An Effectivity Deep Learning Optimization Model to Traditional Batak Culture Ulos Classification

Rizki Muliono[1], Mayang Septania Iranita[2], *Rahmad BY Syah[3]
Faculty of Engineering, Informatics Department, Universitas Medan Area, Medan, Indonesia[1, 2, 3]
Excellent Centre of Innovations and New Science, Universitas Medan Area, Medan, Indonesia[1, 2, 3]

***Abstract*—Ulos is one of the Batak culture's traditional heritage fabrics. Ulos cloth is divided into several types, each with a distinct function. Ulos Ragi Hotang, Ulos Pinunsaan, Ulos Tumtuman, Ulos Ragi Hidup, and Ulos Sadum are the five Batak ulos motifs. The Batak ulos motif has evolved over time and is now well-known in other countries. However, many ordinary people have difficulty distinguishing between ulos cloth and other fabrics. This study categorizes the different types of ulos cloth so that it can be used by ordinary people who are unfamiliar with the different types and functions. The Convolutional Neural Network is the method used (CNN). CNN is used to recognize and classify images. CNN's main feature is that it detects feature patches from locations in the input matrix and assembles them into high-level references. The Modular Neural Network (MNN) is then used to break down large and complex computational processes into smaller components, reducing complexity while still producing the desired output. 80% of the data for the training process, 20% for testing. The accuracy value achieved is 97.83%, the loss value is 0.0793, the val loss is 2.1885, and the val accuracy is 0.7429.**

***Keywords*—Ulos; classification; convolutional neural network; modular neural network; deep learning**

I. Introduction

Indonesia has a rich cultural diversity. Language, ethnicity, religion, and community beliefs are all part of the community. The Toba Batak are a tribe found throughout North Sumatra [1]. In everyday life, the Batak tribe maintains culture, as evidenced by the performance of traditional events with a distinct feature, namely, ulos [2].

Ulos are used in Toba Batak traditional ceremonies, with different types and functions depending on their roles. At first glance, Batak Ulos appear to have the same motif or size, but this is not the case [3]. If you want to learn more about the different types of Batak Ulos, you can talk to weavers or those who work with ulos. The motif of the Batak ulos has, of course, evolved over time, and it is now well known not only in Indonesia, but also in other countries [4].

Toba Batak has five different types of ulos. Ulos Ragi Hotang is used in wedding traditions. Ulos Pinuncaan is worn by kings and used in large customary events. The first child of sedition wore Ulos Tumtuman. Ulos Yeast Life as a metaphor for finding happiness in life. Ulos Sadum is commonly given as a gift to officials.

One of the long-standing problems in computer vision that has yet to be solved is the classification of objects in images in general [5]. How to replicate the human ability to understand image information so that computers, like humans, can recognize objects in images. The feature engineering process is generally very limited in that it can only apply to specific datasets and cannot generalize to any type of image. This is due to various differences between images, such as differences in viewing angles, scale, light conditions, object deformation, and so on [6,7].

There are several algorithms that can be used in image processing. They include Nave Bayes, Support Vector Machines, and Neural Networks. A Neural Network is a commonly used algorithm [8]. The workings of neural networks in the human brain inspired the development of neural networks. Digital image processing algorithms have also been developed in tandem with technological advancements. The Convolutional Neural Network is a deep learning development [9,10].

Deep learning is an artificial intelligence (AI) method that teaches computers to process data in ways similar to how the human brain does. Deep learning models are capable of recognizing complex patterns in images, text, sound, and other data in order to generate insights and accurate predictions [11]. Deep learning methods can be used to automate tasks that would otherwise require human intelligence, such as describing an image or transcribing a sound file into text. In digital image recognition, the Convolutional Neural Network method produces the most significant results. This is due to the fact that CNN is based on an image recognition system in the human visual cortex [12].

A modular neural network is a collection of different neural networks that work independently to produce output without interacting with one another. Each neural network performs different sub-tasks by obtaining distinct inputs from other networks [13]. The benefit of this artificial neural network is that it divides large and complex computational processes into smaller components, reducing complexity while still producing the desired output. Interfaces between modules in such a modular architecture can be viewed as "information relays" that encode, delimit, and disseminate critical information [14].

II. Research Methodology

A. State of the Art

The following is a classification model formation scheme for CNN and MNN.

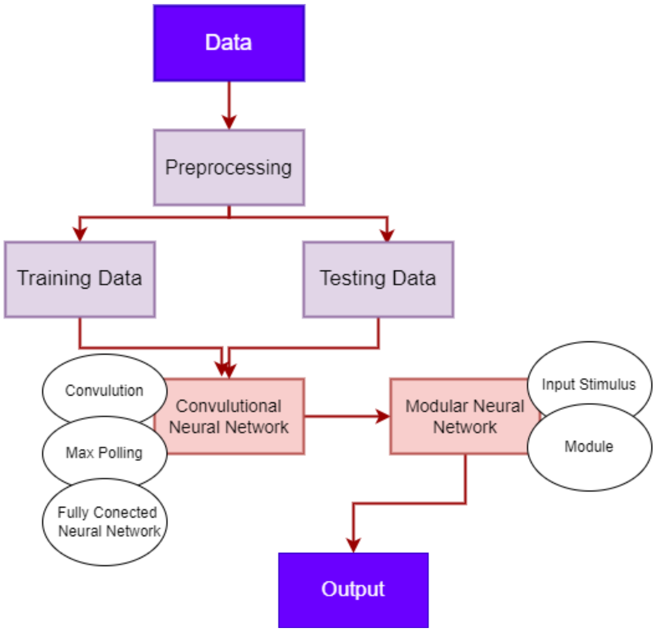

Fig. 1. State of the art.

The stages of the research can be seen in Fig. 1 [15,16].

- Data
- Data pre-processing to sort.
- Training to identify the features of each image, followed by labelled neurons to activate the image, and finally classification. Batch size and epoch are initialization parameters in the training process.
- The image is multiplied by a 32x32 kernel on three filters in the first convolution. The image is multiplied by the 64x64 kernel in the second convolution. The third convolution is an image multiplied by a 64x64 kernel with three filters. The resulting feature will make use of the same padding as well as the ReLu activation function. Using three dense layers, the first and second dense 128 layers, and the third 64 layers, the ReLu activation function was used.
- Testing to determine how well the system detects and calculates the number of images in the image. The fitting models obtained from the previous training stages were compared during validation.
- Using a circular control area around each output, the Modular Neural Network selects the number and location of inputs that affect each output. The control area uses each input's response range to identify the critical input and discard the rest.

B. Data Collection

Images obtained from observations at UD. Ulos and Songket Tamariska Tarutung, North Tapanuli, North Sumatra Province. Images in the JPEG format were captured with a smartphone camera. There are 200 images in total, including Ulos Ragi Hotang, Ulos Pinuncaan, Ulos Tumtuman, Ulos Ragi Hidup, and Ulos Sadum [17]. See .Fig. 2 for types of ulos.

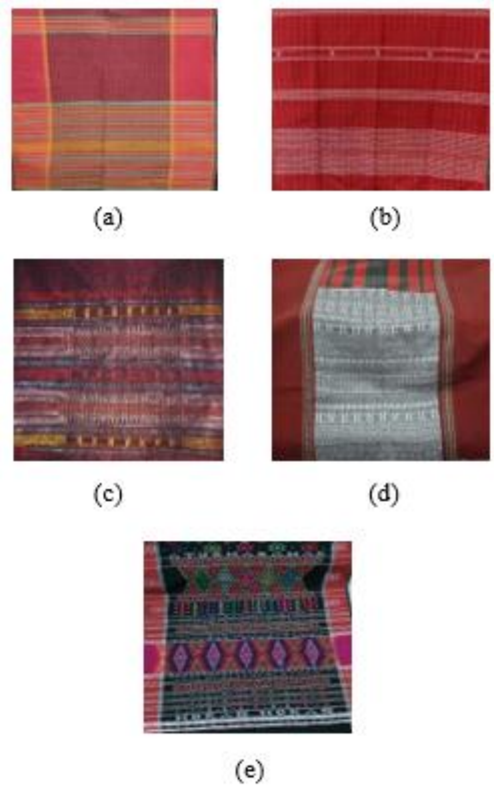

Fig. 2. Types of Ulos. (a) Ulos ragi hotang, (b) Ulos Pinuncaan, (c) Ulos tumtuman, (d) Ulos ragi hidup, (e) Ulos sadum [17].

C. Convolutional Neural Network (CNN)

CNN consists of several layers. The CNN architecture has the following components [18,19,20]:

1) Input layer: This layer is the first to load and enter data, which is then carried over to the next layer.

2) Convolution layer: The convolution layer performs the convolution operation on the output from the previous layer. This layer serves as the foundation for CNN operations.

3) Activation layer: To convert a node's input signal into an output signal. The output results will be carried over to the next layer [21].

$$A(x) = \max(0, x) \tag{1}$$

$$C_i = \frac{a^{xi+\text{LOG}(b)}}{\sum_{k=1}^{S} a^{xi+\text{LOG}(b)}} \tag{2}$$

xi is the strength value of the neuron. Log(b) constant value that can be determined [22].

$$\frac{\partial_{ai}}{\partial_{aj}} = C_i(\delta ij - Cj)\; with\; \delta ij$$

$$= \begin{Bmatrix} 1, if\ i = j \\ 0, if\ i\ \neq j \end{Bmatrix} \quad (3)$$

4) Max polling: minimizing the number of parameters and calculations needed.

5) Dropout: To prevent overfitting problem.

6) Fully connected layer: All neuron activation layers in the previous layer are fully connected.

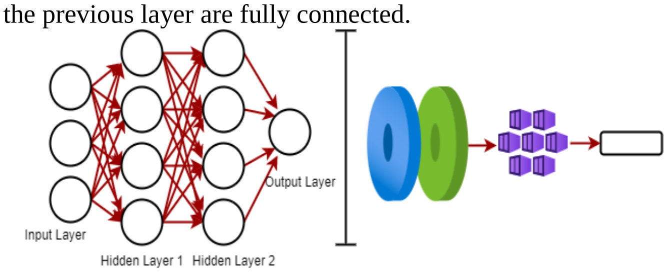

Fig. 3. CNN architecture.

Fig. 3 depicts the relationship between weighted units used to determine the effect of one object on another. CNN has one or more convolution layers that perform the convolution operation on the input before passing the results in the form of output to the next layer [23].

Several learning parameters, including batch size and epoch, must be set. Table I shows an explanation of the parameters [24].

TABLE I. INITIATION PARAMETERS

Parameter	Information
Batch Size	The number of data samples sent to the neural network in a single epoch.
Epoch	The number of rounds completed from the start of the first dataset to the end.

D. Modular Neural Network

Modular is a self-contained system that interacts with the overall architectural function to perform complex tasks. To achieve local computation, an explicit Action must be performed, so that the system being modeled becomes a meaningful function. In eq. (4), (5) and (6), static functions are represented in two dimensions [25, 26].

$$A(c(x), h(x)) = [3x] + [-x^2 + x] \quad (4)$$

$$c(x) = 3x \quad (5)$$

$$h(x) = -x^2 + x \quad (6)$$

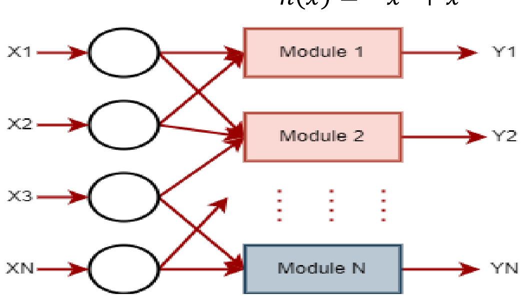

Fig. 4. Modular neural network structure.

A modular neural network structure in Fig. 4 [27].

- Tasks should be broken down into subtasks.
- Modular architecture organization
- Communication between modules

A vector can be used as the target response. On the network, each input vector is represented. Eq. (7) represents the MNN output [28].

$$h = \sum_{i=1}^{k} c_i y_i \quad (7)$$

c_i shows the weight on the number of network outputs resulting from the response, as explained by the deterministic transformation. Eq. (8) can be used to express the response [29].

$$a = R_i(x) + \epsilon_i \qquad i = 1,2,3, \ldots\ldots n \quad (8)$$

III. RESULT AND DISCUSSION

Training data can account for up to 80% of the total. In the training process, the iteration parameter is set to 100 epochs, and the batch size is set to 32. The training procedure is carried out and is repeated 100 times in order to obtain the desired feature extraction. Table II shows the amount of training and testing data.

TABLE II. TRAINING DATA AND TESTING DATA

Distribution of Testing & Training Data	Amount of data
Training 80%	100
Testing 20%	35

The test employs 35 test data, with 7 images in each Ulos category. The model is used directly in the process of calculating accuracy. fit() method by calling the validation folder's directory and then executing it with the CNN method.

TABLE III. RESULTS BASED ON PARAMETERS

Epoch	Loss	Accuracy	Val Loss	Val Accuracy	Time
50	0.3315	0.8768	1.4779	0.6571	4 minute 13 second
75	0.2361	0.9203	0.9707	0.6571	20 minute 45 second
100	0.0793	0.9783	2.1885	0.7429	18 minute 40 second

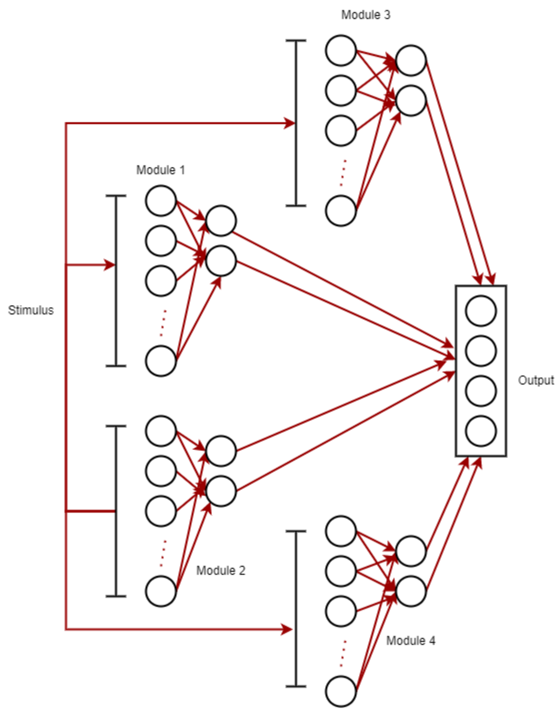

Fig. 5. MNN modeling paradigm [26].

MNN modeling paradigm is shown in Fig. 5. Table III shows the goal of determining model parameters: to compare the best models based on parameter values. It is re-optimized based on these parameters by dividing the smallest modules into the greatest number. The best accuracy results are found in the 100th epoch, which is 97.83%.

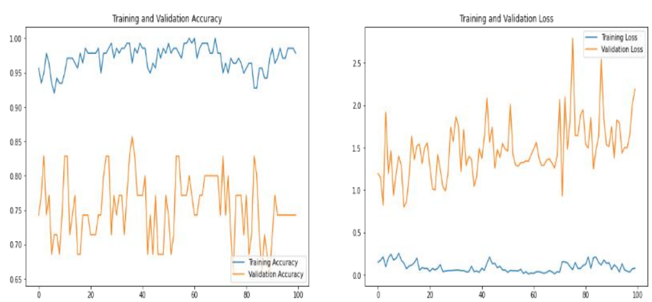

Fig. 6. Accuracy and loss at epoch 100.

Fig. 6 shows that the training and validation graphs on accuracy have the same pattern as the training and validation graphs on loss.

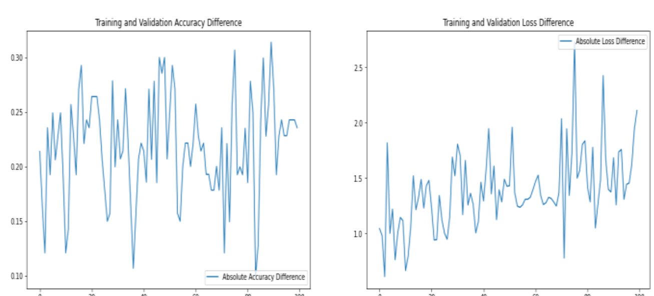

Fig. 7. Visualization of trends in trains and tests in epoch 100.

The accuracy and loss graphs have diverse patterns, as seen in Fig. 7. Due to insufficient data, this occurs and the pattern keeps repeating.

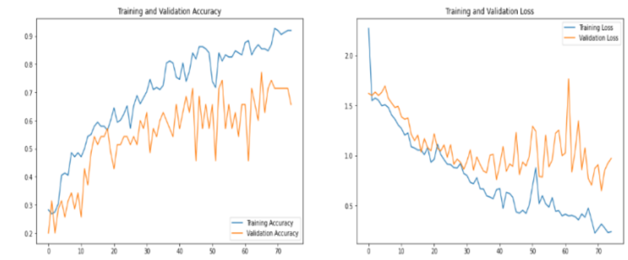

Fig. 8. Accuracy and loss at epoch 75.

An accuracy validation chart that follows the pattern of training results is shown in Fig. 8. Similar results can be obtained with the loss validation graph that mimics the training outcomes.

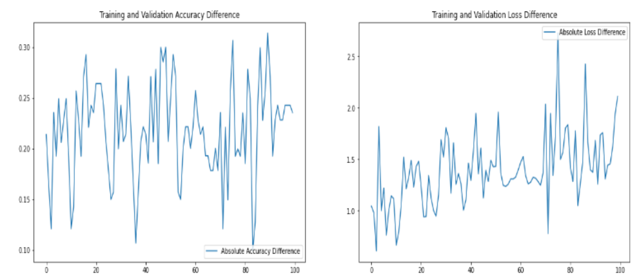

Fig. 9. Visualization of trends in trains and tests in epoch 75.

Fig. 9 shows the accuracy and loss graphs have different patterns. This is due to too little data so the pattern repeats.

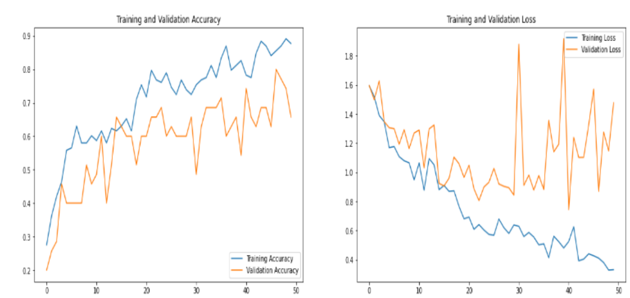

Fig. 10. Accuracy and loss at epoch 50.

Fig. 10 displays a validation graph for accuracy and a validation graph for losses that follow the pattern of training results.

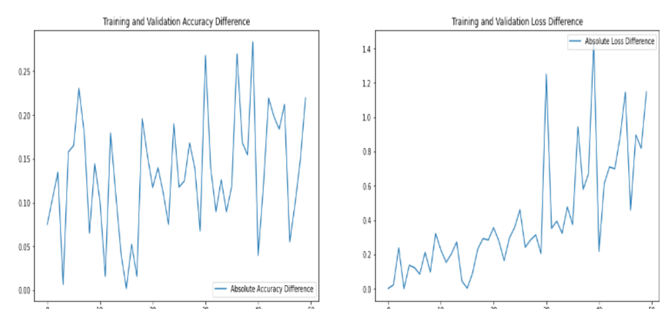

Fig. 11. Visualization of trends in trains and tests in epoch 50.

Fig. 11 shows the accuracy and loss graphs have different patterns due to too little data.

IV. Conclusion

The testing procedure is by using 35 test data with 7 images in each Ulos category. Because CNN already has a feature learning phase or process, it can classify objects without the need for additional feature extraction. When the process achieves accuracy, the use of MNN facilitates classification, where the process makes it easier to understand behavior or features in images to make them easier to recognize, and it increases in terms of time. Accuracy was 97.83%, loss was 0.0793, val loss was 2.1885, and val accuracy was 0.7429.

References

[1] E. Fatmawati, "Strategies to grow a proud attitude towards Indonesian cultural diversity," Linguistics and Culture Review, vol. 5, no. S1, pp. 810–820, Sep. 2021, doi: 10.21744/lingcure.v5nS1.1465.

[2] T. M. Panggabean and A. Christy Barus, "Combining Local and Global Descriptors Through Rotation Invariant Texture Analysis for Ulos Classification," in 2019 7th International Conference on Robot Intelligence Technology and Applications (RiTA), Nov. 2019, pp. 153–159. doi: 10.1109/RITAPP.2019.8932823.

[3] C. Nugroho, I. K. Nurhayati, K. Nasionalita, and R. M. U. Malau, "Weaving and Cultural Identity of Batak Toba Women," J Asian Afr Stud, vol. 56, no. 6, pp. 1165–1177, Sep. 2021, doi: 10.1177/0021909620958032.

[4] Zulkifli and M. Ridwan, "Revitalization of the traditional values lost due to the commodification of art/crafts: a case study of Bataknese traditional Ulos," Asian Ethnicity, vol. 20, no. 4, pp. 541–554, Oct. 2019, doi: 10.1080/14631369.2019.1608812.

[5] X. Feng, Y. Jiang, X. Yang, M. Du, and X. Li, "Computer vision algorithms and hardware implementations: A survey," Integration, vol. 69, pp. 309–320, Nov. 2019, doi: 10.1016/j.vlsi.2019.07.005.

[6] A. O. Salau and S. Jain, "Feature Extraction: A Survey of the Types, Techniques, Applications," in 2019 International Conference on Signal Processing and Communication (ICSC), Mar. 2019, pp. 158–164. doi: 10.1109/ICSC45622.2019.8938371.

[7] A. Al-Kaff, D. Martín, F. García, A. de la Escalera, and J. María Armingol, "Survey of computer vision algorithms and applications for unmanned aerial vehicles," Expert Syst Appl, vol. 92, pp. 447–463, Feb. 2018, doi: 10.1016/j.eswa.2017.09.033.

[8] X. Li et al., "A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches," Artif Intell Rev, vol. 55, no. 6, pp. 4809–4878, Aug. 2022, doi: 10.1007/s10462-021-10121-0.

[9] S. Sutrisno, N. Khairina, R. B. Y. Syah, E. Eftekhari-Zadeh, and S. Amiri, "Improved Artificial Neural Network with High Precision for Predicting Burnout among Managers and Employees of Start-Ups during COVID-19 Pandemic," Electronics (Basel), vol. 12, no. 5, p. 1109, Feb. 2023, doi: 10.3390/electronics12051109.

[10] M. Ramanan et al., "Secure blockchain enabled Cyber- Physical health systems using ensemble convolution neural network classification," Computers and Electrical Engineering, vol. 101, p. 108058, Jul. 2022, doi: 10.1016/j.compeleceng.2022.108058.

[11] K. Choudhary et al., "Recent advances and applications of deep learning methods in materials science," NPJ Comput Mater, vol. 8, no. 1, p. 59, Apr. 2022, doi: 10.1038/s41524-022-00734-6.

[12] Y. Dong, Q. Liu, B. Du, and L. Zhang, "Weighted Feature Fusion of Convolutional Neural Network and Graph Attention Network for Hyperspectral Image Classification," IEEE Transactions on Image Processing, vol. 31, pp. 1559–1572, 2022, doi: 10.1109/TIP.2022.3144017.

[13] Y. Jiao, B. Xue, C. Lu, M. S. Avidan, and T. Kannampallil, "Continuous real-time prediction of surgical case duration using a modular artificial neural network," Br J Anaesth, vol. 128, no. 5, pp. 829–837, May 2022, doi: 10.1016/j.bja.2021.12.039.

[14] J. Kauer-Bonin et al., "Modular deep neural networks for automatic quality control of retinal optical coherence tomography scans," Comput Biol Med, vol. 141, p. 104822, Feb. 2022, doi: 10.1016/j.compbiomed.2021.104822.

[15] S. Tripathy and R. Singh, "Convolutional Neural Network: An Overview and Application in Image Classification," 2022, pp. 145–153. doi: 10.1007/978-981-16-4538-9_15.

[16] Y. Bhatia, A. H. Bari, G.-S. J. Hsu, and M. Gavrilova, "Motion Capture Sensor-Based Emotion Recognition Using a Bi-Modular Sequential Neural Network," Sensors, vol. 22, no. 1, p. 403, Jan. 2022, doi: 10.3390/s22010403.

[17] F. R. Suwanto, D. Kartika, and D. Y. Niska, "Ethnomathematics: An analysis of frieze and crystallographic patterns on Ulos," 2022, p. 110021. doi: 10.1063/5.0113269.

[18] Y.-L. Chang et al., "Consolidated Convolutional Neural Network for Hyperspectral Image Classification," Remote Sens (Basel), vol. 14, no. 7, p. 1571, Mar. 2022, doi: 10.3390/rs14071571.

[19] A. Sellami and S. Tabbone, "Deep neural networks-based relevant latent representation learning for hyperspectral image classification," Pattern Recognit, vol. 121, p. 108224, Jan. 2022, doi: 10.1016/j.patcog.2021.108224.

[20] A. S. Paymode and V. B. Malode, "Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG," Artificial Intelligence in Agriculture, vol. 6, pp. 23–33, 2022, doi: 10.1016/j.aiia.2021.12.002.

[21] Y. Kong, X. Ma, and C. Wen, "A New Method of Deep Convolutional Neural Network Image Classification Based on Knowledge Transfer in Small Label Sample Environment," Sensors, vol. 22, no. 3, p. 898, Jan. 2022, doi: 10.3390/s22030898.

[22] T. Hassanzadeh, D. Essam, and R. Sarker, "EvoDCNN: An evolutionary deep convolutional neural network for image classification," Neurocomputing, vol. 488, pp. 271–283, Jun. 2022, doi: 10.1016/j.neucom.2022.02.003.

[23] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, and Y. Miao, "Review of Image Classification Algorithms Based on Convolutional Neural Networks," Remote Sens (Basel), vol. 13, no. 22, p. 4712, Nov. 2021, doi: 10.3390/rs13224712.

[24] X. Cao, J. Yao, Z. Xu, and D. Meng, "Hyperspectral Image Classification With Convolutional Neural Network and Active Learning," IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 7, pp. 4604–4616, Jul. 2020, doi: 10.1109/TGRS.2020.2964627.

[25] J. Soto, O. Castillo, P. Melin, and W. Pedrycz, "A New Approach to Multiple Time Series Prediction Using MIMO Fuzzy Aggregation Models with Modular Neural Networks," International Journal of Fuzzy Systems, vol. 21, no. 5, pp. 1629–1648, Jul. 2019, doi: 10.1007/s40815-019-00642-w.

[26] M. Dindin, Y. Umeda, and F. Chazal, "Topological Data Analysis for Arrhythmia Detection Through Modular Neural Networks," 2020, pp. 177–188. doi: 10.1007/978-3-030-47358-7_17.

[27] M. Amer and T. Maul, "A review of modularization techniques in artificial neural networks," Artif Intell Rev, vol. 52, no. 1, pp. 527–561, Jun. 2019, doi: 10.1007/s10462-019-09706-7.

[28] R. Atefinia and M. Ahmadi, "Network intrusion detection using multi-architectural modular deep neural network," J Supercomput, vol. 77, no. 4, pp. 3571–3593, Apr. 2021, doi: 10.1007/s11227-020-03410-y.

[29] M. K. Dahouda and I. Joe, "Neural Architecture Search Net-Based Feature Extraction With Modular Neural Network for Image Classification of Copper/ Cobalt Raw Minerals," IEEE Access, vol. 10, pp. 72253–72262, 2022, doi: 10.1109/ACCESS.2022.3187420.