Computer Engineering and Applications Vol. 12, No. 1, February 2023

Segmentation of Skin Lesions Using Convolutional Neural Networks

Firdaus Firdaus[1], Muhammad Fachrurrozi[2], Muhammad Naufal Rachmatullah[1], Siti Nurmaini[1], Dewi Chayanti[1], Annisa Darmawahyuni[1], Anggun Islami[1], Ade Iriani Sapitri[1], and Bambang Tutuko[1]

[1]Intelligent System Research Group, Universitas Sriwijaya, Palembang 30139, Indonesia.
[2]*Informatic Engineering, Faculty of Computer Science, Universitas Sriwijaya, Palembang 30139, Indonesia.*

**firdaus@unsri.ac.id*

ABSTRACT

Skin lesions play a crucial role as the initial clinical symptoms of diseases such as chickenpox and melanoma. By employing digital image processing techniques for skin cancer detection, it becomes feasible to diagnose these conditions without the need for physical contact with the skin. However, the automatic analysis of dermoscopy images, which exhibit characteristics like residue (hair and ruler markers), indistinct borders, varying contrast, and variations in shape and color, poses significant challenges. To overcome these difficulties, effective hair removal through segmentation has been explored extensively in the literature. In this study, we present a skin lesion segmentation system developed using the Convolutional Neural Networks (CNNs) method with the U-Net architecture. The model was constructed and evaluated using the HAM10000 Dataset. The results achieved by the best-performing model were outstanding, with a Pixel Accuracy, Intersection over Union (IoU), and F1 Score of 95.89%, 90.37%, and 92.54%, respectively.

Keywords: Skin lesion, Melanoma, Image processing, Convolutional neural networks, U-Net.

1. INTRODUCTION

Skin cancer has become increasingly prevalent in certain countries, making it one of the most common types of cancer worldwide. Among the different types of skin cancer, melanoma stands out due to its aggressive nature and high mortality rates [1]. In 2019, approximately 59,830 cases of skin cancer were diagnosed in the United States alone [2], while in Indonesia, skin cancer ranks as the third most prevalent type after uterine and breast cancer [3].

Skin lesions serve as the initial clinical indicators of diseases such as chickenpox and melanoma. Early detection of skin diseases can be challenging, especially for inexperienced dermatologists. However, advancements in digital image processing have paved the way for diagnosing these conditions without direct physical contact with the skin. Machine learning techniques have played a vital role in automating various processes in dermoscopy images within the medical field [4]. Nevertheless, automated analysis of dermoscopy images presents several complexities, including residues (such as hair and ruler markers), indistinct borders, varying contrast, shape differences, and color variations, which hinder accurate analysis [5]. The presence of hair within skin lesions can be effectively addressed through dermoscopy image

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

segmentation techniques [6] [7]. Image segmentation, which involves separating the image into foreground (object) and background regions, facilitates further disease analysis [8].

In the literature, segmentation of skin lesions in dermoscopy images has been explored using diverse machine learning methods, such as Trainable Weka Segmentation (TWS) [9], Markov Random Fields (MRFs) [10], Random Forest [11], K-Means [12] and Support Vector Machine (SVM) [13]. However, these methods often necessitate manually defined feature extraction, limiting their segmentation performance [14]. Deep learning methods, such as Recurrent Neural Networks (RNNs), Deep Neural Networks (DNNs), and Convolutional Neural Networks (CNNs), have emerged as a solution to this limitation [15].

The CNNs method effectively automates the segmentation of skin lesions in dermoscopy images, enabling the identification of hidden feature patterns [16]. CNNs consist of three layers: convolution, pooling, and fully connected layers [17]. The convolution layer maps input data features, allowing for feature extraction and subsequent dimension reduction in the pooling layer. The fully connected layer connects all layers of the CNNs architecture into a unified framework [18].

Among the various CNNs architectures, U-Net is commonly used for biomedical image segmentation. The U-Net architecture resembles the letter 'U' and comprises three components: the contracting, bottleneck, and expanding segments. Each segment incorporates two convolution layers (3x3) and a max-pooling layer (2x2). The number of kernels is doubled after each segment, enabling the architecture to effectively capture complex structures. The bottleneck layer acts as an intermediary between the contracting and expanding segments [19].

Based on the aforementioned considerations, this study proposes a CNNs method with the U-Net architecture for segmenting skin lesion images.

2. MATERIAL AND METHODS

The research can be broadly outlined into several main stages. The first stage involves data preparation, which includes gathering and organizing the necessary data. Next, dermoscopy image preprocessing is conducted to prepare the data before proceeding to the next stage. Subsequently, the data is divided into training and testing sets to avoid bias in the training and evaluation process. The next stage involves segmentation using the U-Net architecture within a CNNs. The best model is then determined based on the evaluation results. Finally, an analysis of the obtained results is conducted (Figure 1),

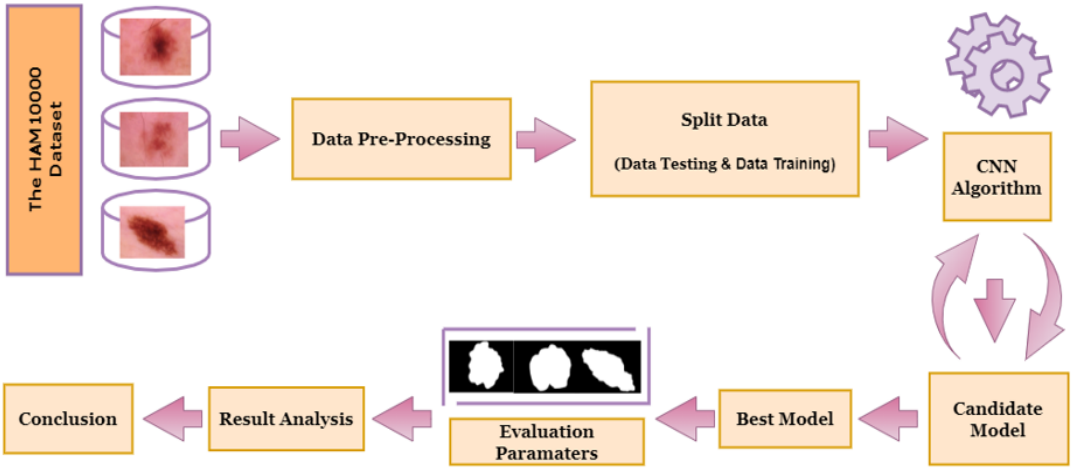
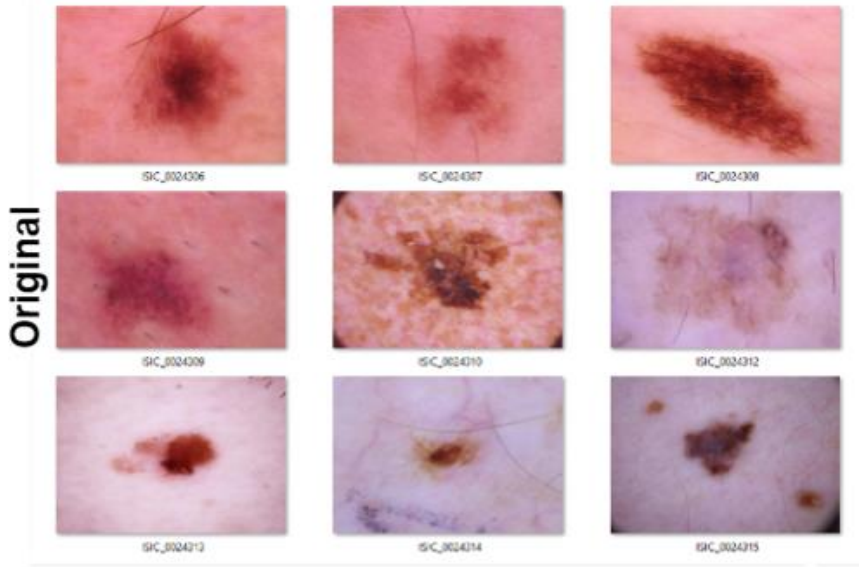


Figure 1. Segmentation Framework

2.1 DATASET

The dataset used in the research in the segmentation process is a dermoscopy image from The HAM10000 Dataset which is available in the International Skin Imaging Collaboration (ISIC) 2018 archive which can be accessed by the public via https://doi.org/10.7910/DVN/DBW86T. The example image is shown in Figure 2. The dataset is 10,015 dermoscopy images used as training data for machine learning.

The HAM10000 dataset itself stands for "Human Against Machine with 10,000 training images". The dataset was collected from two different locations over 20 years: Cliff Rosendahl's skin cancer practice in Queensland, Australia, and the Department of Dermatology at the Medical University of Vienna, Austria [20].



(a)

(b)

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

Figure 2. The example of (a) raw data and (b) data labels from the HAM10000 Dataset

In this study, the entire data available in the HAM10000 dataset was used for seven diagnostic category classes. The total number of data used was 10,015 dermoscopy images (Table 1). This dataset is biased towards the melanocytic nevi class, while the class with the second highest sample count is melanoma.

Table 1. Number of images in HAM10000 Dataset

Skin Lesions Category	Number of images
Actinic keratosis	327
Basal cell carcinoma	514
Benign keratosis-like lesions	1,099
Dermatofibroma	115
Melanoma	1,113
Melanocytic nevi	6,705
Vascular lesions	142
Sum	10,015

2.1 Data Pre-processing

The data preprocessing process is conducted to prepare the images before they are fed into the network system. This stage encompasses image resizing, data augmentation, data digitization, and feature scaling (Figure 3).

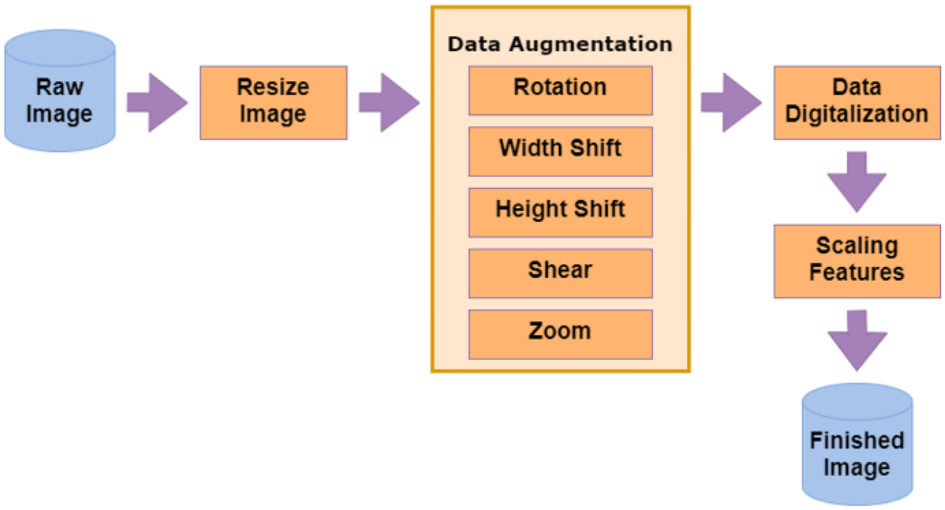


Figure 3. Data Pre-processing Stages

Image resizing involves adjusting the dimensions of the image, including width and height. The original images in the dataset used for this study have a resolution of 450 x 600 pixels. However, these images are too large and computationally intensive, and the network system requires square-shaped dimensions. Therefore, it is necessary to scale and resize the images to 256 x 256 RGB images.

The next step is data augmentation, a technique employed to augment the dataset without compromising its core characteristics, thereby generating data variations. This is achieved by applying random transformations to the data. The data augmentation techniques utilized in this study include rotation, shear, zoom,

width shift, and height shift. The data augmentation process is implemented using an Image Data Generator.

Data digitization involves representing the images numerically to enable computer interpretation. In this study, the images used are 8-bit RGB digital images. Digital images are represented by matrices consisting of rows and columns. During the digitization process, an image with M rows and N columns is obtained, resulting in an MxN matrix.

Feature scaling is the process of normalizing the numeric data in a dataset by dividing the RGB values ranging from 0 to 255 by 255. This transformation ensures that the RGB values fall within the range of 0 to 1. Feature scaling is beneficial for enhancing the learning process. After the data has undergone preprocessing, the next step is to divide the data into two parts: 80% of the data for training and 20% for testing.

2.2 Segmention Model Development

The CNNs skin lesion segmentation process utilizes the U-Net architecture. The U-Net architecture is constructed with various parameters, such as activation functions in the hidden layers, activation functions in the output layer, optimizer, batch size, number of neurons, number of epochs, number of input layers, learning rate, number of hidden layers, and the loss function, as shown in Table 2.

Table 2. General Parameters of U-Net Architecture

Method	Input Layer	Activation Function Hidden	Activation Function Output	Optimatio n	Epoch	Neuron
U-Net	(256,256, 1)	Relu	Sigmoid	Adam	800	512

The parameters that are tuned in this study are the loss function, learning rate, and batch size. The loss functions being experimented with are binary cross-entropy and dice coefficient. The learning rates used are 10^{-4} and 10^{-5}, while the batch sizes tested are 32 and 64. The combination scenarios of the hyperparameters can be seen in Table 3.

Table 3. Hyperparameter tuning scenario.

Model	*Loss Function*	*Learning Rate*	*Batch Size*
Model 1	*Binary Crossentropy*	10^{-4}	64
Model 2	*Binary Crossentropy*	10^{-5}	32
Model 3	*Binary Crossentropy*	10^{-5}	64
Model 4	*Dice Coefficient*	10^{-4}	64
Model 5	*Dice Coefficient*	10^{-5}	32
Model 6	*Dice Coefficient*	10^{-5}	64

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

The performance of the segmentation model is evaluated based on pixel accuracy (equation 1), Intersection over Union (IoU) (Equation 2), and F1 score (Equation 3).

$$Pixel\ Accuarcy = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Posotive + False\ Negative} \quad (1)$$

$$IoU = \frac{True\ Positive}{True\ Positive + False\ Positive + False\ Negative} \quad (2)$$

$$F_1 Score = \frac{True\ Positive}{True\ Positive + \frac{1}{2}(False\ Positive + False\ Negative)} \quad (3)$$

3. RESULTS AND DISCUSSION

Based on the performance of the results of training and testing on the U-Net architecture for image segmentation of the overall skin lesion, a comparison of models, can be seen in Table 4.

Table 4. Evaluation Results of 6 U-Net Models on Testing Data

Model	*Pixel Accuracy*	IoU	*F1 Score*
Model 1	95.86	90.29	92.53
Model 2	95.20	88.82	90.94
Model 3	95.61	89.70	91.73
Model 4	95.67	89.79	92.10
Model 5	95.62	89.72	92.06
Model 6	**95.89**	**90.37**	**92.54**

After testing the test skin lesion images with the 6 best models and shown as a whole as shown in Table 4, it shows that Model 6 is the best model among other candidate models. Model 6 using the loss function dice coefficient parameter, epoch learning rate $10^{-5,}$ and batch size 64 has a Pixel Accuracy value of 95.89%, Intersection over Union (IoU) of 90.37%, and F1 Score of 92.54%.

The results of the experiment using the loss function dice coefficient, epoch learning rate $10^{-5,}$ and batch size 32. The following comparison graph between the accuracy of the model is shown in Figure 4 and the loss model in Figure 5.

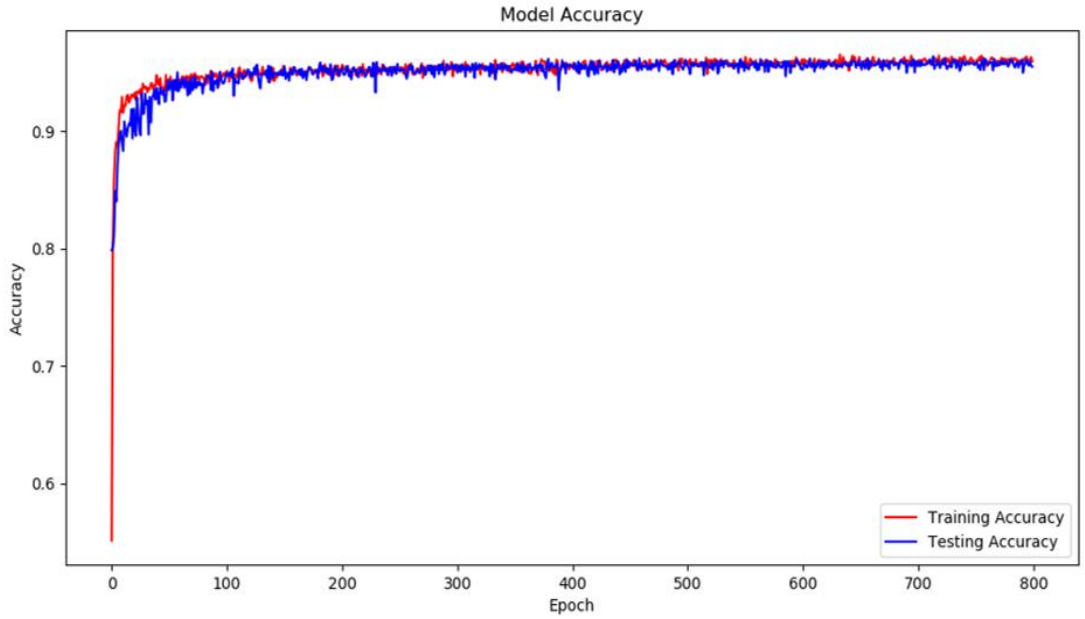


Figure 4. Graph of U-Net Model 5 Training and Testing Process Accuracy

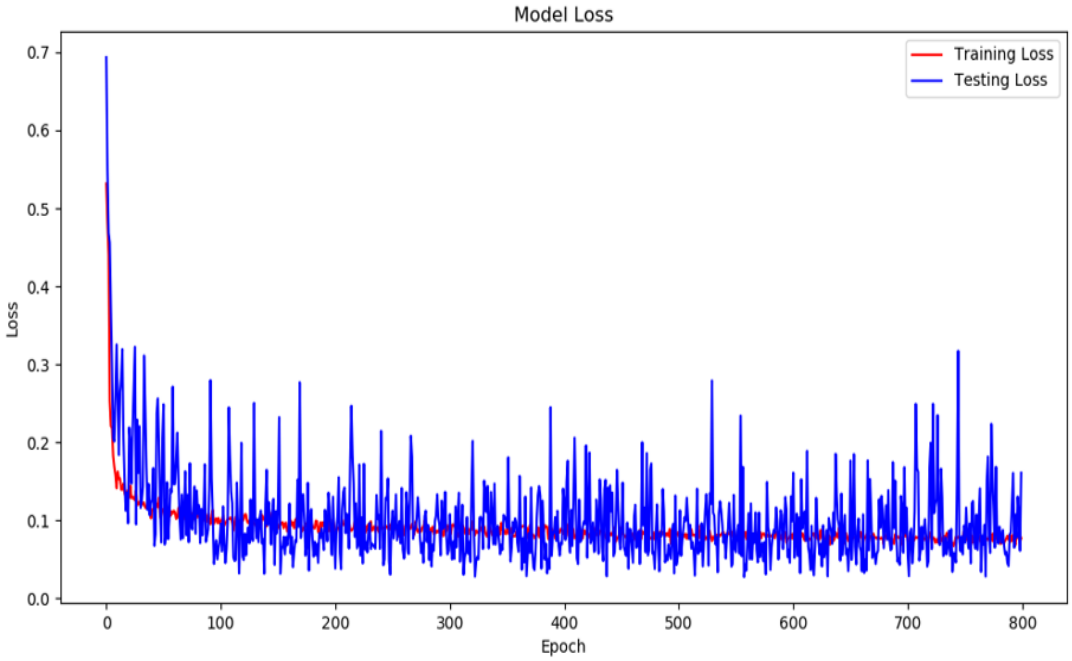


Figure 5. Loss Graph of U-Net Model 6 Training and Testing Process

In the initial epochs, both the training and testing losses decrease rapidly, indicating effective learning. As the number of epochs increases, the training and testing losses continue to decrease, suggesting that the model is learning and generalizing well. A good fit is indicated by the convergence of the training and validation losses, indicating that the model is not overfitting or underfitting the data. This convergence suggests that the model has reached a stable point in its learning process. Overall, with 800 epochs, the learning curve demonstrates a good fit

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

between the training and validation data, indicating that the model has successfully learned and can generalize well to unseen examples.

Several other studies that have applied the CNNs method with various architectures for segmenting skin lesions with dermoscopy images in recent years are compared with the results of the best segmentation model in this study, which are summarized in Table 5.

Table 5. Comparison of Result with Other Relevant Research

ResearchMe	Method	Dataset	Acc	IoU	F1 Score
Salih & Viriri [21]	SRM+MRF	ISIC 2018	0.92	0.79	0.88
Jin et. al. [22]	CKDNet	ISIC 2018	0.93	0.79	0.87
Arora et. al. [23]	Attn_U-Net+GN	ISIC 2018	0.95	0.83	0.91
Proposed*	U-Net	ISIC 2018	0.95	0.90	0.92

From the table, it is evident that the U-Net model proposed in this study outperforms other architectures explored in the referenced studies. Comparing the metrics, the U-Net model in this study achieves an accuracy of 0.95, IoU of 0.90, and an F1 Score of 0.92.

Based on these comparisons, it can be concluded that the U-Net model used in this study demonstrates superior performance in skin lesion segmentation when compared to other CNN architectures examined in the referenced studies. The U-Net model achieves higher IoU and F1 Score values, indicating improved accuracy and better overall segmentation quality.

4. CONCLUSION

Based on the results and analysis of research that has been carried oust on the segmentation process, the comparison of the performance of skin lesion segmentation using the CNNs can be concluded that the CNNs method with the U-Net architecture can be used to segment dermoscopy images of skin lesions properly. The best model performance resulting from CNNs segmentation on U-Net architecture with the best parameters from model 6 has a Pixel Accuracy value of 95.89%, Intersection over Union (IoU) of 90.37%, and F1 Score of 92.54%. The U-Net model in this study is superior to other architectures that have been studied by other studies using the CNNs method.

Based on the research that has been done, it has limitations such as the process of fine-tuning parameters. So, a wider study is needed, such as applying algorithms for fine-tuning the right parameters, where fine-tuning is very time-consuming in the training process with limited computational resources, making efficient implementation very necessary. The application of another CNNs architecture for the task of segmenting skin lesion images that can help improve model performance is better for future needs.

REFERENCES

[1] R. B. Oliveira, J. P. Papa, A. S. Pereira, and J. M. R. S. Tavares, "Computational methods for pigmented skin lesion classification in images: review and future trends," *Neural Computing and Applications*, vol. 29, no. 3.

Springer London, pp. 613–636, Feb. 01, 2018, doi: 10.1007/s00521-016-2482-6.

[2] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," *CA. Cancer J. Clin.*, 2019, doi: 10.3322/caac.21551.

[3] S. Wilvestra, S. Lestari, and E. Asri, "Studi Retrospektif Kanker Kulit di Poliklinik Ilmu Kesehatan Kulit dan Kelamin RS Dr. M. Djamil Padang Periode Tahun 2015-2017," *J. Kesehat. Andalas*, 2018, doi: 10.25077/jka.v7i0.873.

[4] R. Sumithra, M. Suhil, and D. S. Guru, "Segmentation and classification of skin lesions for disease diagnosis," in *Procedia Computer Science*, Jan. 2015, vol. 45, no. C, pp. 76–85, doi: 10.1016/j.procs.2015.03.090.

[5] Ş. Öztürk and U. Özkaya, "Skin Lesion Segmentation with Improved Convolutional Neural Network," *J. Digit. Imaging*, vol. 33, no. 4, pp. 958–970, Aug. 2020, doi: 10.1007/s10278-020-00343-z.

[6] S. Joseph and J. R. Panicker, "Skin lesion analysis system for melanoma detection with an effective hair segmentation method," in *Proceedings - 2016 International Conference on Information Science, ICIS 2016*, Feb. 2017, pp. 91–96, doi: 10.1109/INFOSCI.2016.7845307.

[7] E. A. Al-Mansour and A. Jaffar, "A study on automatic segmentation and classification of skin lesions in dermoscopic images," in *Oncology: Breakthroughs in Research and Practice*, vol. 2–2, IGI Global, 2016, pp. 559–569.

[8] A. Sharma, R. Chaturvedi, U. K. Dwivedi, S. Kumar, and S. Reddy, "Firefly algorithm based effective gray scale image segmentation using multilevel thresholding and entropy function," *Int. J. Pure Appl. Math.*, 2018.

[9] I. Arganda-Carreras *et al.*, "Trainable Weka Segmentation: A machine learning tool for microscopy pixel classification," *Bioinformatics*, vol. 33, no. 15, pp. 2424–2426, Aug. 2017, doi: 10.1093/bioinformatics/btx180.

[10] N. S. M. Raja, V. Rajinikanth, S. L. Fernandes, and S. C. Satapathy, "Segmentation of breast thermal images using kapur's entropy and hidden markov random field," *J. Med. Imaging Heal. Informatics*, 2017, doi: 10.1166/jmihi.2017.2267.

[11] B. Kang and T. Q. Nguyen, "Random Forest with Learned Representations for Semantic Segmentation," *IEEE Trans. Image Process.*, 2019, doi: 10.1109/TIP.2019.2905081.

[12] X. Zheng, Q. Lei, R. Yao, Y. Gong, and Q. Yin, "Image segmentation based on adaptive K-means algorithm," *Eurasip J. Image Video Process.*, 2018, doi: 10.1186/s13640-018-0309-3.

[13] T. Ramakrishnan and B. Sankaragomathi, "A professional estimate on the computed tomography brain tumor images using SVM-SMO for classification and MRG-GWO for segmentation," *Pattern Recognit. Lett.*, 2017, doi: 10.1016/j.patrec.2017.03.026.

[14] F. A. Khan, U. Voß, M. P. Pound, and A. P. French, "Volumetric Segmentation of Cell Cycle Markers in Confocal Images Using Machine Learning and Deep Learning," *Front. Plant Sci.*, 2020, doi: 10.3389/fpls.2020.01275.

[15] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*. 2015, doi:

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

10.1038/nature14539.

[16] R. Mishra and O. Daescu, "Deep learning for skin lesion segmentation," in *Proceedings - 2017 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2017,* Dec. 2017, vol. 2017-January, pp. 1189–1194, doi: 10.1109/BIBM.2017.8217826.

[17] Z. Zhao and A. Kumar, "Accurate Periocular Recognition under Less Constrained Environment Using Semantics-Assisted Convolutional Neural Network," *IEEE Trans. Inf. Forensics Secur.*, 2017, doi: 10.1109/TIFS.2016.2636093.

[18] Q. Zhang, M. Zhang, T. Chen, Z. Sun, Y. Ma, and B. Yu, "Recent advances in convolutional neural network acceleration," *Neurocomputing*, 2019, doi: 10.1016/j.neucom.2018.09.038.

[19] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015, doi: 10.1007/978-3-319-24574-4_28.

[20] P. Tschandl, C. Rosendahl, and H. Kittler, "Data descriptor: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Sci. Data*, vol. 5, no. 1, pp. 1–9, Aug. 2018, doi: 10.1038/sdata.2018.161.

[21] O. Salih and S. Viriri, "Skin lesion segmentation using stochastic region-merging and pixel-based markov random field," *Symmetry (Basel).*, 2020, doi: 10.3390/SYM12081224.

[22] Q. Jin, H. Cui, C. Sun, Z. Meng, and R. Su, "Cascade knowledge diffusion network for skin lesion diagnosis and segmentation," *Appl. Soft Comput.*, vol. 99, p. 106881, Feb. 2021, doi: 10.1016/j.asoc.2020.106881.

[23] R. Arora, B. Raman, K. Nayyar, and R. Awasthi, "Automated skin lesion segmentation using attention-based deep convolutional neural network," *Biomed. Signal Process. Control*, vol. 65, p. 102358, Mar. 2021, doi: 10.1016/j.bspc.2020.102358.