



Transfer Learning-Based Skin Tumor Identification Improvement

Shafira Maharani¹, Dina Tri Utari^{2*}

Department of Statistics, Universitas Islam Indonesia, Yogyakarta 55584, Indonesia

Corresponding Author Email: dina.t.utari@uii.ac.id

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/mmep.111020>

ABSTRACT

Received: 9 July 2024

Revised: 4 September 2024

Accepted: 14 September 2024

Available online: 31 October 2024

Keywords:

skin tumor, CNNs, transfer learning, MobileNetV2, VGG16, VGG19

This study addresses the critical challenge of accurately identifying skin disorders as benign, malignant, or non-tumors, essential for timely and successful treatment. Early identification can greatly minimize tumor development and cut fatality rates. Given the high costs involved with standard medical detection approaches, this research addresses using sophisticated Convolutional Neural Networks (CNNs) with transfer learning to categorize skin malignancies efficiently. Specifically, the study assesses the performance of MobileNetV2, VGG16, and VGG19 architectures. The primary objective is to find which model has the maximum accuracy in classifying skin cancers. Our findings reveal that while a standard CNN reached an accuracy of 62.2%, the transfer learning models greatly outperformed it, with MobileNetV2 achieving the highest accuracy at 93.9%, followed by VGG19 at 90.0% and VGG16 at 88.9%. These results imply that MobileNetV2 is the most successful solution for this task since it consistently obtained a prediction accuracy of 90% for both in-dataset and out-of-dataset images.

1. INTRODUCTION

Tumors, irrespective of their classification as benign or malignant, represent abnormal proliferation within the body that may manifest varying levels of severity. Benign tumors are generally considered less harmful as they do not develop into cancer and can be effectively controlled. In contrast, malignant tumors, known as cancer, pose a greater risk, often leading to cancerous growth and potentially resulting in death. Malignant tumors display swift expansion, infiltration of neighboring tissues, and the ability to metastasize to distant body sites, making them highly lethal [1, 2].

Cancer, identified as malignant tumors, is widely recognized as a primary factor contributing to mortality on a global scale. The most perilous type of cancer to human life, characterized by its rapid proliferation, is skin cancer, encompassing melanoma and non-melanoma. The upcoming year, 2020, will witness a substantial surge in new cases of skin cancer and fatalities attributed to this ailment. Indonesia grapples with a pronounced prevalence of both forms of skin cancer, melanoma, and non-melanoma, indicating a significant health challenge in the region [3].

Many researchers have explored using advanced deep learning techniques, such as convolutional neural networks (CNNs). The CNNs perform well in classifying images similar to the training dataset. However, they often struggle with images that have tilt or rotation. To improve future technology, the proposal suggests using fully CNNs to classify 3D images, which could address these challenges and enhance classification accuracy in such scenarios [4].

A study compared different Deep Convolutional Neural

Network (DCNNs) models for classifying MR brain images. The pre-trained InceptionV3 DCNN model used in this approach achieved a classification accuracy of 99.82% on the Figshare MRI dataset, outperforming other models and setting a high classification standard [5]. Moreover, the study [6] aimed to analyze the impact of different data modalities (text, images, and a combination of both) using Deep Learning (DL) models for emotion analysis. Pre-trained models like ResNet152V2, VGG19, and EfficientNetB7 were used for image classification, while CNNs and CNNs+LSTM were applied for text. EfficientNetB7 achieved the best image performance and CNNs with Glove embedding for text. An early fusion technique combining CNNs and EfficientNetB7 produced strong results in multimodal analysis. The proposed model outperformed baseline studies, achieving high accuracy and F1-macro scores, demonstrating its effectiveness for meme classification.

The study highlights the effectiveness of using CNNs for seed classification in agriculture. To enhance model robustness, a modified VGG architecture with two 0.5 dropout layers after the dense layer was employed. Transfer learning and fine-tuning techniques improved performance, reduced computational demands, and shortened training time, making the seed classification process more efficient and effective [7].

Based on the previous study, using CNNs and transfer learning models has demonstrated promising results in various applications. To confront the challenges of diagnosing and classifying skin neoplasms, the CNNs has exhibited promising outcomes in effectively distinguishing between benign and malignant skin tumors, thereby assisting in promptly identifying and categorizing such neoplasms. Studies have

indicated that CNNs integrating transfer learning have attained notable precision in categorizing skin ailments, including melanoma. The integration of pre-existing models and sophisticated architectures such as MobileNet, VGG-16, and ResNet-50 has facilitated researchers in achieving a substantial degree of accuracy in skin tumor classification tasks [8-10].

The implementation of CNNs in conjunction with transfer learning has significantly improved the accuracy of skin lesion classification by utilizing a robust pre-trained model and optimizing feature extraction. Various investigations have proven accuracy levels above 70% using models such as MobileNetV2, VGG16, and VGG19. Notably, MobileNetV2 distinguishes itself by effectively balancing precision and computational efficiency. These models are critical in advancing the differentiation of benign, malignant, and non-tumorous skin lesions, laying a solid groundwork for prospective investigations in this domain [11].

Nevertheless, although prior research has investigated the use of CNNs and transfer learning in skin cancer diagnosis, there is a lack of detailed comparative analysis of models such as MobileNetV2, VGG16, and VGG19. This work addresses this information gap by thoroughly evaluating these models, emphasizing their classification accuracy, computational efficiency, and practical usefulness in clinical environments. The originality of this study resides in its methodical assessment of these models, providing fresh perspectives on their advantages and constraints and directing future investigation and implementation in the field of medical image analysis.

Moreover, the application of transfer learning in medical imaging, namely in skin cancer categorization, has played a crucial role in surmounting obstacles associated with restricted training datasets. Using pre-trained models, researchers can extract significant characteristics from medical images, improving the classification accuracy of CNNs [12-15]. This study expands upon prior research by not only evaluating the efficacy of CNNs models but also examining the influence of

transfer learning on the capacity of these models to distinguish between benign, malignant, and non-tumorous skin lesions.

In contrast to previous studies that often assess a single model or a restricted range of topologies, our work thoroughly evaluates three distinct CNNs architectures. The present study provides novel perspectives on each model’s comparative merits and limitations regarding classification accuracy, computational efficiency, and resilience to diverse datasets. This allows us to identify the most effective model for skin tumor classification, which is critical for clinical application. Moreover, we assess in-dataset performance and out-of-dataset images, demonstrating the robustness and generalizability of the transfer learning model, which consistently achieves high accuracy across different datasets. A key focus of our work is to provide dependable and scalable solutions in medical imaging, especially in settings with limited resources where the consequences of misclassification might be significant.

Subsequent sections of this paper will provide detailed explanations of the methodology used, show the outcomes of our comparative analysis, and examine the consequences of our findings for future scholarship and clinical application. This organizational framework aims to offer a coherent and rational progression of information, directing the reader through the study’s goals, methodologies, findings, and conclusions.

2. METHODS

2.1 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a specialized form of deep neural network devised to analyze structured grid data, such as images. The architecture comprises strata such as convolutional layers, pooling layers, and fully connected layers [16] as shown in Figure 1.

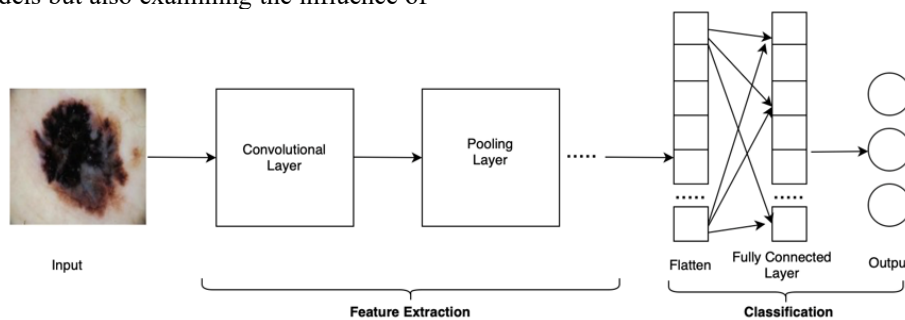


Figure 1. CNNs architecture

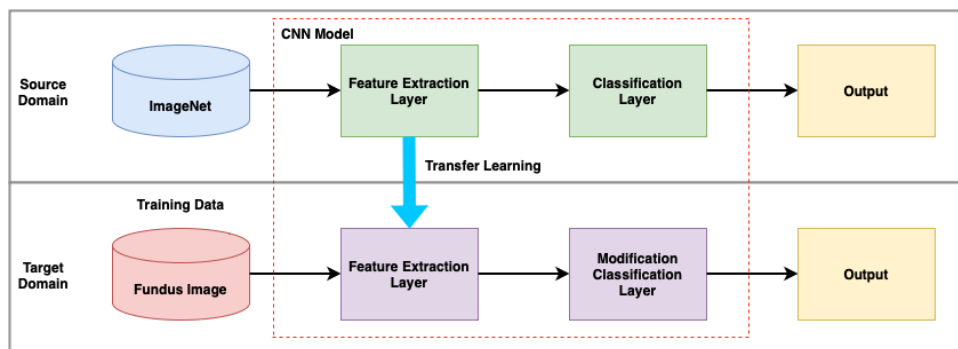


Figure 2. CNNs transfer learning illustration

Convolutional layers employ filters to record spatial hierarchies and patterns, encompassing edges, textures, and forms. Pooling layers decrease the dimensions in the data, improving computing efficiency and reducing the likelihood of overfitting.

Although CNNs are the fundamental structure for analyzing visual data, transfer learning is a method that can be used to modify a pre-trained model for a different purpose. CNNs differs from transfer learning because it does not need extensive datasets or substantial computing resources to train from scratch. Instead, it utilizes an existing model, decreasing training time and data resource requirements. The CNNs are deployed from the beginning to acquire knowledge of patterns from unprocessed data, while transfer learning is applied after a model has already been trained on a similar job.

2.2 Transfer learning

CNNs utilize transfer learning to use pre-existing models for relevant tasks by excluding the final fully connected layer and introducing a new classification layer tailored to the new dataset. The methodology uses convolution and pooling layers while disregarding fully connected and SoftMax layers, deeming them superfluous for pre-existing models [17, 18]. Subsequently, the model undergoes fine-tuning by utilizing the new dataset containing distinct classes [19]. Adjusting weights in transfer learning mirrors that of the initial project. Typically, ImageNet weights are utilized, offering an extensive repository of labelled object images to advance computer vision research. The availability of training data plays a pivotal role in the classification system, with the assessment of accuracy being carried out using validation data.

Transfer learning is beneficial in scenarios where data is scarce, enabling efficient learning even with limited datasets, in contrast to conventional machine learning and CNNs approaches that require substantial datasets to achieve high accuracy [20, 21]. Figure 2 visually represents the application of CNNs and CNNs models incorporating transfer learning methodologies.

In Figure 2, a CNNs is illustrated as it employs the ImageNet dataset to extract features and the classification layers, thereby leading to data categorization into specific classes. In contrast, transfer learning entails the utilization of a targeted dataset, where the feature extraction layer adopts a pre-trained CNNs model from the ImageNet dataset. Consequently, the classification layer is adjusted to align with the categorized dataset. Transfer learning with CNNs involves using pre-trained architectural models, such as those trained on ImageNet, to achieve heightened classification accuracy. This strategy facilitates enhanced performance in classification assignments compared to commencing training from the beginning with smaller datasets.

2.3 MobileNetV2 architecture model

Transfer learning in CNNs pertains to using a pre-trained model, such as MobileNetV2, that has undergone initial training on vast datasets like ImageNet. MobileNetV2 represents a progression of the MobileNet framework, presenting enhancements compared to its precursor. An overview of the MobileNetV2 architecture is presented in Table 1.

Table 1. MobileNetV2 architecture [22]

Input	Operator	<i>t</i>	<i>c</i>	<i>n</i>	<i>s</i>
224 ² ×3	conv2d	-	32	1	2
112 ² ×32	bottleneck	1	16	1	1
112 ² ×16	bottleneck	6	24	2	2
56 ² ×24	bottleneck	6	32	3	2
28 ² ×32	bottleneck	6	64	4	2
14 ² ×64	bottleneck	6	96	3	1
14 ² ×96	bottleneck	6	160	3	2
7 ² ×160	bottleneck	6	320	1	1
7 ² ×320	conv2d 1×1	-	1280	1	1
7 ² ×1280	avgpool 7×7	-	-	1	-
1×1×1280	conv2d 1×1	-	<i>k</i>	-	-

Table 1 utilizes the symbol *n* to represent the act of repeating something *n* times, and the symbol *c* indicates the magnitude of the output channels. The parameter *s* represents the block stride, while *t* is the expansion factor, which increases the number of channels. The input data for this structural design has dimensions of 224×224. Moreover, the system includes a bottleneck component. There are a total of 7 bottleneck blocks. The first bottleneck occurs once, the second bottleneck occurs twice, the third bottleneck occurs thrice, the fourth bottleneck occurs four times, the fifth bottleneck thrice, the sixth bottleneck thrice, and the seventh bottleneck occurs once. MobileNetV2's architecture comprises 17 bottleneck layers [22]. The bottleneck residual block in the MobileNetV2 framework includes an extra layer, as specified in Table 2.

Table 2. Bottleneck MobileNetV2 [22]

Input	Operator	Output
$h \times w \times k$	1×1 conv2d, ReLU6	$h \times w \times (tk)$
$h \times w \times tk$	3×3 dwse $s=s$, ReLU6	$\frac{h}{s} \times \frac{w}{s} \times tk$
$\frac{h}{s} \times \frac{w}{s} \times tk$	linear 1×1 conv2d	$\frac{h}{s} \times \frac{w}{s} \times k'$

Table 2 illustrates the configuration of every bottleneck within the architecture of MobileNetV2. Every bottleneck comprises three principal elements: a convolutional layer with Rectified Linear Unit (ReLU) activation, a depth wise separable convolution with ReLU activation, and a linear convolution layer.

This architecture integrates inverted residuals and linear bottlenecks to improve the propagation of features and decrease the complexity of the model. This architecture enables MobileNetV2 to sustain a high level of accuracy while maintaining a compact model size. Compared to many traditional CNNs, MobileNetV2 is more efficient as it delivers more accuracy with less processing time. This high-efficiency level is especially advantageous for applications that need immediate processing, such as mobile applications and embedded devices [23, 24].

2.4 VGG16 architecture model

Another prominent architectural paradigm employed in CNNs transfer learning is VGG16, developed by the Visual Geometry Group at Oxford University. VGG16 is distinguished by its utilization of 3×3 filters throughout all layers, facilitating the extraction of intricate features at a relatively reduced computational expense. The architectural layout of VGG16 is depicted in Figure 3.

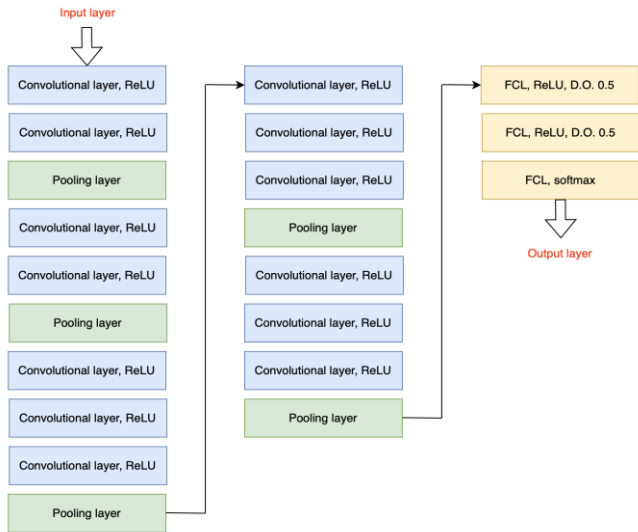


Figure 3. VGG16 architecture

The architectural design’s depth enables it to accurately capture complex patterns and characteristics from images, making it incredibly efficient for jobs such as medical image analysis, where even minor variations can be crucial [25, 26].

The VGG16 model has demonstrated exceptional performance in transfer learning challenges, including fine-tuning pre-trained models for applications. This method’s popularity for different image classification problems stems from its capacity to extract comprehensive feature representations [27, 28].

2.5 VGG19 architecture model

The architectural design of VGG19, an expansion of VGG16 by incorporating three more convolutional layers, therefore increasing the overall number of layers to 19. It upholds the utilization of 3×3 filters throughout its various layers, mirroring the approach taken in VGG16 [29]. The VGG19 architecture is visually depicted in Figure 4.

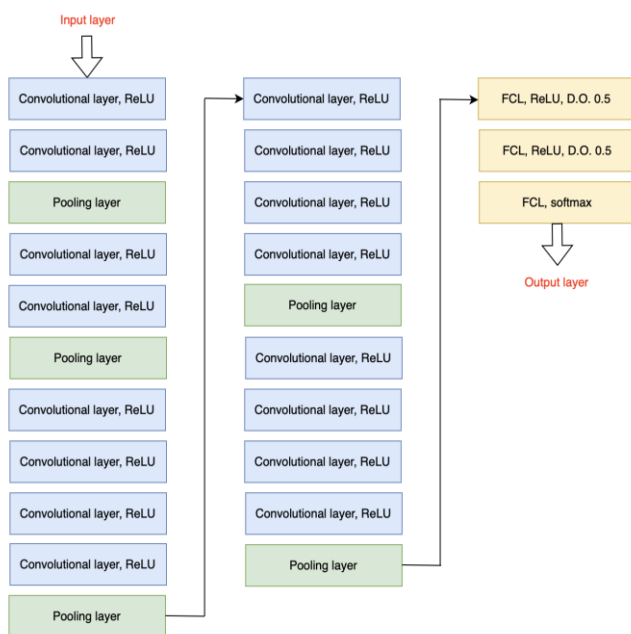


Figure 4. VGG19 architecture

VGG19 is composed of a total of 47 layers, featuring an initial layer with dimensions of 224×224 pixels, 16 convolutional layers, 18 ReLU activation functions, five max-pooling layers, three fully connected layers, two dropout layers, one SoftMax activation function, and a concluding output layer.

In image classification tasks, including more layers in VGG19 enhances accident accuracy compared to other methods. Previous research has demonstrated that VGG19 frequently surpasses VGG16 in several applications, especially fine-grained categorization tasks [25, 27].




Like its predecessor, VGG16, VGG19 is very efficient in transfer learning. The pre-trained weights of this model, trained on extensive datasets such as ImageNet, allow it to apply effectively to novel tasks, enhancing its usefulness for researchers and practitioners [27, 30].

3. RESULTS AND DISCUSSIONS

3.1 Data overview

The images we employed in this investigation encompassed two classifications of cutaneous neoplasms: benign and malignant. These particular images were procured from the International Skin Imaging Collaboration (ISIC) website. Furthermore, images portraying cutaneous neoplasms from male individuals aged between 40 and 80 years were encompassed, both with and without a familial background of cutaneous neoplasms. A total of 600 images were accounted for, comprising 300 images of benign cutaneous neoplasms and 300 images of malignant cutaneous neoplasms. Furthermore, we also used a dataset of non-neoplastic skin that included images of various non-neoplastic skin conditions. The image representing normal skin was acquired from kaggle.com and acknowledged to the user Joydip Paul. Images illustrating skin with atrophic scars, contusions, acne papules, and acne pustules were obtained from kaggle.com and attributed to the username Kukuh Prakoso. The collection of non-neoplastic skin specimens encompassed 300 images, comprising 37 images of normal skin, 57 images of atrophic scarred skin, 57 images of contused skin, 88 images of skin with acne papules, and 61 images of skin with pustules.

Table 3. Variable definition

Variable	Description	Example
Benign	Patients aged 40-60 years old, male, and whether there is a family history of skin tumors.	
Malignant	It contains images of malignant skin tumors or skin cancer.	
Non-tumors	It contains images of non-tumor skin, such as healthy skin, pockmarks, rashes, acne papules, and acne pustules.	

In this research endeavour, the investigator employed three distinct research variables, categorized explicitly as benign, malignant, and non-tumor skin. These research variables have been systematically organized and presented in Table 3.

3.2 Training model

The present investigation utilized a dataset comprising 300 images per category, encompassing benign, malignant, and non-tumor skin lesions, culminating in 900 images. Following this, the dataset was divided into three subsets: 80% for training, 20% for testing, and 10% for the training subset designated for validation purposes. Consequently, were 648 images earmarked for training, 180 for testing, and 72 for validation. The pixel intensities of the images underwent rescaling from a range spanning 0 to 255 to a standardized

interval of 0 to 1. This normalization procedure was implemented to enhance the convergence and robustness of the model throughout the training phase. Furthermore, the images were uniformly resized to 224×224 pixels and converted to the RGB color space to faithfully capture the broad spectrum of colors in skin lesion images. A batch size 64 was employed for the training, testing, and validation collections.

In the CNNs methodology, the transfer learning models MobileNetV2, VGG16, and VGG19 each exhibit unique architectural designs. Throughout the training stage, the input dataset undergoes processing via the feature extraction stratum followed by the classification stratum. The distinct configurations of these feature extraction and classification strata employed in the CNNs methodology are delineated in Figure 5.

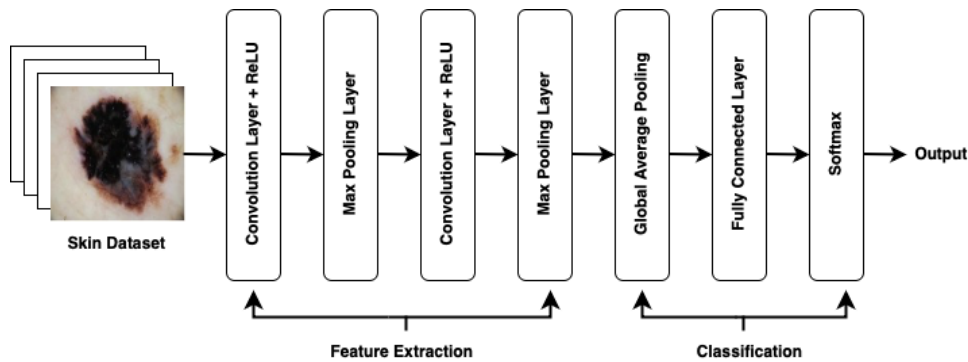


Figure 5. CNNs training model

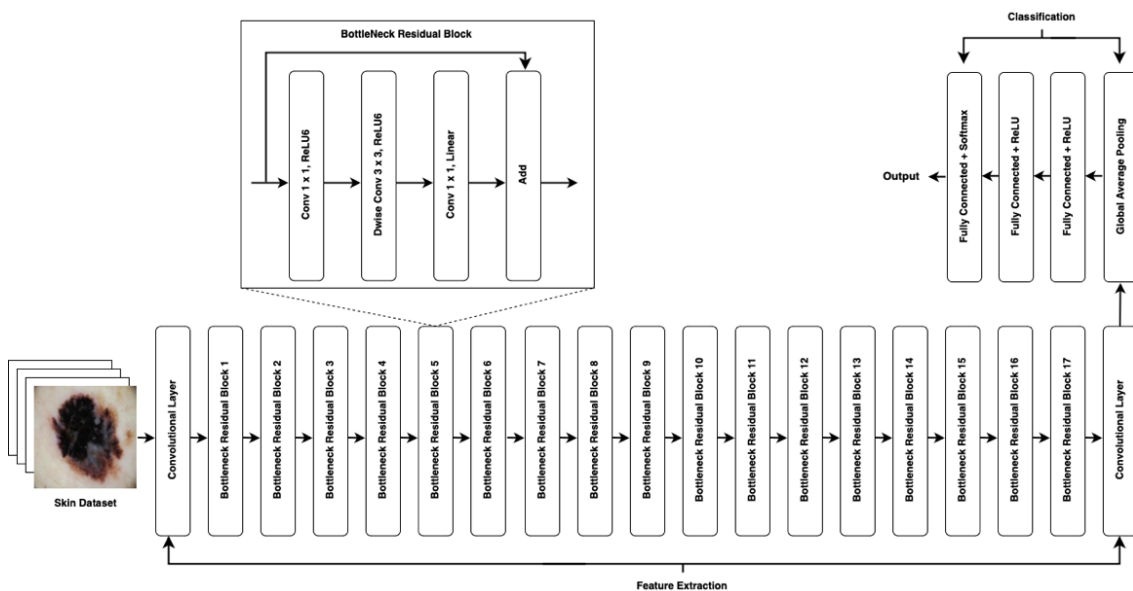


Figure 6. MobileNetV2 training model

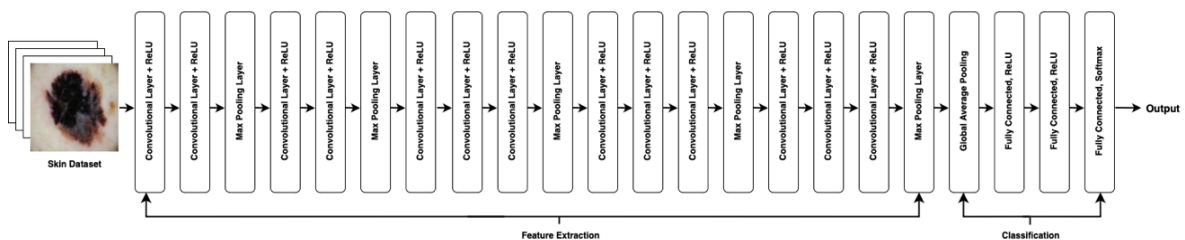


Figure 7. VGG16 training model

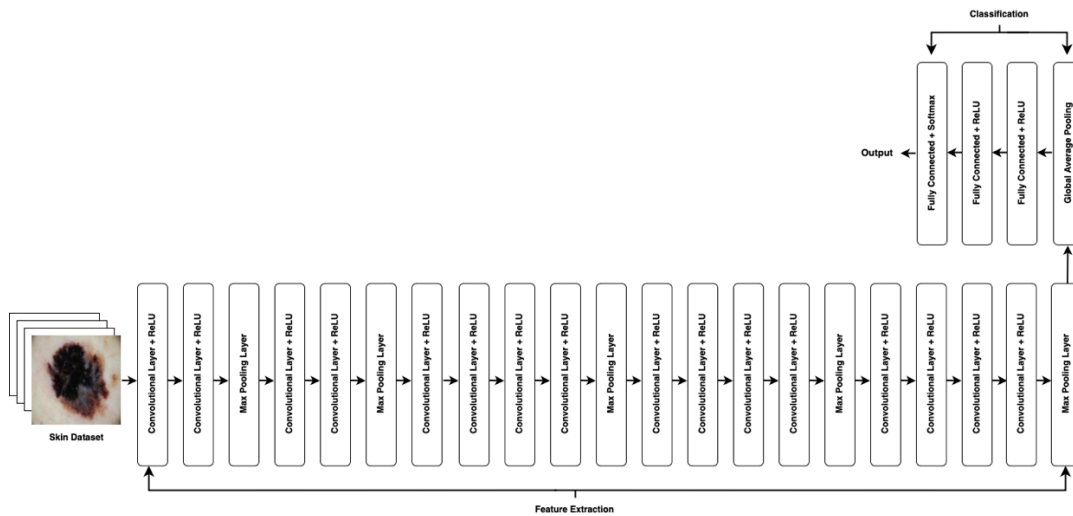


Figure 8. VGG19 training model

The CNNs methodology entails the utilization of multiple pre-determined layers. In this study, we used skin images with dimensions of 224×224 pixels as input and a kernel size 3×3 . Afterward, the input is processed using a convolutional layer that consists of 16 filters, a 3×3 kernel, and the ReLU activation function. This specific layer is designed to extract prominent characteristics from the input images. The next layer in the architecture is a max pooling layer with a 2×2 kernel, which is used to decrease the dimensionality of the image.

After the initial stage, a subsequent convolutional layer is added with 32 filters, a 3×3 kernel, and the ReLU activation function. Another max pooling layer is implemented with a 2×2 kernel sequentially. After two cycles of convolution and pooling, a global average pooling layer is applied to produce a feature matrix with a single average value for each channel. A compact layer with a SoftMax activation function produces probability values, which helps classify images into specific categories such as benign, malignant, and non-tumor skin types.

In CNNs with transfer learning, the MobileNetV2 architecture features layers different from those in the VGG16 and VGG19 models. The specific layers within the feature extraction and classification components of the MobileNetV2 architecture are distinct and are illustrated in Figure 6.

In the CNNs with the transfer learning, the MobileNetV2 architecture model takes as input skin images measuring 224×224 pixels, using a 3×3 kernel. This input is processed through the feature extraction section, which utilizes the MobileNetV2 model. Consequently, the layers in this section follow the structure of MobileNetV2, consisting of 17 bottleneck layers. Modifications to the classification layer were made to accommodate three classes: benign, malignant, and non-tumor skin. The classification layer includes two fully connected (dense) layers with the ReLU activation function and one dense layer with the SoftMax activation function.

In contrast, the VGG16 architecture features different layers from those in MobileNetV2 and other models. The specific layers in the feature extraction and classification sections of the VGG16 model are depicted in Figure 7.

The VGG16 architecture model processes skin images measuring 224×224 pixels with a 3×3 kernel. This input is passed to the feature extraction section, which adopts the VGG16 model. Consequently, the feature extraction layers follow the VGG16 structure, consisting of 13 convolutional

layers and five pooling layers. The classification layer includes two fully connected (dense) layers with the ReLU activation function and one dense layer with the SoftMax activation function.

Similarly, the VGG19 architecture model has layers that closely resemble those in the VGG16 model. The specific layers in the feature extraction and classification sections for the VGG19 model are illustrated in Figure 8.

The VGG19 architecture model analyses skin images with dimensions of 224×224 pixels using a 3×3 kernel. The input is passed to the feature extraction section, which utilizes the VGG19 model. Therefore, the layers in this section adhere to the VGG19 architecture, consisting of 16 convolutional layers and five pooling layers. The classification layer comprises two fully connected (dense) layers utilizing the ReLU and one dense layer employing the SoftMax activation function.

In this study, we utilize the Adam optimizer to adjust the weights and biases in the model, aiming to minimize the loss function during training. The maximum number of epochs is set to 100. After comprehensively evaluating four methods, the MobileNetV2 emerged as the best model. It produced the most stable and smooth accuracy graph across epochs, the highest accuracy value at the end of training, and overall superior model performance, affirming the validity and reliability of our findings.

For each method, it utilizes a different number of parameters during model training. These parameters play a crucial role in enhancing model performance. The research results highlight the impact of the number of parameters on the accuracy and loss values for both the training and validation data at the end of the training period.

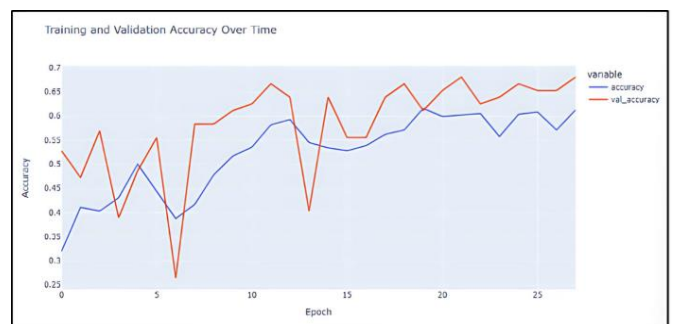


Figure 9. CNNs accuracy

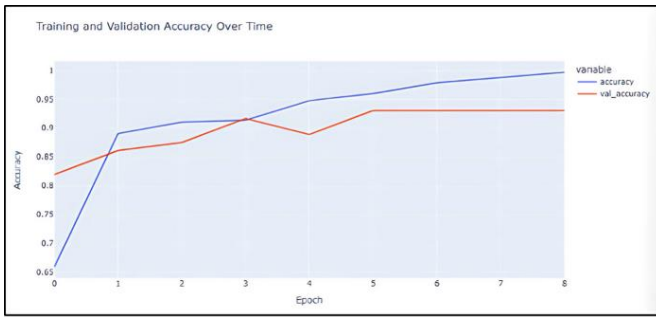


Figure 10. MobileNetV2 accuracy

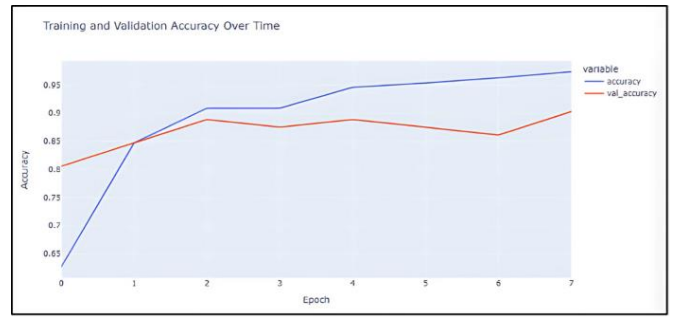


Figure 12. VGG19 accuracy

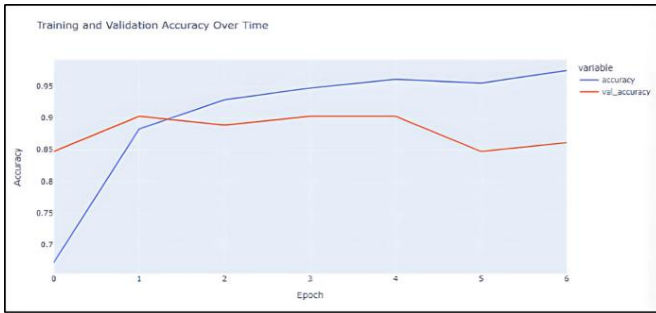


Figure 11. VGG16 accuracy

Figures 9-12 shows that among the four training approaches, the CNN transfer learning approach of the MobileNetV2 architecture model yielded the most stable and smooth graph in each epoch, the highest accuracy value at the end of training, and achieved solid model performance. MobileNetV2 generates a more consistent and seamless accuracy value graph for each epoch compared to alternative approaches. Furthermore, the training and validation data have an accuracy rate of around 95% upon completion. This approach yields a very high level of accuracy. By achieving a rather high accuracy value, the resultant model will exhibit excellent performance in classifying data into suitable and more precise categories.

Table 4. Model accuracy

	CNNs	MobileNetV2	VGG16	VGG19
Number of parameters	5,187	2,438,851	14,797,251	20,106,947
Train accuracy	61.3%	99.7%	97.5%	97.4%
Validation accuracy	68.1%	93.1%	86.1%	90.3%
Train loss	0.9	0.0	0.1	0.1
Validation loss	0.8	0.2	0.3	0.4

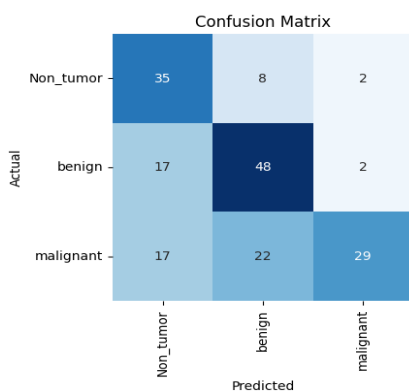


Figure 13. CNNs metric

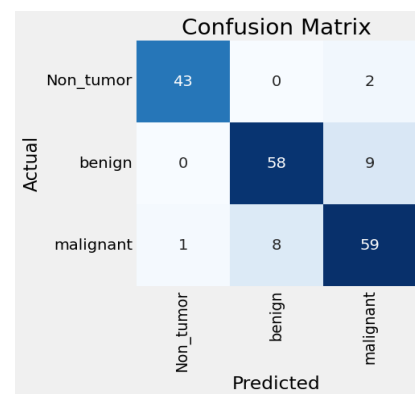


Figure 15. VGG16 metric

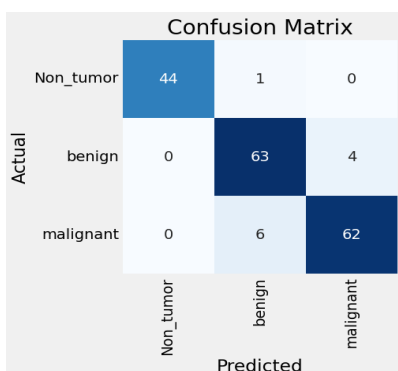


Figure 14. MobileNetV2 metric

The outcomes of model training are presented in Table 4, containing details regarding the total quantity of parameters utilized and their impact on accuracy and loss metrics. The methodologies are arranged in increasing order based on the number of parameters they possess are CNNs, MobileNetV2, VGG16, and VGG19. Conversely, the accuracy outcomes for training and validation datasets post-training, organized in ascending sequence, are CNNs, VGG16, VGG19, and MobileNetV2. The loss values for training and validation, ranked in descending sequence, are as follows: CNNs, VGG16, VGG19, and MobileNetV2. Its superior accuracy and minimal loss distinguish the most suitable model. The CNNs implementing the MobileNetV2 architecture showcased superior accuracy and minimal loss under these circumstances.

		Predicted		
		Non_tumor	benign	malignant
Actual	Non_tumor	44	0	1
	benign	0	63	4
	malignant	0	13	55

Figure 16. VGG19 metric

However, a surprising discovery emerged. Despite MobileNetV2 having fewer parameters, both VGG16 and VGG19 attained inferior accuracy and higher loss metrics. This intriguing result challenges the conventional belief that a higher parameter count always leads to enhanced performance for categorizing benign, malignant, and non-tumor skin images. In fact, a model with a reduced parameter count, like MobileNetV2, may demonstrate greater efficacy and accomplish optimal accuracy and loss metrics, especially for simpler tasks and smaller datasets.

The analysis of the four confusion matrices in Figure 13-16 reveals that the MobileNetV2 architectural model boasts the most significant number of accurately corrected images (True Positive). Among all the class prediction metrics, including accuracy, precision, recall, and F1-score, the MobileNetV2 architecture attains the highest overall accuracy. We use the malignant class's accuracy and recall metrics as model performance indicators. In the context of equal or similar class

sizes, accuracy is selected as a metric to evaluate the overall correctness of a model's predictions. Each class consists of 300 samples. The recall metric assesses the model's accuracy in predicting the positive (malignant) class, a critical factor in reducing false negatives. Accurate prediction of the malignant class is crucial in diagnosing benign, malignant, and non-tumor skin tumors. Misclassifying a malignant tumor as benign or non-tumor could hinder the patient from receiving timely treatment for a dangerous pathology such as skin cancer.

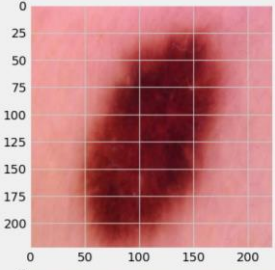
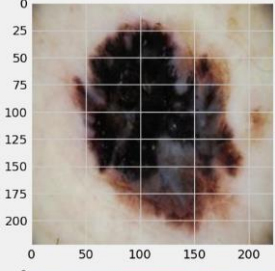
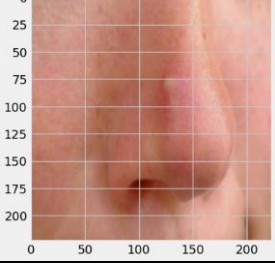
Table 5. Accuracy and loss of test data

	CNNs	MobileNetV2	VGG16	VGG19
Accuracy	62.2%	93.9%	88.9%	90.0%
Loss	0.9	0.2	0.2	0.2

Regarding Table 5, the MobileNetV2 architecture achieves the highest test accuracy and the lowest test loss. It makes the MobileNetV2 model the most effective method for classifying skin images into benign, malignant, and non-tumorous skin. Afterwards, the MobileNetV2 model is utilized to predict skin images from outside the dataset. The MobileNetV2 model incorporates an inverted residual structure equipped with linear bottlenecks. This architecture facilitates the maintenance of a lightweight model while retaining the capability to represent intricate data patterns. Inverting the residuals guarantees that the feature maps in intermediate layers have reduced dimensionality, hence minimizing the required computing operations.

The model's effectiveness enables its implementation on more affordable electronic devices, potentially increasing the accessibility of sophisticated skin lesion diagnosis tools to healthcare providers and patients with limited resources.

Table 6. Probability prediction using MobileNetV2

Image	Benign	Malignant	Non-Tumor	Decision
	90.3%	9.4%	0.3%	Correct (Benign)
	4.6%	95.3%	0.1%	Correct (Malignant)
	0.5%	0.0%	99.5%	Correct (Non-Tumor)

The subsequent phase utilizes the optimized model and architecture to predict skin disease images from external datasets. Table 6 illustrates that the model effectively categorizes each skin illness, attaining an average prediction probability exceeding 90%. This exceptional performance highlights the efficacy and accuracy of the MobileNetV2 model in evaluating intricate image data, establishing it as a dependable instrument for skin disease identification and diagnosis.

4. CONCLUSION

The CNNs strategy manually determined the layers during the feature extraction phase, while transfer learning methods utilized pre-existing layers. The classification layers consisted of flattened or global average pooling layers and fully connected layers customized based on the dataset's classes. The training involved a dataset containing images of benign, malignant, and non-tumor skin lesions, which underwent processing through feature extraction and classification layers to categorize the skin images accurately. The MobileNetV2 architecture demonstrated the highest accuracy, establishing it as the most efficient approach for predicting skin lesion classifications in this investigation.

Moreover, these findings can categorize skin types into benign, malignant, and non-tumorous tumors. These are then compiled into a website for convenient access and utility for many individuals.

In summary, it is crucial to tackle these constraints in future studies. Highlighting that although the present study offers valuable insights into the utilization of MobileNetV2 for skin lesion identification, it is essential to address these obstacles to enhance the clinical usefulness of the model and guarantee its successful implementation in various healthcare settings.

REFERENCES

- [1] Brinker, T.J., Hekler, A., Enk, A.H., et al. (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. <https://doi.org/10.1016/j.ejca.2019.04.001>
- [2] Saravanan, S., Kumar, V.V., Sarveshwaran, V., Indirajithu, A., Elangovan, D., Allayear, S.M. (2022). Computational and mathematical methods in medicine glioma brain tumor detection and classification using convolutional neural network. *Computational and Mathematical Methods in Medicine*, 2022(1): 4380901. <https://doi.org/10.1155/2022/4380901>
- [3] Ferlay, J., Ervik, M., Lam, F., et al. (2020). *Global cancer observatory: Cancer today*. Lyon: International Agency for Research on Cancer, 20182020.
- [4] Lahouaoui, L., Abdelhak, D., Abderrahmane, B., Toufik, M. (2022). Image classification using a fully convolutional neural network CNN. *Mathematical Modelling of Engineering Problems*, 9(3): 771-778. <https://doi.org/10.18280/mmep.090325>
- [5] Bulla, P., Anantha, L., Peram, S. (2020). Deep neural networks with transfer learning model for brain tumors classification. *Traitement du Signal*, 37(4): 593-601. <https://doi.org/10.18280/ts.370407>
- [6] Aslam, N., Khan, I.U., Albahussain, T.I., Almousa, N.F., Alolayan, M.O., Almousa, S.A., Alwhebi, M.E. (2022). MEDeep: A deep learning based model for memotion analysis. *Mathematical Modelling of Engineering Problems*, 9(2): 533-538. <https://doi.org/10.18280/mmep.090232>
- [7] Setiawan, W., Saputra, M.A., Koeshardianto, M., Rulaningtyas, R. (2024). Transfer learning and fine tuning in Modified VGG for haploid diploid corn seed images classification. *Revue d'Intelligence Artificielle*, 38(2): 483-490. <https://doi.org/10.18280/ria.380211>
- [8] Bechelli, S., Delhommelle, J. (2022). Machine learning and deep learning algorithms for skin cancer classification from dermoscopic images. *Bioengineering*, 9(3): 97. <https://doi.org/10.3390/bioengineering9030097>
- [9] Trinh, H., Nguyen, H., Vo, T., Nguyen, D. (2023). Tiny convolution contextual neural network: A lightweight model for skin lesion detection. In *Fifteenth International Conference on Machine Vision (ICMV 2022)*, SPIE, 12701: 246-252. <https://doi.org/10.1117/12.2679699>
- [10] Xu, X. (2023). Research on image features classification based on graph convolutional neural network. In *Sixth International Conference on Computer Information Science and Application Technology (CISAT 2023)*, 12800: 976-981. <https://doi.org/10.1117/12.3004260>
- [11] Taşar, B. (2023). SkinCancerNet: Automated classification of skin lesion using deep transfer learning method. *Traitement du Signal*, 40(1): 285-295. <https://doi.org/10.18280/ts.400128>
- [12] Srinivas, C., KS, N.P., Zakariah, M., Alothaibi, Y.A., Shaukat, K., Partibane, B., Awal, H. (2022). Deep transfer learning approaches in performance analysis of brain tumor classification using MRI images. *Journal of Healthcare Engineering*, 2022(1): 3264367. <https://doi.org/10.1155/2022/3264367>
- [13] Hassan, A.M., El-Mashade, M.B., Aboshosha, A. (2022). Deep learning for cancer tumor classification using transfer learning and feature concatenation. *International Journal of Electrical and Computer Engineering*, 12(6): 6736. <https://doi.org/10.11591/ijece.v12i6.pp6736-6743>
- [14] Alanazi, M.F., Ali, M.U., Hussain, S.J., et al. (2022). Brain tumor/mass classification framework using magnetic-resonance-imaging-based isolated and developed transfer deep-learning model. *Sensors*, 22(1): 372. <https://doi.org/10.3390/s22010372>
- [15] Huynh, B.Q., Li, H., Giger, M.L. (2016). Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. *Journal of Medical Imaging*, 3(3): 034501. <https://doi.org/10.1117/1.JMI.3.3.034501>
- [16] Khan, S., Rahmani, H., Shah, S.A.A., Bennamoun, M., Medioni, G., Dickinson, S. (2018). A guide to convolutional neural networks for computer vision. *Synthesis Lectures on Computer Vision*, 8(1): 1-207. <https://doi.org/10.1007/978-3-031-01821-3>
- [17] Gatys, L., Ecker, A., Bethge, M. (2016). A neural algorithm of artistic style. *Journal Vis*, 16(12): 326. <https://doi.org/10.1167/16.12.326>
- [18] Mordvintsev, A., Olah, C., Tyka, M. (2015). Inceptionism: Going deeper into neural networks. *Google Research Blog*, 20(14): 5. <https://api.semanticscholar.org/CorpusID:69951972>
- [19] Liu, S., Tian, G. (2019). An indoor scene classification method for service robot based on CNN feature. *Journal of Robotics*, 2019(1): 8591035.

- <https://doi.org/10.1155/2019/8591035>
- [20] Efremova, D.B., Sankupellay, M., Konovalov, D.A. (2019). Data-efficient classification of birdcall through convolutional neural networks transfer learning. In 2019 Digital Image Computing: Techniques and Applications (DICTA), Perth, WA, Australia, pp. 1-8. <https://doi.org/10.1109/DICTA47822.2019.8946016>
- [21] Lei, Q., Hu, W., Lee, J. (2021). Near-optimal linear regression under distribution shift. In International Conference on Machine Learning, pp. 6164-6174.
- [22] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, City, UT, USA, pp. 4510-4520. <https://doi.org/10.1109/CVPR.2018.00474>
- [23] Togatorop, P.R., Pratama, Y., Monica Sianturi, A., Sari Pasaribu, M., Sangmajadi Sinaga, P. (2023). Image preprocessing and hyperparameter optimization on pretrained model MobileNetV2 in white blood cell image classification. IAES International Journal of Artificial Intelligence, 12(3): 1210. <https://doi.org/10.11591/ijai.v12.i3.pp1210-1223>
- [24] Ye, M., Zhang, J. (2023). Mobip: A lightweight model for driving perception using MobileNet. Frontiers in Neurorobotics, 17: 1291875. <https://doi.org/10.3389/fnbot.2023.1291875>
- [25] Devasia, J., Goswami, H., Lakshminarayanan, S., Rajaram, M., and Adithan, S. (2023). Deep learning classification of active tuberculosis lung zones wise manifestations using chest X-rays: A multi-label approach. Scientific Report, 13(1): 887. <https://doi.org/10.1038/s41598-023-28079-0>
- [26] El Asnaoui, K., Chawki, Y. (2021). Using X-ray images and deep learning for automated detection of coronavirus disease. Journal of Biomolecular Structure and Dynamics, 39(10): 3615-3626. <https://doi.org/10.1080/07391102.2020.1767212>
- [27] Yeoh, P.S.Q., Lai, K.W., Goh, S.L., Hasikin, K., Wu, X., Li, P. (2023). Transfer learning-assisted 3D deep learning models for knee osteoarthritis detection: Data from the osteoarthritis initiative. Front Bioeng Biotechnology, 11. <https://doi.org/10.3389/fbioe.2023.1164655>
- [28] Ilhan, H. O., Serbes, G., and Aydin, N. (2022). Decision and feature level fusion of deep features extracted from public COVID-19 data-sets. Applied Intelligence, 52(8): 8551-8571. <https://doi.org/10.1007/s10489-021-02945-8>
- [29] Chowdary, G.J., Yogarajah, P. (2023). Nucleus segmentation and classification using residual SE-UNet and feature concatenation approach incervical cytopathology cell images. Technology in Cancer Research & Treatment, 22. <https://doi.org/10.1177/15330338221134833>
- [30] Karac, A. (2023). Predicting COVID-19 Cases on a large chest X-Ray dataset using modified pre-trained CNN architectures. Applied Computer Systems, 28(1): 44-57. <https://doi.org/10.2478/acss-2023-0005>