENet) [19], and Residual

Network Next Generation Weakly Supervised Learning (ResNeXt WSL) [20] to conduct a multi-class classification on the International Skin Imaging Collaboration (ISIC) 2019 dataset. A clipping strategy is applied to the images to cope with multiple model input resolutions. Furthermore, a loss balance approach has been used to deal with imbalanced datasets. Srinivasu et al. [21] have introduced a deep convolutional neural network (DCNN) that integrates MobileNetV2 [22] with Long Short-Term Memory (LSTM) [23] for classifying lesions on the HAM10000 dataset. The MobileNetV2 model demonstrated high efficiency and accuracy, making it suitable for lightweight computational devices. The proposed model excels in maintaining stateful information, leading to more precise predictions. Additionally, a grey-level co-occurrence matrix is employed to evaluate disease progression. The performance of this approach has been compared with other advanced models, including Fine-Tuned Neural Networks (FTNN) [24], CNN, Very deep Convolutional Networks for Large-Scale Image Recognition [25] developed by the Visual Geometry Group (VGG), and a modified convolutional neural network architecture. Utilizing the HAM10000 dataset, the proposed method outperformed these models with over 85% accuracy.

3. Materials and Methods

Melanoma skin cancer is one of the most crucial skin disorders. Early diagnosis of melanoma is critical for successful treatment of the disease. Today, melanoma is diagnosed by a specialist physician. Computer-aided systems are also being developed to improve the diagnostic process. Deep learning-based approaches (CNN, etc.) are widely used in these systems. In the CNN model, there are various operations performed in layers. This study evaluates the effects of using different numbers of CNN convolution layers and filters on classification performance.

3.1. Dataset

The dataset used in the study includes a total of 10,605 image data. In the training set, there are a total of 9,605 images, of which 5,000 are labeled as benign tumors and 4,605 as melanoma. The test set includes a total of 1,000 images of which 500 are benign tumor images and 460 melanomas. In experimental studies, 70% of the data for the two classes was utilized for training, while 30% was used for testing. Some sample images of the dataset are shown in Figure 2.

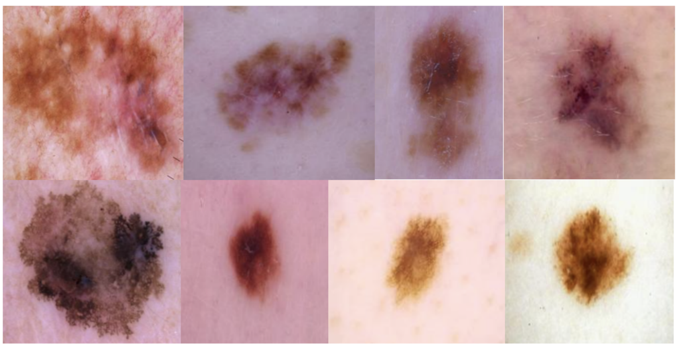

Figure 2. First and second rows show malignant and bening samples, respectively

3.2. Deep Features

Extracting relevant attributes is fundamental for predicting patterns and making informed decisions in dataset analysis. Traditionally, feature extraction involved applying mathematical techniques or utilizing third-party methods, such as Speed-Up Robust Features (SURF) [26] and Scale-Invariant Feature Transform (SIFT) [27], designed to detect and describe local features in images. However, the advent of CNNs has revolutionized this process.

CNNs are advanced, multi-layered neural network architectures specifically designed to automatically learn and extract features from raw data. They consist of several types of layers, including convolutional layers that apply various filters to detect features, pooling layers that reduce the dimension while retaining important information, and fully connected layers that integrate these features for final classification or regression tasks.

The features obtained through these CNNs are known as deep features. These features are termed "deep" because they are generated through multiple layers of convolution and pooling, allowing the network to capture increasingly abstract and complex representations of the data. This hierarchical learning process enables CNNs to discern high-level patterns and intricate details within the data, making them highly effective for tasks, such as image recognition and classification.

3.3. CNN Model Configuration

CNNs represent a sophisticated class of machine-learning architectures designed for image classification tasks. These networks process images through a series of specialized layers. Initially, the image is fed into the input layer. Subsequently, the convolutional layers apply various filters to extract features, while normalization techniques are employed to standardize the aggregated data across different image channels. The values are then processed through an activation layer in which they are compared against a predefined threshold. Following this, sub-sampling is performed through pooling operations, which reduce the spatial dimensions of the data while preserving essential features. This sequence of convolution, normalization, activation, and pooling layers is iterated multiple times to refine the data representations progressively. Finally, the refined data is forwarded through one or more fully connected layers, which integrate the features for comprehensive analysis. Figure 3 shows the layer structures for a basic CNN model.

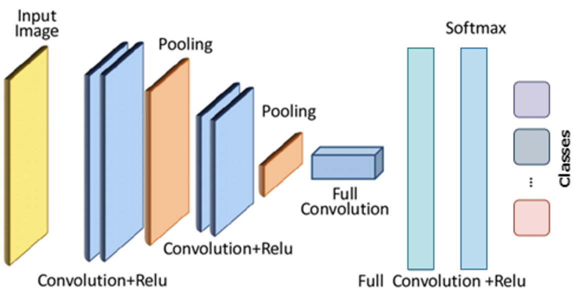

Figure 3. Basic CNN structure

We used the algorithm RMSprop (Root Mean Square Propagation) [28] for optimization, a key hyperparameter that adapts the learning rate based on recent gradients. RMSprop stabilizes training and accelerates convergence, essential for training our deep CNN model. The output from the fully connected layers is then classified using a soft-max function, which provides probabilistic predictions for each class. A key characteristic of CNNs is that the features extracted from the data typically improve in quality as the network depth increases. This network deepening enhances its ability to capture complex and hierarchical patterns within the image data, leading to more accurate classification results.

3.4. CNN Model with Variable Convolution Layering

This study developed various configurations of CNNs to assess the impact of convolutional layers and the number of filters on model performance. A critical aspect of these configurations is the choice of filter size and the number of filters used in each convolutional layer. Filter size determines the dimensions of the receptive fields that the network uses to extract features from the input data, influencing the granularity of the detected patterns. Meanwhile, the number of filters controls the number of distinct features that can be learned and captured at each layer, thereby affecting the network's ability to recognize and represent complex patterns.

Table 1 details the parameters for filter size (FS) and number of filters (NoF) employed across the convolutional layers in these CNN models. By varying these parameters, we aimed to evaluate their effects on the network's performance and feature extraction capabilities, providing insights into how different configurations impact the overall effectiveness of the CNN in handling and classifying image data.

Table 1. Convolution layers' parameters

Layer	FS	NoF
Conv1	[3 3]	8
Conv2	[3 3]	16
Conv3	[3 3]	32
Conv4	[3 3]	64
Conv5	[3 3]	128

The models created have the convolution layers specified in Table 2. All convolution layers have the same padding ratio.

Table 2. CNN models' convolution layers

Model	Convolution Layers
CNN1	Conv1
CNN2	Conv1, Conv2
CNN3	Conv1, Conv2, Conv3
CNN4	Conv1, Conv2, Conv3, Conv4
CNN5	Conv1, Conv2, Conv3, Conv4, Conv5

The models have a maximum pooling layer after each convolution layer. Maximum pooling computes the average of the elements present in the region of the feature map covered by the filter. In the maximum pooling layer for all models, we used the pool size of [2 2] and the stride value of [2 2]. All models have in common a fully connected layer, a SoftMAX (Soft Maximum) [29] activation layer and a classification layer.

4. Experimental Results

Five different CNN models were designed within the experiments' scope, and experiments were carried out. The CNN model image input size is $64 \times 64 \times 3$ and in Red-Green-Blue (RGB) space. The hyper-parameter settings are common to all CNN models and are provided in Table 3.

Table 3. Common hyper-parameters for CNN models

Training Parameters
Max Epochs = 30
Initial Learn Rate = 0.01
Optimizer = 'Rmsprop'
Shuffle = 'every-epoch'
Validation Frequency = 30
Execution Environment = GPU

Figures 4 and 5 show the accuracy versus loss function values of CNN models with one and five convolution layers. While the accuracy value increases slightly in most iterations, the loss function tends to decrease. The decrease in the loss function at each iteration indicates that the model learns the patterns and relationships in the training data better. The parameters were updated to minimize the error. Thus, the improvement of the model is observed.

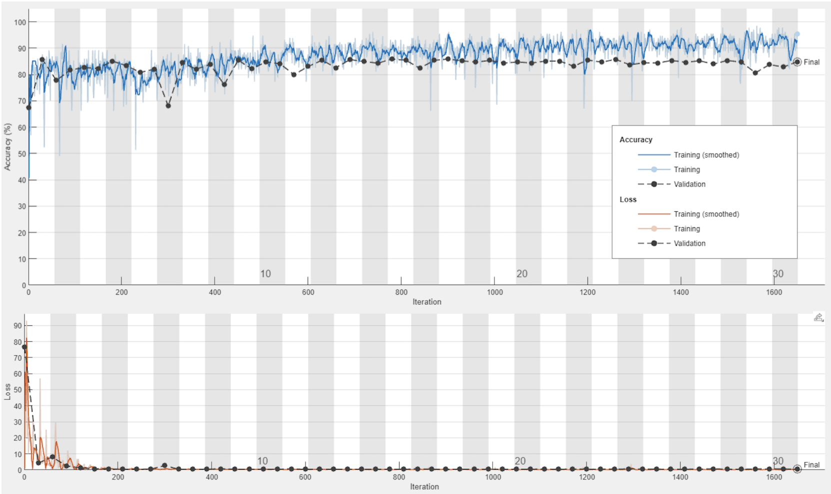

Figure 4. CNN model performance with a single convolution layer

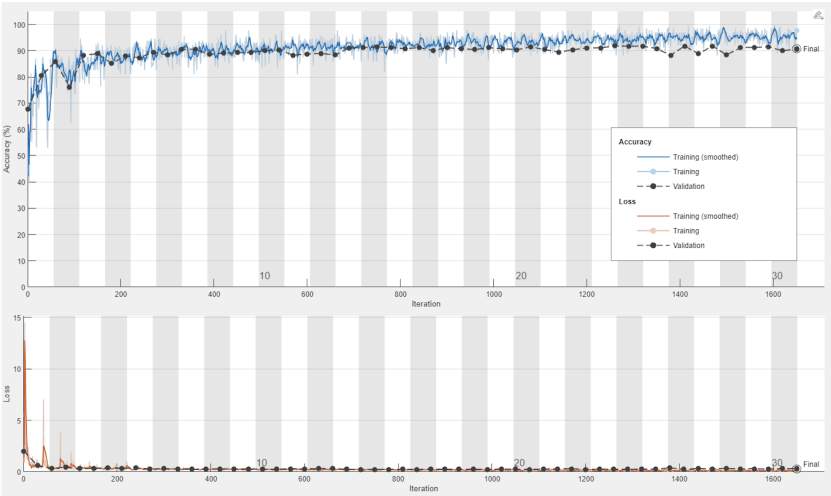

Figure 5. CNN model performance with five convolution layers

It can be observed Figures 4 and 5 that adding more convolution layers and filters improves classification performance. The confusion matrices' values for the CNN models are shown in Table 4. Confusion matrices summarize the performance of a classification model by presenting the counts of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) predictions.

Table 4. Confusion matrices' values for CNN models

Model	TP	FP	TN	FN
CNN1	1693	293	1238	233
CNN2	1707	180	1351	219
CNN3	1719	134	1397	207
CNN4	1770	144	1387	156
CNN5	1757	146	1385	169

The results obtained by the performance metrics Accuracy (Acc), Sensitivity (Sen), Specificity (Spe), and F1-Score (F1) using the confusion matrices of CNN models are provided in Table 5.

Table 5. Results obtained by the performance metrics using the confusion matrices of CNN models

Model	Acc	Sen	Spe	F1
CNN1	84,78%	0,879	0,809	0,866
CNN2	88,46%	0,886	0,882	0,895
CNN3	90,14%	0,893	0,912	0,910
CNN4	91,32%	0,919	0,906	0,92
CNN5	90,89%	0,912	0,905	0,918

As can be observed from Table 5, while increasing the number of convolution layers and the number of filters improved the performance up to a certain point, the performance of the final model decreased compared to those of the previous model. Doubling the number of filters in each convolution layer compared to the previous one improved the performance to a certain extent.

5. Discussion

Accurate detection of skin diseases is vital in healthcare due to its significant impact on patient outcomes. Among various skin conditions, melanoma—a severe form of skin cancer—is particularly critical. Early detection of melanoma greatly enhances treatment success by preventing further spread and more severe complications. A comparison with similar studies is made in Table 6 below. As can be seen from the table, a significant accuracy value has been obtained with the proposed simple CNN model. It should be stated that the convolution layers and the number of filters in these layers are effective for the final performance, in the proposed model.

Table 6. Comparison with featured studies

Study	Method	Data	Performance
[5]	Watershed Segmentation, kNN (k-Nearest Neighbors) [30], RandomForest [31], SVM	1000	89.43% Acc
[6]	ResNet-50, XGBoost (eXtreme Gradient Boosting) [32], Statistical Analysis	11444	89.00% Acc
[7]	Caffe CNN [33], Sparse Coding [34], SVM	2624	73.90% Acc
[14]	Pre-trained CNNs [35], Ensemble [36], Majority Voting [37]	2000	89.10% Acc
This Study	Configurable Simple CNN Model	10605	91.32% Acc

This study used advanced machine learning techniques to diagnose skin lesions, particularly distinguishing between benign moles and malignant melanoma.

We developed a CNN model, which was initially simple but improved by adding more layers and filters. Increasing the number of convolutional layers and filters enhances the model's ability to learn complex features. Despite these improvements, the increased complexity of the model introduces challenges such as over-fitting and higher computational costs. To address these challenges, future research should explore strategies such as integrating multi-modal data, leveraging transfer learning to reduce training time, and employing advanced regularization techniques. Developing real-time diagnostic tools and mobile applications could also facilitate broader and more accessible melanoma screening. Future studies can enhance the efficacy and practicality of early melanoma detection and improve overall healthcare outcomes.

6. Conclusion

Early and precise diagnosis of melanoma is essential for successful treatment and prevention of disease progression. This study explored the application of CNNs to differentiate between benign moles and malignant melanoma, achieving a high accuracy rate of 91%. By systematically enhancing the model architecture with additional layers and filters, we demonstrated the effectiveness of CNN-based approaches in accurately identifying skin lesions. These findings confirm that machine learning models can significantly contribute to the early detection and management of melanoma when properly developed and optimized. Integrating such technology into clinical practice could transform how skin diseases are diagnosed, leading to more efficient and accurate interventions. Future research should aim to further improve model accuracy by expanding the dataset and exploring complementary algorithms to enhance diagnostic reliability and scalability.

Author Contributions

All the authors equally contributed to this work. They all read and approved the final version of the paper.

Conflicts of Interest

All the authors declare no conflict of interest.

Ethical Review and Approval

No approval from the Board of Ethics is required.

References

[1] A. Masood, A. Al-Jumaily, *Computer-aided diagnostic support system for skin cancer: A review of techniques and algorithms*, International Journal of Biomedical Imaging 2013 (1) (2013) 323268 22 pages.

[2] International Agency for Research on Cancer, https://www.iarc.who.int, Accessed on 15 May 2024.

[3] American Cancer Society, Key Statistics for Melanoma Skin Cancer (2024), https://www.cancer.org/cancer/melanoma-skin-cancer/about/key-statistics.html, Accessed on 15 May 2024.

[4] K. Thomsen, L. Iversen, T. L. Titlestad, O. Winther, *Systematic review of machine learning for diagnosis and prognosis in dermatology*, Journal of Dermatological Treatment 31 (5) (2020) 496–510.

[5] J. R. Quinlan, *Induction of decision trees*, Machine Learning 1 (1986) 81–106.

[6] C. Cortes, V. Vapnik, *Support-vector networks*, Machine Learning 20 (3) (1995) 273–297.

[7] F. Rosenblatt. *The perceptron: A probabilistic model for information storage and organization in the brain*, Psychological Review 65 (6) (1958) 386—408.

[8] N. Codella, J. Cai, M. Abedini, R. Garnavi, A. Halpern, J. R. Smith, *Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images*, in L. Zhou, L. Wang, Q. Wang, Y. Shi (Eds.), Machine Learning in Medical Imaging, 2015, pp. 118–126.

[9] G. Litjens, C. I. Sanchez, N. Timofeeva, M. Hermsen, I. Nagtegaal, I. Kovacs, C. H. van de Kaa, P. Bult, B. ban Ginneken, J. van der Laak, *Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis*, Scientific Reports 6 (2016) Article Number 26286 11 pages.

[10] S. Liu, W. Cai, S. Pujol, R. Kikinis, D. Feng, *Early diagnosis of Alzheimer's disease with deep learning*, in G. Wang, B. He (Eds.), IEEE 11th International Symposium on Biomedical Imaging, Beijing, 2014, pp. 1015–1018.

[11] G. Wang, W. Li, M. A. Zuluaga, R. Pratt, P. A. Patel, M. Aertsen, T. Vercauteren, *Interactive medical image segmentation using deep learning with image-specific fine-tuning*, IEEE Transactions on Medical Imaging 37 (7) (2018) 1562–1573.

[12] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, H. Adeli, *Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals*, Computers in Biology and Medicine 100 (2018) 270–278.

[13] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, S. Thrun, *Dermatologist-level classification of skin cancer with deep neural networks*, Nature 542 (2017) 115–118.

[14] B. Harangi, *Skin lesion classification with ensembles of deep convolutional neural networks*, Journal of Biomedical Informatics 86 (2018) 25–32.

[15] J. Yap, W. Yolland, P. Tschandl, *Multimodal skin lesion classification using deep learning*, Experimental Dermatology 27 (11) (2018) 1261–1267.

[16] K. He, X. Zhang, S. Ren, J. Sun, *Deep residual learning for image recognition*, in T. Tuytelaars, F.-F. Li, R. Bajcsy (Eds.), 29th IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, 2016, pp. 770–778.

[17] N. Gessert, M. Nielsen, M. Shaikh, R. Werner, A. Schlaefer, *Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data*, MethodsX 7 (2020) 100864 8 pages.

[18] M. Tan, Q. V. Le, *EfficientNet: Rethinking model scaling for convolutional neural networks*, in K. Chaudhuri, R. Salakhutdinov (Eds.), 36th International Conference on Machine Learning, California, 2019, pp. 6105–6114.

[19] J. Hu, L. Shen, G. Sun, *Squeeze-and-excitation networks*, in M. S. Brown, B. Morse, S. Peleg (Eds.), IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018, pp. 7132–7141.

[20] D. Mahajan, R. Girshick, V. Ramanathan, H. Kaiming, M. Paluri, Y. Li, A. Bharambe, L. V. D. Maaten, *Exploring the limits of weakly supervised pretraining*, in V. Ferrari, M. Hebert, C. Sminchisescu, Y. Weiss (Eds.), 15th European Conference on Computer Vision, Munich, 2018, pp. 181–196.

[21] P. N. Srinivasu, J. G. SivaSai, M. F. Ijaz, A. K. Bhoi, W. Kim, J. J. Kang, *Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM*, Sensors 21 (8) (2021) 2852 27 pages.

[22] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L. Chen, *MobileNetV2: Inverted residuals and linear bottlenecks*, in M. S. Brown, B. Morse, S. Peleg (Eds.), IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018, pp. 4510–4520.

[23] S. Hochreiter, J. Schmidhuber, *Long short-term memory*, Neural Computation 9 (8) (1997) 1735–1780.

[24] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, *How transferable are features in deep neural networks*, in Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, K. Q. Weinberger (Eds.), 28th International Conference on Neural Information Processing Systems, Cambridge, 2014, pp. 3320–3328.

[25] K. Simonyan, A. Zisserman, *Very deep convolutional networks for large-scale image recognition*, in Y. Bengio, Y. LeCun (Eds.), International Conference on Learning Representations, San Diego, 2015, pp. 1–14.

[26] H. Bay, T. Tuytelaars, L. Van Gool, *SURF: Speeded up robust features*, in A. Leonardis, H. Bischof, A. Pinz (Eds.), 9th European Conference on Computer Vision, Graz, 2006, pp. 404–417.

[27] D. G. Lowe, *Distinctive image features from scale-invariant keypoints*, International Journal of Computer Vision 60 (2) (2004) 91–110.

[28] M. C. Mukkamala, M. Hein, *Variants of RMSProp and Adagrad with Logarithmic Regret Bounds*, in D. Precup, Y. W. Teh (Eds.), 34th International Conference on Machine Learning, Sydney, 2017, pp. 2545–2553.

[29] C. M. Bishop, Pattern recognition and machine learning, Springer, New York, 2006.

[30] E. Fix, J. L. Hodges, *Discriminatory analysis, nonparametric discrimination: Consistency properties*, Technical Report USAF School of Aviation Medicine (1951) Dayton.

[31] T. K. Ho, *Random decision forests*, in C. Y. Suen (Ed.), 3rd International Conference on Document Analysis and Recognition, Montreal, 1995, pp. 278–282.

[32] T. Chen, C. Guestrin, *XGBoost: A scalable tree boosting system*, in B. Krishnapuram, M. Shah (Eds.), 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, 2016, pp. 785–794.

[33] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell. *Caffe: Convolutional architecture for fast feature embedding*, in K. A. Hua, Y. Rui, R. Steinmetz (Eds.), 22nd ACM International Conference on Multimedia, Orlando , 2014, 675–678.

[34] B. A. Olshausen, D. J. Field, *Emergence of simple-cell receptive field properties by learning a sparse code for natural images*, Nature 381 (1996) 607–609.

[35] A. Krizhevsky, I. Sutskever, G. E. Hinton, *ImageNet classification with deep convolutional neural networks*, in F. Pereira, C. J. Burges, L. Bottou, K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 25, Lake Tahoe, 2012, pp. 1097–1105.

[36] T. G. Dietterich, *Ensemble methods in machine learning*, in J. Kittler, F. Roli (Eds.), Multiple Classifier Systems, Cagliari, 2000, pp. 1–15.

[37] M. Kearns, Thoughts on Hypothesis Boosting (1988), https://www.cis.upenn.edu/~mkearns/, Accessed 10 May 2024.