JOIV : Int. J. Inform. Visualization, 8(2) - May 2024 606-612

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

Classification of Dermoscopic Images Using CNN-SVM

Agus Eko Minarno [a,*], Muhammad Fadhlan [a], Yuda Munarko [a,b], Didih Rizki Chandranegara [a]

[a] *Department of Information Technology, Universitas Muhammadiyah Malang, Jl. Raya Tlogomas 246, Malang, 65144, Indonesia*
[b] *Auckland Bioengineering Institute, University of Auckland, Auckland, 1010, New Zealand*

Corresponding authors: [*]aguseko@umm.ac.id

***Abstract*— Traditional machine learning methods like GLCM and ABCD rules have long been employed for image classification tasks. However, they come with inherent limitations, primarily the need for manual feature extraction. This manual feature extraction process is time-consuming and relies on expert domain knowledge, making it challenging for non-experts to use effectively. Deep learning methods, specifically Convolutional Neural Networks (CNN), have revolutionized image classification by automating the feature extraction. CNNs can learn hierarchical features directly from the raw pixel values, eliminating the need for manual feature engineering. Despite their powerful capabilities, CNNs have limitations, mainly when working with small image datasets. They may overfit the data or struggle to generalize effectively. In light of these considerations, this study adopts a hybrid approach that leverages the strengths of both deep learning and traditional machine learning. CNNs are automatic feature extractors, allowing the model to capture meaningful image patterns. These extracted features are then fed into a Support Vector Machine (SVM) classifier, known for its efficiency and effectiveness in handling small datasets. The results of this study are encouraging, with an accuracy of 0.94 and an AUC score of 0.94. Notably, these metrics outperform Abbas' previous research by a significant margin, underscoring the effectiveness of the hybrid CNN-SVM approach. This research reinforces that SVM classifiers are well-suited for tasks involving limited image data, yielding improved classification accuracy and highlighting the potential for broader applications in image analysis.**

***Keywords*— Convolutional neural network; feature extraction; classification; support vector machine; acral melanoma.**

Manuscript received 27 Sep. 2023; revised 29 Dec. 2023; accepted 6 Feb. 2024. Date of publication 31 May. 2024.

I. Introduction

Acral Melanoma (AM) is a melanoma that appears on human skin; it is one of the four main histological subtypes [1] which is a common disease in patients from Asia, Africa, Hispanic (Latin America) regions, accounting for 2-3% of all melanoma cases [1][2]. Acral melanoma usually appears on the palms, soles, and near the nail area [1][2]. Acral Melanoma is still little studied and researched, even though this subtype of acral melanoma has a worse, more difficult prognosis (predictions about acral melanoma) and has a high mortality rate [2]. In addition, acral melanoma is challenging to diagnose at an early stage because many dermatological conditions have to be considered. In contrast, when it can be detected early, it can be cured by surgical excision of the lesion [1]. So, early detection and classification are needed for acral melanoma.

Several studies have been conducted regarding the classification of dermoscopic images [3][15]. First, a study by Yu et al. [16] used a modified 16-layer VGG-16 fine-tuned model on 724 dermoscopic images divided into 350 acral melanoma images from 81 patients and 374 benign nevi images from 194 patients. The data was tested by dividing it into two groups in this study. Group A, with 175 acral melanoma images and 187 benign nevi images, was tested in group B with the same amount of data, and vice versa. The accuracy obtained was 83.51% (Group A) and 80.23% (Group B). The research by [17] used the CNN method with eight 2D Convolutional Layers with the convolutional layer filters 16, 16, 32, 32, 64, 64, 128, 128 with the first input layer of (20,20). The pooling used is average pooling and max pooling. There are two fully connected layers with the number of units, respectively 256 and 128, using a dropout of 0.2. The dataset used is the ISIC 2018 dataset, with a total of 10015 images measuring 450 x 600 pixels. Then, the dataset was grouped into only two classes: the malignant class with 841 pictures and the benign class with 1356 images. The results obtained in this study were an accuracy of 75%. Then, the precision, recall, and f1 scores are 0.80, 0.82, and 0.81, respectively.

Subsequent research was conducted by Abbas et al. [18] using the 7-layer Deep CNN method with 724 dermoscopic

images divided into 350 acral melanoma image data and 374 benign nevi image data. Data is divided into 70% for training, 20% for validation, and 10% for testing. This study obtained an accuracy of 91%. Another study was conducted by Brinker et al. [19] using ResNet50 CNN with training data consisting of 1888 melanomas and 10490 nevi, validation data comprised of 210 melanomas and 1049 nevi, and testing data containing 20 melanomas and 80 nevi. In that study, the sensitivity was 74.1%, and the specificity was 86.5%.

Then, a study by Murugan et al. [20] uses shape, GLCM, and ABCD Rule to extract features and then uses KNN, Random Forest, and SVM for the classifier. Accuracy with the ABCD Rule-SVM Classifier method is the highest, 89.43%. Traditional machine learning methods, such as GLCM, ABCD rules, and so on, have limitations, requiring manual feature extraction [21], but manual feature extraction is difficult and time-consuming [20]. The deep learning (CNN) method can automatically extract features in an image [22]. Even though CNN can automatically extract features, it is terrible at training on small image data [16]. It performs better on many image data [22], so it can cause bias, decreased accuracy, and errors in classification.

A support Vector Machine (SVM) is a classifier with advantages over accuracy and robustness even if the training sample shows a bias [20], [23]. It has also proven to be an effective method for classification problems because it adheres to the structure principle of minimizing risk to maximize generalizability [24]. Besides having robustness, a SVM can also handle overfitting training data [25]. Then, SVM is one of the good choices for small datasets [26]. Therefore, in this study, a combination of CNN automatic feature extraction and a Support Vector Machine classifier will be used to deal with slight, biased, and overfitting classification problems in dermoscopic image data so that higher accuracy will be obtained.

II. Material and Method

A. Research Stages

This study proposes the CNN method for feature extraction and SVM as a classifier. Several stages are carried out in this research. The first stage is the collection of dermoscopic acral melanoma image data. The second stage is preprocessing the image. The third stage is feature extraction from the CNN model. The fourth stage is classification using the SVM Classifier from the features obtained in the third or feature extraction stage. The final stage is evaluation using the confusion matrix, classification report, and ROC Curve (AUC Score).

B. Work Environment

The work environment in this study is the Python programming language version 3.9 with Jupyter Notebook. Then, a laptop with Intel Core i7-9750H CPU @ 2.60GHz (12 CPUs), 16384 MB RAM, and NVIDIA GeForce GTX 1650 GPU.

C. Dataset

The dataset used in this study comes from [3] in the form of dermoscopic images collected from several hospitals around Yonsei University Health System, Seoul, Korea. The total number of images is 724, with details of 350 images of acral melanoma and 374 images of benign nevi. The amount of data between classes can be considered balanced. All image data has a size of 2560 x 1920 pixels. Then, in this study, the data set was divided into three ratios for the three initial scenarios: 90% train data & 10% test data, 80% train data & 20% test data, 70% train data & 30% test data. Fig. 1 and Fig. 2 are examples of acral melanoma and benign nevi.

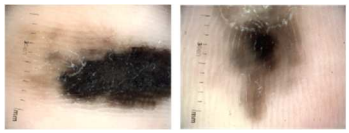

Fig. 1 Examples of acral melanoma

Fig. 2 Examples of benign nevi

D. Preprocessing

At this stage, the step is to resize the image data to 224 x 224 pixels. Then, remove the ruler markings and hairs from the image data. Then, normalize the image data so that the scale is the same and each class is labeled in the image data.

E. Feature Extraction with CNN

CNN works like a regular neural network, but in CNN, there is a convolution layer, a two-dimensional filter, or higher convoluted with the input layer. CNN learns and filters patterns from an image so that CNN gets features in the form of feature maps [27], [28], [29].

At this stage, CNN performs feature extraction to obtain features that will later be useful for classification with the SVM Classifier. The CNN model design for this stage is a 2D Convolutional Layer with sequential filters 32, 64, 128, 256, and 512. Kernel size, strides, and the activation function used are the same for each 2D Convolutional Layer, namely 3, 1, and ReLU. MaxPooling2D with pooling size (2,2) will pass the 2D convolutional layer. Then, flatten it and give a dropout of 0.5. Fig. 3 is the CNN model architecture for feature extraction in this study.

F. SVM Classifier and Grid Search

After obtaining the image features from the feature extraction stage, these features are then subjected to training or classification using the SVM Classifier. SVM has quite a convincing performance in predicting or classifying a class [30]. This SVM can handle high-dimensional feature spaces [31], making it suitable for this study. The SVM Classifier uses several parameters: kernel, C, and gamma. Then, use Grid Search to find the best parameter estimator. Table I shows the parameters used to find the best parameters.

TABLE I
COMPARISON PARAMETERS IN GRID SEARCH SVM

Parameter	Values
Kernel	linear, poly, rbf, sigmoid
C	0.1, 1, 10, 100,
Gamma	0.1, 0.01, 0.001, 0.0001

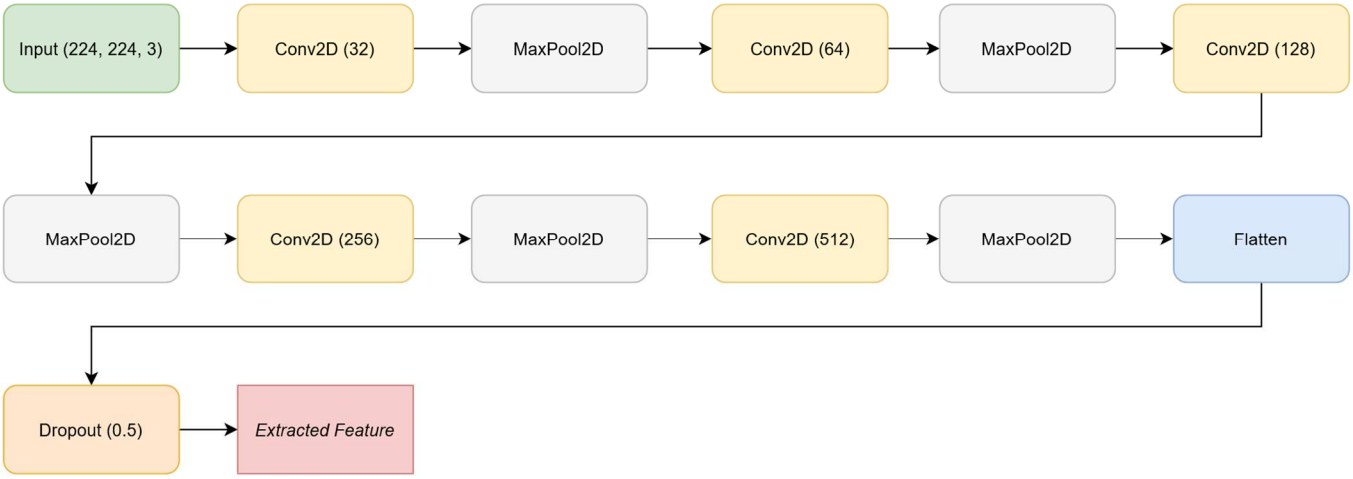

Fig. 3 Architecture of CNN for feature extraction

G. Evaluation

The model's evaluation uses a confusion matrix, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The classification report contains the accuracy, precision, recall, and f1 score values. Finally, the ROC curve with the AUC score is included.

H. Test Scenario

The tests carried out in this study used six different scenarios based on differences in dataset comparisons. Also, a grid search that compared several parameters (Kernel, C, Gamma) to the classification results in accuracy, precision, recall, f1 score, and ROC curve with AUC score. Table II shows the details of the test scenarios in this study.

In the first scenario, testing is carried out by dividing the dataset by 90% training data and 10% testing data for the model proposed in Fig. 4 with the best parameters of the parameters in Table I. Then, in the second scenario, testing is carried out by dividing a dataset of 80% training data and 20% testing data for the model proposed in Fig. 4 with the best parameters in Table I. Then, in the third scenario, testing is carried out by dividing the dataset by 70% training data and 30% data testing of the model proposed in Fig. 4 with the best parameters in Table I. Furthermore, in the fourth scenario, testing is carried out with the pre-trained model VGG-19, which divides the dataset based on the results of the previous best scenario. Likewise, the parameters used for the SVM classifier are the best parameters of the parameters in Table I. Finally, in scenarios 5 and 6, the proposed model is tested with the best dataset comparison performance results in scenarios 1 to 3 and added with hidden layers with a different number of units, namely 2048 for scenario 5 and 512 for scenario 6.

III. Result and Discussion

After the dataset is obtained, the preprocessing images are carried out, as explained in the research methodology. The results are shown in Fig. 5 and Fig. 6. The preprocessed image dataset is separated into train data and test data, and the separation is based on the test scenarios mentioned. The results of this study are based on six (6) test scenarios that have been tested. In addition, in six (6) test scenarios, a deeper search has been carried out on the parameter values specified above with the help of a grid search to obtain the best parameter in each scenario. After that, the classification results in accuracy, precision, recall, f1 score, and AUC score-ROC curve are compared with the results of previous reference studies.

TABLE II
Test Scenario Details

Scenario	Details
1	The Proposed model with a 90:10 dataset
2	The Proposed model with an 80:20 dataset
3	The Proposed model with a 70:30 dataset
4	The VGG-19 Pretrained Model with the best dataset ratio
5	The proposed model with 2048 hidden layer units
6	The proposed model with 512 hidden layer units

A. Results of the Splitting Dataset

After preprocessing the data, the dataset is separated into several ratios for the three initial scenarios. The data sets were separated with several ratios, namely 90:10, 80:20, and 70:30.

B. Scenario 1: The proposed model with a 90:10 dataset

In the first scenario, the test is carried out by dividing the dataset with 90% training data and 10% testing data for the model proposed, with the best parameters obtained being C of 0.1, gamma of 0.1, and the kernel linear. Fig. 4 shows the results of the connection matrix and the AUC Score-ROC curve from the first scenario.

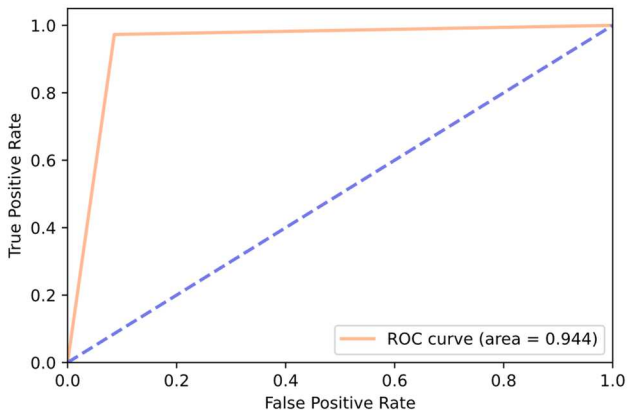

Fig. 4 Result of AUC-ROC score of scenario 1

C. *Scenario 2: The Proposed Model with an 80:20 Dataset*

In the second scenario, the test is carried out by dividing the dataset with 80% training data and 20% testing data for the model proposed, with the best parameters obtained being C of 0.01, gamma of 0.1, and the kernel linear. Fig. 5 shows the results of the connection matrix and the AUC Score-ROC curve from the second scenario.

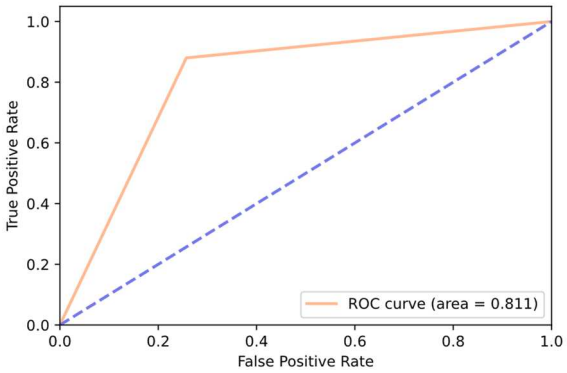

Fig. 5 Result of AUC-ROC score of scenario 2

D. *Scenario 3: The Proposed Model with a 70:30 Dataset*

In the third scenario, the test is carried out by dividing the dataset with 70% training data and 30% testing data for the model proposed, with the best parameters obtained being C of 1000, gamma of 0.00001, and the kernel RBF. Fig. 6 shows the results of the connection matrix and the AUC Score-ROC curve from the third scenario.

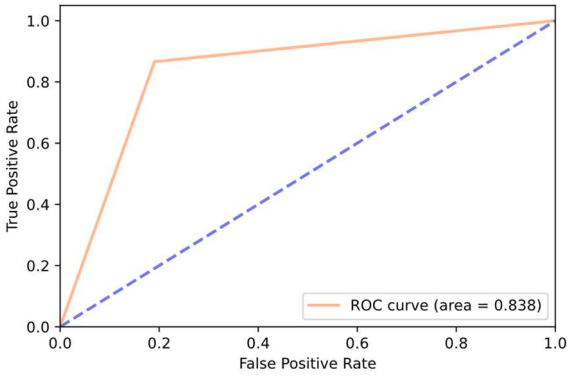

Fig. 6 Result of AUC-ROC score of scenario 3

E. *Scenario 4: VGG-19 Pre-trained Model with Best Dataset Ratio*

In this fourth scenario, testing was carried out with the pre-trained model VGG-19. Then, the distribution of the dataset based on the results of the previous best scenario, namely scenario 1, is 90:10. By fine-tuning the hyperparameters, the optimal combination was identified as a regularization parameter (C) of 10, a gamma value of 0.001, and an RBF kernel. These parameter settings played a crucial role in achieving the best performance by balancing model complexity and generalization capability. Fig. 7 shows the results of the connection matrix and the AUC Score-ROC curve from the fourth scenario.

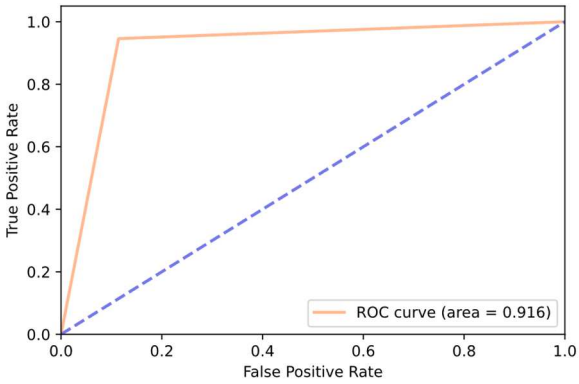

Fig. 7 Result of AUC-ROC score of scenario 4

F. *Scenario 5: The Proposed Model with 2048 Hidden Layer Units*

In this fifth scenario, testing is carried out with the proposed model, but a hidden layer with 2048 units is added. Then, the distribution of the dataset is based on the results of the previous best scenario, namely scenario 1, with a portion of 90:10. The best parameters obtained are C of 0.1, gamma of 0.1, and the kernel is linear. Fig. 8 shows the results of the connection matrix and the AUC Score-ROC curve from the fifth scenario.

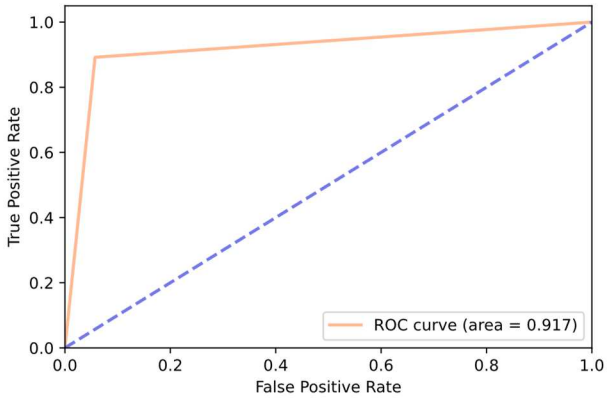

Fig. 8 Result of AUC-ROC score of scenario 5

G. *Scenario 6: The proposed Model with 512 Hidden Layer Units*

In this sixth scenario, an advanced approach was employed to improve the performance of the proposed model further. Building upon the previous best scenario, scenario 1, additional experimentation was conducted by introducing a hidden layer with 512 units. This adjustment enhanced the model's capacity to capture complex patterns and relationships within the dataset. To ensure the reliability of the evaluation, the dataset distribution for this scenario was based on the outcomes of scenario 1, maintaining a partition ratio of 90:10 for training and testing, respectively. By fine-tuning the hyperparameters, the best combination was determined to be a regularization parameter (C) of 0.01, a gamma value of 0.1, and a linear kernel. These parameter settings were instrumental in achieving optimal performance and minimizing potential overfitting or underfitting issues. The outcomes of the sixth scenario are visualized in Figure 9, which presents the connection matrix and the AUC Score-ROC curve.

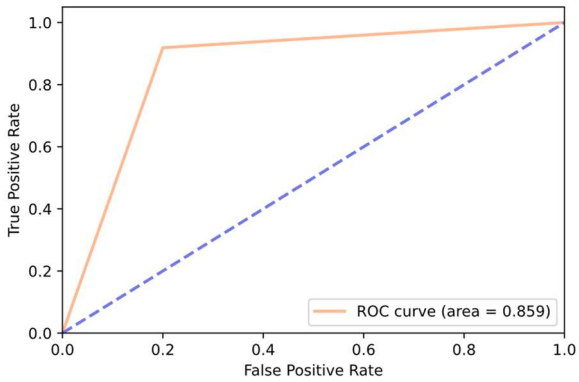

Fig. 9 Result of AUC-ROC score of scenario 6

H. *Analysis of Results*

After testing several scenarios, there appear to be differences in the results from the classification report, confusion matrix, and ROC curve (AUC score). The first scenario is the best because it gets the best results compared to other scenarios in terms of accuracy, precision, recall, F1-Score, and AUC scores of 0.94, 0.95, 0.94, 0.94, and 0.94, respectively. Table III shows details of the evaluation of the results of all scenarios. The highest dataset division, 90% training data, and 10% testing data get the best results in scenario 1. In scenario 2, when training data is reduced by 10% to 80% and data testing is increased by 10% to 20%, the results decrease both from accuracy, precision, recall, F1-Score, and AUC scores to 0.81, 0.82, 0.81, 0.81, and 0.811, respectively.

Furthermore, in scenario 3, when the training data is reduced by 20% from 90% to 70% and the testing data is increased by 20% from 10% to 30%, the results increase compared to scenario 2 in terms of accuracy, precision, recall, F1—Score, and AUC scores 0.84, 0.84, 0.84, 0.84, and 0.83, respectively. However, when compared to scenario one, the results are still decreasing.

In scenario 4, the VGG-19 pre-trained model gets good results from accuracy, precision, recall, F1-Score, and AUC scores, which are 0.92, 0.92, 0.92, 0.92, and 0.91, respectively. However, the results are still under scenario 1. Furthermore, Scenario Five and Scenario Six are tested using the same scenario as Scenario One. The portion used is 90:10. Then, it is added to the hidden layer with a different number of units to see whether the number of features will affect the SVM classifier and its accuracy. In scenario five, the 2048 hidden layer units obtained accuracy, precision, recall, and F1-Score to AUC scores of 0.92, 0.92, 0.92, 0.92, and 0.91, respectively. The different number of hidden layer units (2048) lowers the performance results compared to scenario one.

Scenario six was tested by reducing the number of hidden layer units to 512. The accuracy, precision, recall, and F1-Score results to AUC scores were 0.86, 0.87, 0.86, 0.86, and 0.85, respectively. The performance results decrease when the number of hidden layer units decreases. From this, the division of the portion of the dataset, both training data and testing data, can affect the accuracy, precision, recall, F1-Score, and AUC score obtained from model testing.

However, the more significant the training portion compared to the testing portion, the better the results obtained. It is seen in scenarios 1, 2, and 3. Scenario 1, with a portion of 90:10, gets the best results, but if the portion is changed to 80:20, where the training data is less than scenario 1, the results are the lowest. Then, if the portion of the training data is lowered again, like in scenario 3 to 70:30, the results obtained increase compared to scenario 2. In addition, using the VGG-19 pre-trained model also shows good results because previous training has been carried out with the ImageNet dataset. Nevertheless, the results are not better or still below the first scenario.

Finally, the number of features trained on the SVM classifier affects accuracy, precision, recall, F1-Score, and AUC score results. This is evidenced in scenario one, which has 12800 features with an accuracy of 0.94. In contrast, scenario five, which has 2048 features, gets a lower accuracy of 0.92, and scenario six, which has 512 features, gets an even lower accuracy of 0.86. However, in scenario 4, the number of features is 25088, and the performance results are not as great as scenario 1 or scenario 5, even though the performance results are almost similar to scenario 5. The SVM classifier may search for an optimal number of features to get the best performance results.

TABLE III
RESULTS OF ALL PROPOSED SCENARIOS

Model	Acc	Macro Avg			AUC
		Precision	Recall	F1-Score	
Scenario 1	**0.94**	**0.95**	**0.94**	**0.94**	**0.94**
Scenario 2	0.81	0.82	0.81	0.81	0.81
Scenario 3	0.84	0.84	0.84	0.84	0.83
Scenario 4	0.92	0.92	0.92	0.92	0.91
Scenario 5	0.92	0.92	0.92	0.92	0.91
Scenario 6	0.86	0.87	0.86	0.86	0.85

I. *Comparison with prior research*

This sub-chapter summarizes several previous studies used as a reference for this research. This summary contains several things: the author, the dataset used, the data distribution, the method used, and the results obtained. Table IV shows details regarding the summary comparison of this study with previous studies. This study uses the same dataset as [16][18]but differs from the distribution of the portion of the dataset. Research [16] uses training data and data testing only, while research [18] uses training data, data validation, and data testing.

The results of research by Yu et al. obtained an accuracy of 0.83 and 0.80. Then, AUC scores of 0.84 and 0.81 with the same number of datasets (without augmentation). However, the dataset distribution proportion differs from that of this study, where the method used was VGG-16. Meanwhile, the research by Abbas et al. obtained an accuracy of 0.91 and an AUC score of 0.90 with an augmented image dataset of 4344 image data, 2100 acral melanoma images, and 2244 benign nevi images using the Deep ConvNet method.

The results of the accuracy and AUC of the three research scenarios, namely scenarios 1, 4, and 5 with the proposed CNN-SVM method, are greater than those of research from [16], [18]. Based on performance comparisons with the same dataset. It can be said that the CNN method as feature extraction and SVM as a classifier has succeeded in outperforming the results of a study conducted by [16], [18] with the same dataset but different methods and portions of the data distribution.

TABLE IV
SUMMARY OF COMPARISON RESULTS

Author	Data Distribution	Method	Results	
			AUC	Acc
[16]	Group A (362 images) dan Group B (362 images).	VGG-16	Group A (0.84) and Group B (0.81)	Group A (0.83) and Group B (0.80)
[18]	Augmented images 4344 images, 2100 acral melanoma, and 2244 benign nevi, 70% training, 20% validation, and 10% testing.	Deep ConvNet	0.90	0.91
Proposed	652 images of acral melanoma and 72 images of benign nevi (90:10)	CNN-SVM	**0.94**	**0.94**

The results of this research using the CNN-SVM method are better than those of previous studies with the same dataset. This study obtained an accuracy of 0.94 or 94%. Then, an AUC score of 0.94 compared to prior research, namely Abbas et al., has an accuracy of 0.91 and an AUC score of 0.90. It is also higher than the research by Yu et al., which has an accuracy of 0.83 (group A) and 0.80 (group B). Then, an AUC score of 0.84 or 84% (group A) and 0.81 or 81% (group B). The number of features trained on the SVM classifier and the distribution of dataset portions affect the results of accuracy and AUC scores. Finally, the model with scenarios 1, 4, and 5 achieved high accuracy and outperformed previous studies.

IV. CONCLUSION

This paper presents the classification of dermoscopic images using CNN and SVM. The number of features trained on the SVM classifier and the distribution of dataset portions affect the results of accuracy and AUC scores. Finally, the model with scenarios 1, 4, and 5 achieved high accuracy and outperformed previous studies. This research still needs further improvement in the image preprocessing section because the difference between acral melanoma and benign nevi is less visible. It can be researched and focused on preprocessing images with more specific and better methods so that models more easily process images.

REFERENCES

[1] P. Basurto-Lozada et al., "Acral lentiginous melanoma: Basic facts, biological characteristics and research perspectives of an understudied disease," Pigment Cell & Melanoma Research, vol. 34, no. 1, pp. 59–71, Jun. 2020, doi: 10.1111/pcmr.12885.

[2] Y. A. Chen et al., "Translational pathology, genomics and the development of systemic therapies for acral melanoma," Seminars in Cancer Biology, vol. 61, pp. 149–157, Apr. 2020, doi:10.1016/j.semcancer.2019.10.017.

[3] A. Kalaivani and S. Karpagavalli, "A Deep Ensemble Model for Automated Multiclass Classification Using Dermoscopy Images," 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), Mar. 2022, doi:10.1109/iccmc53470.2022.9753708.

[4] X. He, Y. Wang, S. Zhao, and C. Yao, "Deep metric attention learning for skin lesion classification in dermoscopy images," Complex & Intelligent Systems, vol. 8, no. 2, pp. 1487–1504, Jan. 2022, doi:10.1007/s40747-021-00587-4.

[5] Y. Wang et al., "A comparative study of melanocytic nevi classification with dermoscopy and high-frequency ultrasound," Skin Research and Technology, vol. 28, no. 2, pp. 265–273, Dec. 2021, doi:10.1111/srt.13123.

[6] T. Vaiyapuri, P. Balaji, Shridevi. S, H. Alaskar, and Z. Sbai, "Computational Intelligence-Based Melanoma Detection and Classification Using Dermoscopic Images," Computational Intelligence and Neuroscience, vol. 2022, pp. 1–12, May 2022, doi:10.1155/2022/2370190.

[7] S. Ayas, "Multiclass skin lesion classification in dermoscopic images using swin transformer model," Neural Computing and Applications, vol. 35, no. 9, pp. 6713–6722, Nov. 2022, doi: 10.1007/s00521-022-08053-z.

[8] F. Grignaffini et al., "Machine Learning Approaches for Skin Cancer Classification from Dermoscopic Images: A Systematic Review," Algorithms, vol. 15, no. 11, p. 438, Nov. 2022, doi:10.3390/a15110438.

[9] B. Shetty, R. Fernandes, A. P. Rodrigues, R. Chengoden, S. Bhattacharya, and K. Lakshmanna, "Skin lesion classification of dermoscopic images using machine learning and convolutional neural network," Scientific Reports, vol. 12, no. 1, Oct. 2022, doi:10.1038/s41598-022-22644-9.

[10] S. Bechelli and J. Delhommelle, "Machine Learning and Deep Learning Algorithms for Skin Cancer Classification from Dermoscopic Images," Bioengineering, vol. 9, no. 3, p. 97, Feb. 2022, doi: 10.3390/bioengineering9030097.

[11] E. Somfai et al., "Handling dataset dependence with model ensembles for skin lesion classification from dermoscopic and clinical images," International Journal of Imaging Systems and Technology, vol. 33, no. 2, pp. 556–571, Nov. 2022, doi: 10.1002/ima.22827.

[12] R. O. Ogundokun et al., "Enhancing Skin Cancer Detection and Classification in Dermoscopic Images through Concatenated MobileNetV2 and Xception Models," Bioengineering, vol. 10, no. 8, p. 979, Aug. 2023, doi: 10.3390/bioengineering10080979.

[13] Y. Nie, P. Sommella, M. Carratù, M. O'Nils, and J. Lundgren, "A Deep CNN Transformer Hybrid Model for Skin Lesion Classification of Dermoscopic Images Using Focal Loss," Diagnostics, vol. 13, no. 1, p. 72, Dec. 2022, doi: 10.3390/diagnostics13010072.

[14] Y. Wang, Y. Wang, J. Cai, T. K. Lee, C. Miao, and Z. J. Wang, "SSD-KD: A self-supervised diverse knowledge distillation method for lightweight skin lesion classification using dermoscopic images," Medical Image Analysis, vol. 84, p. 102693, Feb. 2023, doi:10.1016/j.media.2022.102693.

[15] A. C. Foahom Gouabou, R. Iguernaissi, J.-L. Damoiseaux, A. Moudafi, and D. Merad, "End-to-End Decoupled Training: A Robust Deep Learning Method for Long-Tailed Classification of Dermoscopic Images for Skin Lesion Classification," Electronics, vol. 11, no. 20, p. 3275, Oct. 2022, doi: 10.3390/electronics11203275.

[16] C. Yu et al., "Acral melanoma detection using a convolutional neural network for dermoscopy images," PLOS ONE, vol. 13, no. 3, p. e0193321, Mar. 2018, doi: 10.1371/journal.pone.0193321.

[17] Luqman Hakim, Z. Sari, and H. Handhajani, "Klasifikasi Citra Pigmen Kanker Kulit Menggunakan Convolutional Neural Network," Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 5, no. 2, pp. 379–385, Apr. 2021, doi: 10.29207/resti.v5i2.3001.

[18] Q. Abbas, F. Ramzan, and M. U. Ghani, "Acral melanoma detection using dermoscopic images and convolutional neural networks," Visual Computing for Industry, Biomedicine, and Art, vol. 4, no. 1, Oct. 2021, doi: 10.1186/s42492-021-00091-z.

[19] T. J. Brinker et al., "Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task," European Journal of Cancer, vol. 113, pp. 47–54, May 2019, doi: 10.1016/j.ejca.2019.04.001.

[20] A. Murugan, S. A. H. Nair, and K. P. S. Kumar, "Detection of Skin Cancer Using SVM, Random Forest and kNN Classifiers," Journal of Medical Systems, vol. 43, no. 8, Jul. 2019, doi: 10.1007/s10916-019-1400-8.

[21] T. Saba, "Computer vision for microscopic skin cancer diagnosis using handcrafted and non-handcrafted features," Microscopy Research and

Technique, vol. 84, no. 6, pp. 1272–1283, Jan. 2021, doi:10.1002/jemt.23686.

[22] M. A. Khan, M. Y. Javed, M. Sharif, T. Saba, and A. Rehman, "Multi-Model Deep Neural Network based Features Extraction and Optimal Selection Approach for Skin Lesion Classification," 2019 International Conference on Computer and Information Sciences (ICCIS), Apr. 2019, doi: 10.1109/iccisci.2019.8716400.

[23] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," Neurocomputing, vol. 408, pp. 189–215, Sep. 2020, doi: 10.1016/j.neucom.2019.10.118.

[24] J. Gu, L. Wang, H. Wang, and S. Wang, "A novel approach to intrusion detection using SVM ensemble with feature augmentation," Computers & Security, vol. 86, pp. 53–62, Sep. 2019, doi:10.1016/j.cose.2019.05.022.

[25] M. Nawaz et al., "Melanoma localization and classification through faster region-based convolutional neural network and SVM," Multimedia Tools and Applications, vol. 80, no. 19, pp. 28953–28974, Jun. 2021, doi: 10.1007/s11042-021-11120-7..

[26] M.-W. Huang, C.-W. Chen, W.-C. Lin, S.-W. Ke, and C.-F. Tsai, "SVM and SVM Ensembles in Breast Cancer Prediction," PLOS ONE, vol. 12, no. 1, p. e0161501, Jan. 2017, doi:10.1371/journal.pone.0161501.

[27] S. Khan, H. Rahmani, S. A. A. Shah, and M. Bennamoun, "A Guide to Convolutional Neural Networks for Computer Vision," Synthesis Lectures on Computer Vision, vol. 8, no. 1, pp. 1–207, Feb. 2018, doi:10.2200/s00822ed1v01y201712cov015.

[28] Iffat. Zafar, Giounona. Tzanidou, Richard. Burton, Nimesh. Patel, and Leonardo. Araujo, *Hands-On Convolutional Neural Networks with TensorFlow : Solve Computer Vision Problems with Modeling in TensorFlow and Python.* Birmingham: Packt Publishing Ltd, 2018.

[29] Md. R. Karim, Mohit. Sewak, and Pradeep. Pujari, *Practical Convolutional Neural Networks : Implement advanced deep learning models using Python*, 1st ed. Packt Publishing, 2018.

[30] B. Santosa and A. Umam, *Data Mining dan Big Data Analytics Teori dan Impelementasi Menggunakan Python & Apache Spark*, 2nd ed. Penebar Media Pustaka, 2018.

[31] Shai. Shalev-Shwartz and Shai. Ben-David, *Understanding machine learning : from theory to algorithms*. Cambridge University Press, 2014.