IT Journal Research and Development (ITJRD)
Vol.8, No.2, March 2024, E-ISSN : 2528-4053 | P-ISSN : 2528-4061
DOI : 10.25299/itjrd.2023.12690

Accuracy Potential of The Convolutional Neural Network (CNN) in Recognizing Traditional Clothing

Herwinsyah, Dery Yuswanto Jaya
Department of Science and Technology, Sunan Kalijaga State Islamic University
herwinsy@gmail.com, dery.education@gmail.com

Article Info

Article history:

Received Apr 14, 2023
Revised Aug 08, 2023
Accepted Dec 08, 2023

Keyword:

Accuracy Potential
Convolutional Neural Network
Traditional Clothing
Recognizing

ABSTRACT

The diversity of cultures in Indonesia is proof that Indonesia is a country that is rich in cultural diversity. Many foreign tourists who want to know about culture in Indonesia are not directly proportional to the media to introduce culture in Indonesia. Therefore, this study aims to classify images of traditional clothing by detecting images of traditional clothing sent to the application to determine the name of the traditional soldier. These images will be converted into vectors and processed to find the closest similarity level. The Deep Learning method which currently has the most significant results in image recognition is the Convolutional Neural Network (CNN). The analysis carried out resulted in an accuracy of 0.7934 with an epoch of 20 and a data set of 700 data. The accuracy value is 0.7934 which is a large enough number to determine the correct classification of image objects. This is proven by testing on 10 different images and only 1 image is inaccurate with 90% accuracy.

Corresponding Author:

Herwinsyah
Department of Science and Technology
Sunan Kalijaga State Islamic University
Jl. Marsda Adisucipto Yogyakarta 55281, Yogyakarta, Indonesia
Email : herwinsy@gmail.com

1. INTRODUCTION

Indonesia is the largest archipelagic country in the world consisting of various ethnicities and ethnic groups. According to the 2010 BPS census on the official website of the Indonesian Information Portal, there were 1,340 ethnic groups spread across 34 provinces in Indonesia [1]. The diversity of these ethnic groups makes Indonesia have a diversity of regional languages, customs, traditional dances and a diversity of traditional clothing which are still preserved today as the identity of the Indonesian state. All these traditions have been passed down from generation to generation. The many traditions in each region make Indonesia have a variety of traditional clothes that are used for different events such as wedding ceremonies, funeral ceremonies, traditional dance costumes and several other traditional ceremonies [2].

This great diversity sometimes makes Indonesian citizens themselves ignorant of other cultures in other regions, so it takes the application of technology to accommodate this. The application of human knowledge to computers is usually done with several algorithms[3] in machine learning[4], deep learning and so on. Today, machine learning underpins many of the applications we use every day, from product recommendations to speech recognition, as well as some of the applications we don't use every day, including self-driving cars. This is the basis of a new approach to computing where we don't write programs but collect data. The idea is to automatically learn task algorithms from data[5]. Meanwhile, Deep learning[6] is a form of machine learning that allows computers to learn from experience and understand the world in terms of a hierarchy of concepts. Because the computer gains knowledge from experience, it is not necessary for a human computer operator to formally specify all the knowledge the computer needs. The hierarchy of concepts allows the computer to learn complex concepts by building them from simpler concepts; a graph of these hierarchies would be many layers deep[7].

Convolutional Neural Network (CNN)[8], architecture has been around for over two decades compared to other neural network models such as multi-layer performance perceptron (MLP)[9], CNN designed to accommodate multiple arrays as input, and then process the input using convolution Operators in the local field by mimicking eye perception Picture. Therefore, it shows excellent performance in solving computer vision problems such as image classification[10], recognition and understanding. He also effective for areas such as speech Recognition that requires the spectral representation of correlated speech Presentation, physical VLSI design, multimedia Compression compared to traditional DCT transformation and pressure measurement methods, and cancer detection from a changing range of conditions Pictures. There were also many top players Fever playing a Go match with alphaGo recently huh implemented CNN.

1.a 1.b 1.c

1.d 1.e 1.f

Figure 1. Traditional Clothing Images of Each Dance

Several studies in the field of machine learning and deep learning that have been carried out previously raise various themes. For example, Latah conducted research to identify human actions based on existing videos [11]. In this study, he used the 3D Convolutional Neural Networks (CNNs) method. In another study conducted by Hanafi, he developed a system to capture the contextual meaning of consumer product reviews to generate item latent factors combined with probabilistic matrix factorization [12]. In his research he uses the Dynamic Convolutional Neural Network (DCNN)[13] method and the evaluation results with Root Mean Squared Error (RMSE). The research conducted by Balasundaram on the detection of breast cancer cells by analyzing histopathological images qualitatively is clear [14]. The study used the Deep Convolutional Neural Network method to detect suspicious tissue growth and the exact nature of suspected cancer cells can be confirmed through histological assessment. Histology, in a general sense, is the microscopic analysis of living tissue. More than 70% of disease, especially tumor cell, is first diagnosed by histopathology.

Research on the potential accuracy of Convolutional Neural Networks (CNN) in recognizing traditional Indonesian clothing remains limited. Prior investigations have predominantly concentrated on object recognition within images. However, the intricate and multifaceted nature of traditional clothing, as a vital component of Indonesia's diverse cultural heritage, has yet to be comprehensively studied. This study seeks to address this gap by delving into the potential accuracy of CNN in recognizing traditional Indonesian attire. Through the application of CNN methodology to clothing recognition, we aspire to bolster the advancement of image recognition technology. This, in turn, can facilitate a deeper comprehension and heightened appreciation of the cultural opulence woven into Indonesia's traditional clothing.

This research aims to identify traditional dances based on the traditional clothes used. This identification utilizes the Convolutional Neural Network (CNN) method by detecting images of traditional clothing sent to the application to determine the name of the traditional dance. These images will be converted into vectors and processed to find the closest level of similarity. There are many traditional dances that exist throughout Indonesia, in this study we only focused on 3 traditional dances, namely the Saman dance from Aceh, the Baksa Kembang dance from Banjarmasin, South Kalimantan and the Jaipong dance from West Java.

2. RESEARCH METHOD

The method used in this research involves collecting image data, processing the image data by standardizing the size and properties of the images, dividing the data into two parts: training data and testing data, then creating a machine learning model, followed by training and validating the model, conducting testing, and evaluating the results. The entire process is illustrated in Figure 2 below

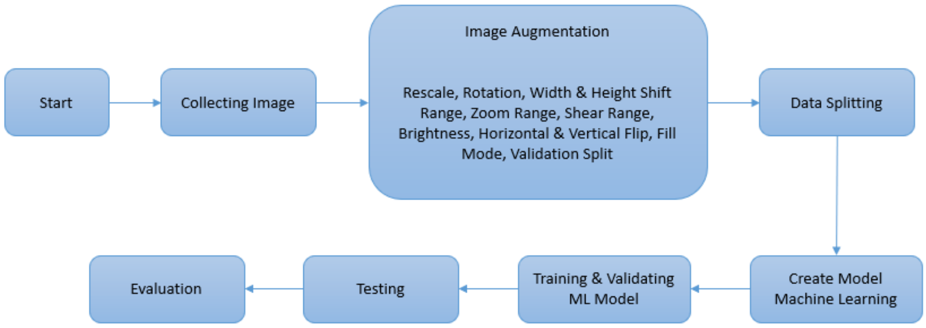


Figure 2. Research Flowchart

2.1. Collecting Images

Machine Learning development relies on data for training machine learning models. In this study, a dataset consisting of images from three dance categories, namely Jaipong Dance, Baksakambang Dance, and Saman Dance, is used. Each of these categories contains an equal amount of data. To address this issue, the author utilizes image augmentation techniques to augment the dataset[15].

The image data used in this research is processed to standardize the size and properties of the images, and then divided into two sets: training data and testing data. A machine learning model is created using this dataset, and the model is trained and validated. Subsequently, testing is conducted to evaluate the performance of the model. The results of this evaluation are used to assess the accuracy and effectiveness of the model in recognizing traditional costumes used in Jaipong Dance, Baksakambang Dance, and Saman Dance.

2.2. Image Augmentation

Image augmentation is a method used to augment data for training machine learning models without the need to search for new data. The principle used by image augmentation is to duplicate an existing image with a slight variation so that the image looks like a new image to the machine. The variations are as follows.

a. Rescale, meaning that the image will be re-scaled before being inserted into the next process. the goal is not to scale the images so large as to overload the machine learning model.
b. Rotation, means that the image will be rotated to generate new data to train the model
c. Width and height shift range is a process where the image will be shifted vertically and horizontally to produce new data. To better understand this process can be seen in Figure 3.

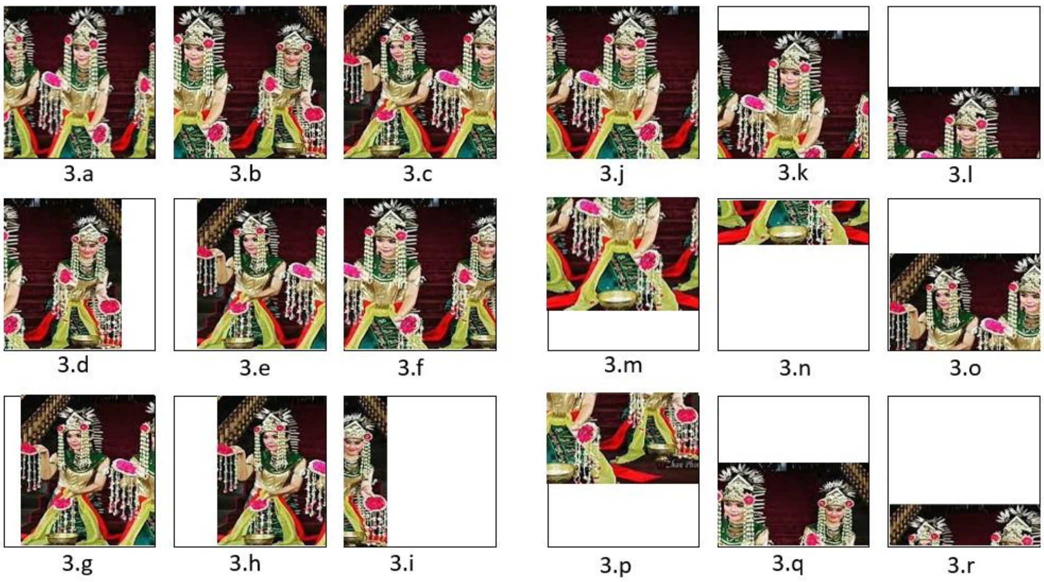


Figure 3. Width & Height Shift Range Examples

d. Zoom range, meaning that the existing image is enlarged and reduced randomly to generate new data.
e. Shear range, meaning that the image will be tilted to produce a new image shape. This technique is different from rotation, in the shear range, the machine will change one axis and stretch the shadow at a certain angle which is called the shear angle. This creates a kind of 'stretch' in the image, which is not visible in rotation. This process can be seen in Figure 4.
f. Brightness, means that the existing image data will be changed to change the brightness level to produce new image data.
g. Horizontal and Vertical flip, meaning that the image data will be flipped vertically and horizontally to produce new data.
h. Fill mode, meaning that the existing image data can be changed according to the fill mode we choose. some options of fill mode are nearest, reflect, wrap, constant.
i. validation split, in this parameter we can set the division of the dataset used as data validation.

Figure 4. Shear Range Examples

2.2. Data Splitting

During the development of machine learning models, it is essential to split the augmented dataset into training data and validation data. This step allows for evaluating the performance and accuracy of the model. Typically, the training data is larger than the validation data to ensure that the model learns from a diverse set of examples.

Next, the machine learning model is created using the training data. The model is then trained using various algorithms and techniques, depending on the specific requirements of the research. This training process involves optimizing the model's parameters and adjusting its weights to achieve the best possible performance.

2.2. Creating Machine Learning Model

The sequential CNN (Convolutional Neural Network) is the chosen machine learning model for this research. CNNs offer several advantages, such as the ability to automatically extract significant features from each image without human intervention. This feature extraction capability makes CNNs ideal for image processing tasks as they can learn to recognize patterns and structures within the images on their own.

Moreover, CNNs are known for their efficiency in terms of memory and complexity. Compared to other neural network methods, CNNs require fewer parameters, making them computationally efficient. This makes them well-suited for large-scale datasets and real-time applications where efficiency is crucial.

The sequential CNN model used in this research has been trained and optimized using the augmented dataset and validated with the validation data. The results obtained from testing the model on unseen data will be used to evaluate its performance and effectiveness. The advantages of CNNs, such as their ability to automatically extract features and their efficiency in terms of memory and complexity, make them a suitable choice for this research and contribute to the overall success of the machine learning development process.

2.2. Model Training And Validation

The training process for the machine learning model begins at this stage using the available dataset. The model undergoes repeated training iterations until it achieves the desired accuracy, which can be controlled through the epoch parameter. During training, the model studies the dataset and learns to identify patterns from the existing images. This enables the model to extract patterns from new images and match them with the learned patterns from the training data, allowing the model to make predictions on the type of dance from the new images.

The training process involves the model adjusting its internal parameters based on the dataset, which helps it learn the underlying patterns and relationships within the data. This allows the model to develop a deeper understanding of the dance categories and their distinguishing features.

Once the model has completed the training process, it will be able to make predictions on new, unseen images with a certain level of accuracy. The accuracy of the model depends on the quality and size of the dataset used for training, as well as the hyperparameters and architecture of the model. The training process is a critical step in the machine learning development process as it empowers the model to make accurate predictions on new data based on the patterns it has learned from the training dataset.

2.2. Testing

After the training process, the next step is to evaluate the performance of the model. This is done by testing the trained model with a set of new images that were not included in the training dataset. Typically, around 10 new dance images are used for this testing process. The purpose of this testing is to assess how accurately the model can make predictions on unseen data. During the testing process, the trained model will be presented with the new images and will attempt to predict the type of dance for each image. The accuracy of the model's predictions is then evaluated by comparing the predicted dance type with the actual dance type of the images.

The testing results provide valuable insights into the performance of the model and its ability to generalize to new, unseen data. If the model achieves a high accuracy rate in predicting the dance types of the new images, it indicates that the model has learned the underlying patterns from the training dataset and can make accurate predictions on unseen data. However, if the accuracy is low, further adjustments and fine-tuning may be needed to improve the model's performance.

2.2. Evaluation

After completing the testing process, the author conducts a comprehensive evaluation of the model's performance. This evaluation involves analyzing the accuracy of the model's predictions, identifying any errors or limitations, and assessing the overall effectiveness of the model in achieving its intended goals. The results of this evaluation provide valuable insights for future model development.

Based on the evaluation results, the author can draw conclusions and make recommendations for improving the model in future iterations. This may involve refining the training dataset, adjusting the model's parameters, or incorporating additional techniques to enhance its performance. The evaluation process is crucial in identifying areas for improvement and guiding the author in making informed decisions for further model development.

By learning from the strengths and weaknesses of the current model, the author can make informed decisions to improve the model's accuracy and effectiveness, ultimately contributing to the advancement of the field of dance recognition through machine learning.

3. RESULTS AND ANALYSIS

In order to implement the CNN algorithm, we are using keras from tensorflow in the python package. Keras is an open source software library that provides a Python interface for artificial neural networks. We build our models using some of the layers functions built into keras so that our models have greater simplicity and flexibility. It describes the building blocks of CNN as an easy-to-use python library, provides routines for calculating linear convolutions with filter banks, feature collection, and more. In this way, keras enables rapid prototyping of the new CNN algorithm; at the same time, it supports efficient computation on CPU and GPU which makes it possible to train complex models on large data sets. coupled with the existence of google collab we can make CPU Memory more efficient.

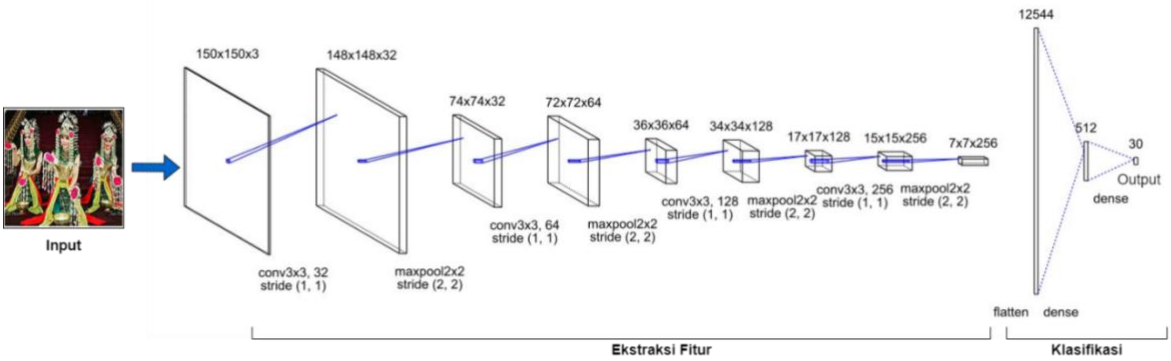


Figure 5. Architecture of our CNN Algorithm

Convolutional Neural Networks (CNN) are today's state-of-the-art architecture for image classification tasks. As shown in Fig.4 our proposed 2-D Convolutional Neural Network (CNN) model was designed using Keras for the image recognition task of traditional dances. Preparing the data is the first step of our approach. Before building the network, we need to prepare the train data and test data, aggregate the data, combine the labels, and reshape it to the appropriate size. We store datasets of normalized data (single precision and zero mean), labels, and other miscellaneous (meta) information. As shown in Figure 1. For modeling, input is needed in the form of a dataset.

The dataset used in this experiment is in the form of 2100 images of traditional dances with 3 classes namely "baksakambang", "saman" and "jaipong". each class amounted to 700 images. Before the data is processed to the next stage, pre-processing of data is carried out, namely performing image augmentation using the imageDataGenerator() function with the aim of increasing sample data. The configurations used are Rescale, rotation range, width_shift_range, height_shift_range, zoom_range, shear_range, brightness_range, vertical_flip, fill_mode and validation_split while the values are respectively 1./255, 30, 0.2, 0.2, 0.4, 0.2, [0,1] , True, True, nearest, and 0.2

Out of 80% of the total dataset, 1,680 labeled images were used as training data, while the remaining 20%, namely 420 images, were used as test data. We have conducted several tests with the highest performance values on the division of 80% training data and 20% test data.

Building and compiling the model is the fourth and fifth step. To create a CNN we have to initialize the CNN architecture and then define important initialization parameters e.g. batch size, number of epochs, learning rate etc.

```
[ ] model = tf.keras.models.Sequential([
      tf.keras.layers.Conv2D(32, (3,3), activation = 'relu', input_shape= (150,150,3)),
      tf.keras.layers.MaxPooling2D(2,2),
      tf.keras.layers.Conv2D(64,(3,3), activation= 'relu'),
      tf.keras.layers.MaxPooling2D(2,2),
      tf.keras.layers.Conv2D(128,(3,3), activation= 'relu'),
      tf.keras.layers.MaxPooling2D(2,2),
      tf.keras.layers.Conv2D(256,(3,3), activation= 'relu'),
      tf.keras.layers.MaxPooling2D(2,2),
      tf.keras.layers.Flatten(),
      tf.keras.layers.Dense(512, activation= 'relu'),
      tf.keras.layers.Dropout(0.1),                                   #m
      tf.keras.layers.Dense(3, activation= 'softmax')
    ])
```

Figure 6. CNN Layer in Python Code

In initializing the CNN architecture, you can use the Sequential function provided by Keras tensorflow. the configurations used are Conv2D, Maxpooling2D, Flatten, Dense and Dropout The general modeling sequence is in Fig. 4. Input in the form of images will be carried out convolutionally with several configuration parameters. The parameters and configuration values are the activation function in the form of a Rectified Linear Unit (ReLU), a filter value of 32 and a filter dimension of 3x3, the input shape and type are RGB or Black/white. Then, the convolved image will be pooled to reduce the size of the image as much as possible. Here we also try to reduce the total number of nodes in the next layer. Pooling in the form of a 2×2 matrix as a minimum pixel loss and precise region where features are allocated. Then do the same steps by multiplying the filter value by two until the filter value becomes 256. Next is the flattening process. The pooling data we have is in the form of a 2-dimensional array which is then converted into a single vector one-dimensional data. Next is the dense process. Dense is a function to add fully connected layers. units indicates the number of nodes that must exist in the hidden layer, the value is between the number of input nodes and output nodes. Meanwhile, we use RelU for the activation function. To reduce the possibility of overfitting and speed up the learning process, Dropout is applied with inputs of 0 and 1. Dropout is a technique in which randomly selected neurons are ignored during training.

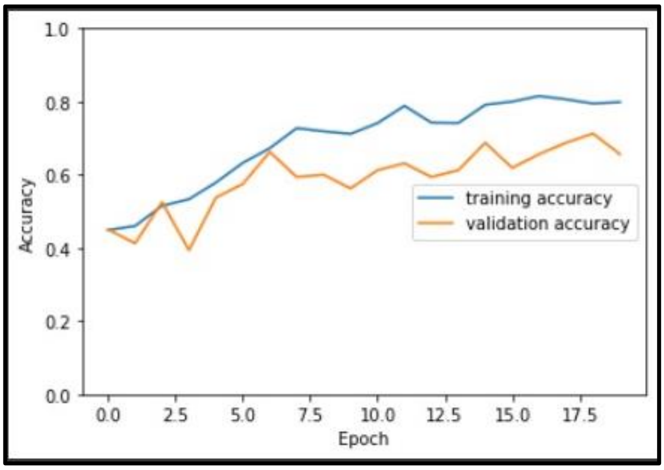


Figure 7. Comparison of Train and Validation Accuracy

Next, initialize the output layer. The output layer that we define is in the form of one node because this classification is included in the binary classification. We only use one node, and the soft max activation function on the last layer. The process of building the CNN model is complete, then it will compile the model.

The batch size determines the number of samples for the training phase of the CNN. The CNN will process all the training data, but only in increments of the specified batch size. We can use batch size for computational efficiency, and its value will be dependent on the user's available hardware. An epoch is a successful forward pass and a backward pass through the network. It's usually beneficial to set its value high and then to reduce it once if one is satisfied with the convergence at a particular state (chosen epoch) in the network. Learning rate is a very sensitive parameter that pushes the model towards convergence. Finding its best value will be an experimental process unless one invokes more powerful techniques such as batch normalization. In our experiment, we use batch size 20,several epochs 25 and learning rate 0.0001 for maximum accuracy.

We build our CNN by creating each layer individually as shown in fig 5. Afterward, we will invoke objective and error layers that will provide a graphical visualization of the training and validation convergence after completing each epoch. Keras initializes the weights by using Gaussian distribution.

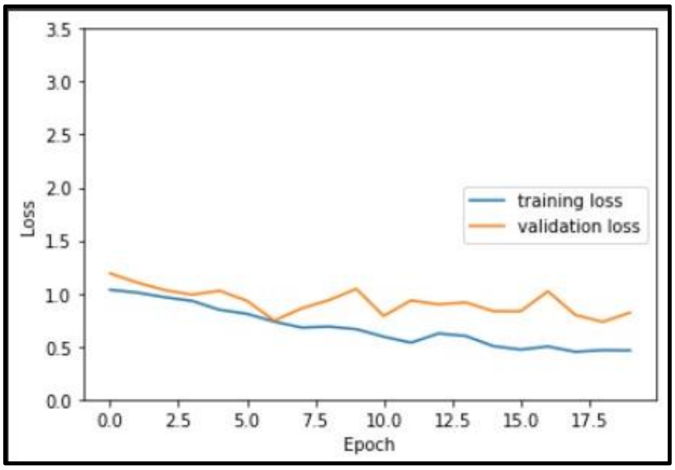


Figure 8. Comparison of Train and Validation Loss

The sixth step is the training and evaluating the model. Training a CNN requires computing the derivative of the loss concerning the network parameters. We calculate the derivatives using an algorithm called back propagation which is a memory-efficient implementation of the chain rule for derivatives. We built the model and performed a random gradient descent training according to the Stochastic Gradient Descent (SGD) training algorithm. We have used SGD training algorithm to adjust the weight of the connection between neurons so that the loss reaches a minimum value or stops after several epochs.

Table 1. Comparison Training and Validation

Epoch	**Comparison Value**			
	Training Loss	***Training Accuracy***	***Validation Loss***	***Validation Accuracy***
1	31,11%	40,87%	31,40%	49,38%
2	27,78%	52,13%	29,32%	44,37%
3	26,34%	55,37%	27,54%	58,75%
4	24,97%	61,61%	28,24%	49,38%
5	23,48%	62,37%	24,01%	62,50%
6	20,27%	69,01%	25,97%	50,00%
7	19,34%	71,13%	18,94%	73,12%
8	21,41%	67,47%	30,45%	52,50%
9	19,38%	70,79%	30,73%	53,75%
10	17,79%	73,50%	23,15%	61,87%
11	16,98%	76,02%	26,38%	62,50%
12	13,87%	79,34%	20,29%	68,75%
13	17,77%	74,23%	25,62%	59,38%
14	17,12%	74,12%	26,15%	61,25%
15	14,38%	79,12%	23,80%	68,75%
16	13,43%	80,00%	23,77%	61,87%
17	14,27%	81,50%	29,17%	65,62%
18	12,84%	80,61%	22,77%	68,75%
19	13,29%	79,46%	20,91%	71,25%
20	13,22%	79,85%	23,40%	65,62%

We can start CNN training by providing the training data, the model built, and the current data set. When training a CNN, only the data specified for training plays a role in minimizing errors in the CNN. We feed the training data through the network for both forward and backward feeds. Validation data is only used to see how CNN responds to new similar data. We do not use validation data to train the network. After that, we store the CNNs that are trained and ready for the testing phase.

During training, there is a comparison of the accuracy values of training and validation shown in Figure 6. The graph shows that the validation value does not always exceed the value of training in each epoch. The highest validation accuracy value is 73,12% while the highest training accuracy value is 81,50%.

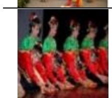
In addition to the accuracy value, there is also a loss value in training from training and validation which is shown in Figure 7. The graph shows that the validation loss value is never smaller than the training loss value in each epoch. The lowest validation loss value is 18,94% while the lowest training loss value is 12,84%. The detailed comparison of the accuracy and loss values for each epoch can be seen in Table 1.

Finally, by using the testing data, we can evaluate our model. The following are an example of classification outputs from the model with CNN on the Traditional Indonesian Dance data. The correct guess is shown in the 1st to 5th pictures and the 7th to 10th pictures in table 2 while the wrong guess is shown in the 6th picture

We can determine the test cases that show failed predict the image. The model can't predict some images because of limitations in the input of standard data images. Moreover, missing pixels caused by image compression and image sharpness problems are also reasons for misclassification. The fourth and final step is to save the model in the disk for reuse. We store the trained model in a file format. Hence the saved model can be reused later or easily ported to other environments too. Among 10 test cases, our model misclassifies total of 1 images shown in Table 2. Test accuracy 90% implies that the model is trained well for prediction. Training set size affects the accuracy increases as the number of data increases. The more data in the training set, the smaller the impact of training error and test error and ultimately the accuracy can be improved.

Apart from that, we also use the confusion matrix provided by Python to test the model. We use 3 test parameters for both true labels and predicted labels, namely "Saman", "Baksa Kambang" and "Jaipong". The amount of test data used was 2100 data with the composition of saman, baksa kambang, and jaipong respectively 700, 714 and 686. The test results can be seen in Figure 8, the model can guess correctly on the saman dance image as many as 580, "Baksakambang" as many as 580, while the most in Jaipong is 614. And the model cannot guess correctly as many as 326.

Table 2. Testing Image

Epoch	**Label**	
	Dataset	***Predicted***
	Saman	Saman
	Jaipong	Jaipong
	Baksa Kambang	Baksa Kambang
	Jaipong	Jaipong
	Saman	Saman
	Baksakambang	Saman
	Saman	Saman
	Jaipong	Jaipong
	Baksa Kambang	Baksa Kambang
	Saman	Saman

The accuracy value is the ability of the model to correctly guess the input given. The accuracy value is obtained from the number of correctly guessed models divided by the total number of trials. So that the accuracy value of the model is 84.48%.

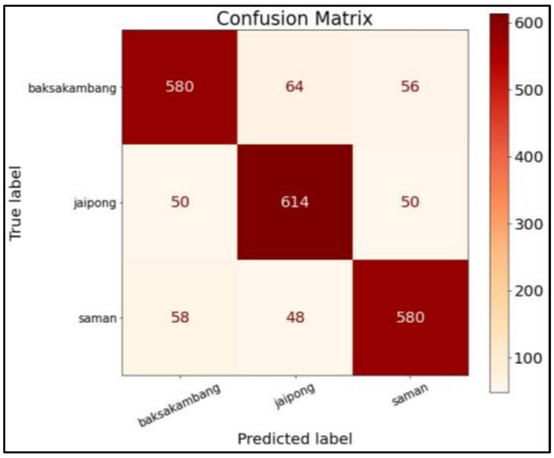


Figure 8. Confusion Matrix of Model CNN

4. CONCLUSION

The utilization of the Convolutional Neural Network (CNN) as a classification method yields a noteworthy accuracy rate of 0.7934. This substantial figure plays a pivotal role in effectively discerning the accurate classification of object images. This assertion finds validation through rigorous testing on a diverse set of 10 distinct images, wherein a mere singular image exhibited inaccuracy, albeit with a commendable accuracy rate of 90%. It is imperative, however, to acknowledge that this accomplishment is attained within the confines of a restricted dataset.

The inherent limitation of the dataset employed underscores the pressing need for an augmentation in the volume of data available. Such augmentation is anticipated to play an instrumental role in further heightening the accuracy levels achieved by the CNN-based classification methodology. By integrating a more expansive and diverse dataset, the potential for a substantial improvement in accuracy becomes discernible. Thus, the trajectory of future research endeavors should be oriented towards procuring and incorporating a richer dataset, thereby unlocking the latent capability of the CNN method to attain even more accurate and reliable classifications.

REFERENCES

[1] Wikipedia, "Indonesia --- Wikipedia{,} Ensiklopedia Bebas." 2023, [Online]. Available: https://id.wikipedia.org/w/index.php?title=Indonesia&oldid=23259874.

[2] Admin, "Suku Bangsa," *Indonesia.go.id*, 2017. https://indonesia.go.id/profil/suku-bangsa/kebudayaan/suku-bangsa (accessed Apr. 15, 2023).

[3] "Introduction to Algorithms, fourth edition - Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein - Google Books." (accessed Apr. 15, 2023).

[4] M. Batta, "Machine Learning Algorithms - A Review," *Int. J. Sci. Res.*, vol. 18, no. 8, pp. 381–386, 2018, doi: 10.21275/ART20203995.

[5] E. Alpaydın, *Machine Learning:The New AI*. 2016.

[6] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523–3542, Jul. 2022, doi: 10.1109/TPAMI.2021.3059968.

[7] Q. Zhang, M. Zhang, T. Chen, Z. Sun, Y. Ma, and B. Yu, "Recent advances in convolutional neural network acceleration," *Neurocomputing*, vol. 323, pp. 37–51, Jan. 2019, doi: 10.1016/J.NEUCOM.2018.09.038.

[8] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, “A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.

[9] Y. Jusman, I. M. Firdiantika, D. A. Dharmawan, and K. Purwanto, “Performance of multi layer perceptron and deep neural networks in skin cancer classification,” *LifeTech 2021 - 2021 IEEE 3rd Glob. Conf. Life Sci. Technol.*, pp. 534–538, Mar. 2021, doi: 10.1109/LIFETECH52111.2021.9391876.

[10] T. He, Z. Zhang, H. Zhang, Z. Zhang, J. Xie, and M. Li, “Bag of Tricks for Image Classification with Convolutional Neural Networks.” pp. 558–567, 2019.

[11] M. Latah and M. Latah, “Human action recognition using support vector machines and 3D convolutional neural networks,” *Int. J. Adv. Intell. Informatics*, vol. 3, no. 1, pp. 47–55, Mar. 2017, doi: 10.26555/ijain.v3i1.89.

[12] Hanafi, N. Suryana, and A. S. H. Basari, “Dynamic convolutional neural network for eliminating item sparse data on recommender system,” *Int. J. Adv. Intell. Informatics*, vol. 4, no. 3, pp. 226–237, Nov. 2018, doi: 10.26555/ijain.v4i3.291.

[13] Y. Chen, X. Dai, M. Liu, D. Chen, L. Yuan, and Z. Liu, “Dynamic Convolution: Attention Over Convolution Kernels.” pp. 11030–11039, 2020.

[14] D. Nurtiyasari, “Daftar Isi: THE APPLICATION OF DEEP NEURAL NETWORK FOR BREAST CANCER CLASSIFICATION,” 2018. https://onesearch.id/Record/IOS1972.article-22237/TOC (accessed Apr. 15, 2023).

[15] I. Kostrikov, D. Yarats, and R. Fergus, “Image Augmentation Is All You Need: Regularizing Deep Reinforcement Learning from Pixels,” Apr. 2020, Accessed: Apr. 15, 2023. [Online]. Available: https://arxiv.org/abs/2004.13649v4.