Journal of Student Research Exploration

https://shmpublisher.com/index.php/josre

p-ISSN 2964-8264 | e-ISSN 2964-1691

Integrating Convolutional Neural Network Features Extraction with Extreme Learning Machines for Image Classification of Pandava Characters in Wayang Kulit

Alfiatul Fitria[1], Budi Prasetiyo[2]

[1,2] Department of Information System, Universitas Negeri Semarang, Indonesia

Article Info

Article history:

Received July 10, 2024
Revised April 2, 2026
Accepted April 14, 2026

Keywords:

Wayang kulit
Extreme learning machines
Image classification
Convolutional neural networks
Cultural heritage preservation

ABSTRACT

This research focuses on the utilization of image processing techniques—the Convolutional Neural Networks (CNNs) and Extreme Learning Machine (ELMs)—to classify the characters of Wayang Kulit automatically. The Pandava characters or casts are classified in accordance with the characters from traditional Indonesian puppets, commonly known as shadow puppets. The focus is to introduce such rich cultural heritage to younger generations by using technology. Prior research has utilized classification of characters using Convolutional Neural Networks (CNNs), Extreme Learnings Machines (ELMs), and Support Vector Machines (SVMs), which led to varied accuracy levels. In our subsequent experiments, three proposed models, with varying underlying model assumptions, were evaluated. The proposed models generated moderate accuracies ranging from 39 to 52%. The results suggest that our models have room for further development to enhance their performance. Strategies from parameter tuning to the in-depth analysis of the confusion matrix are discussed. Above all, the research is geared towards ensuring the appreciation and preservation of traditional cultural heritage in this digital era.

1. Introduction

Wayang kulit is a traditional Indonesian art form originating from the Javanese ethnic group. It is recognized as a visual art form that embodies numerous values and profound philosophies. The characters in wayang kulit represent Hindu and Javanese mythology, crafted using dried and processed buffalo skin to hold colours and textures [1]. The early

[1] ***Corresponding Author:***
Budi Prasetiyo,
Department of Information System,
Universitas Negeri Semarang,
Sekaran, Gunung Pati, Semarang, Central Java 50229, Indonesia.
Email: bprasetiyo@mail.unnes.ac.id
DOI: https://doi.org/10.52465/josre.v4i1.407

existence of wayang kulit has been affirmed by the discovery of an inscription from King Balitung (889-911 After Century) [3]. To this day, wayang kulit has been listed by The United Nations Educational, Scientific and Cultural Organization (UNESCO) as a Representative List of the Intangible Cultural Heritage of Humanity in 2003. Wayang kulit is commonly used as a performance medium for enacting stories, primarily derived from the Ramayana and Mahabharata epics [2]. The performances, accompanied by messages about the meaning of life and how to live it, have also made wayang known as an ancient educational medium. This is realized through the variety of characters and traits in wayang kulit that represent common human characteristics, such as wisdom, generosity, greed, and humor.

There are obstacles in the wayang kulit's preservation because public interest is waning, particularly among younger people. The Central Bureau of Statistics for Social and Culture's 2021 data [4] illustrates this, revealing a 23.06% drop in the number of children five years of age and older who attended art exhibitions between 2018 and 2021. The state and its inhabitants share responsibilities for preserving cultural assets, including wayang kulit. One way to instil interest in wayang kulit among the younger generation is by introducing them to the characters in wayang kulit. One of the most popular characters in wayang is the Pandava from the Mahabharata epic. The Mahabharata story itself is widely known among the Javanese people alongside the influence of Hindu culture, which has affected many aspects of Javanese culture. Linguistically (in Sanskrit), Pandava means the five sons of Pandu, the King of Hastinapura [5]. These five sons are known as Puntadewa, Bima, Arjuna, Nakula, and Sadewa. The life stories of the Pandava have been frequently staged and adapted into wayang kulit performances because each character's story carries messages and noble values appreciated by the audience.

The rapid development of technology is reaching various fields. Advanced technology is expected to introduce cultural heritage, such as wayang kulit, particularly the Pandava, to the younger generation. One branch of technology that has developed well is image processing. Image processing involves stages of creation, processing, and analysis to extract information contained in images [6]. Image classification is a form of image processing, which groups images based on their elements, where images are categorized according to shared traits [7]. Popular methods for image classification include K-Nearest Neighbour(KNNs), Convolutional Neural Networks (CNNs), and Support Vector Machines (SVMs). Previous studies have applied these methods to classify wayang kulit images.

Transfer learning has been an important advancement in machine learning since it makes it possible to apply knowledge from one domain to another, improving the effectiveness and performance of models. This method makes use of pre-trained models on huge datasets to boost generalization, making it especially useful when working with sparse data in the target domain. Transfer learning has been shown to be useful in several applications, including speech recognition, natural language processing, and picture classification [20]. Researchers have shown that fine-tuning pre-trained models yields better results than starting from zero with model training.

Convolutional Neural Networks (CNNs) are essential for applying transfer learning, particularly in computer vision applications. CNNs are made to use backpropagation to automatically and adaptively learn the spatial hierarchies of features from input images. Because of these networks' capacity to capture local dependencies and scale invariance, significant progress has been made in picture recognition tasks [21], [22]. Moreover, the integration of Extreme Learning Machines (ELM) with CNNs has further enhanced model performance by offering a robust learning mechanism with faster training times and better generalization capabilities [23]. ELMs, known for their single hidden layer feedforward neural network architecture, provide a unique advantage in handling non-linear data efficiently, thereby complementing the strengths of CNNs in transfer learning scenarios [24], [25].

Pandava character classification has been previously conducted by Wiwit Supriyanti and Dimas Aryo Anggoro using the CNN (Convolutional Neural Network) method. This study achieved the highest accuracy rate of 93% with a loss function of 0.197, tested using 10% of the data for testing. Another study by Fryda Fatmayati et al. classified wayang kulit character types using the ELM (Extreme Learning Model) method, yielding an accuracy of 81%. This study used extracted results from Operation Morphology as input data. The SVM (Support Vector Machine) method was also employed by Muhathir et al. for the same subject, wayang kulit classification. In this study, features were extracted using Generalized Linear Model Classifiers (GLMCs), resulting in an accuracy rate of 83.2%.

The distinction of this research from previous studies lies in its combination of CNN for feature extraction and ELM as the classification machine. ELM is a superior classification method for image classification. Compared to other more traditional classification methods, ELM offers advantages in speed, which is more ideal [11]. Extreme Learning Machine (ELM) assigns random values to the weights between the input and hidden layers, as well as the biases in the hidden layer. These parameters remain fixed during training [12]. For this research, the ELM model will be combined with CNN-based feature extraction. Compared to traditional methods for extracting bearing fault characteristics, deep neural networks, such as Convolutional Neural Networks (CNNs), can automatically extract intrinsic features without requiring expert knowledge[13].In such hybrid classification methods, the choice of classification method is crucial. The superior feature extraction capability of CNN can be maximized with the appropriate classification model. The optimization of CNNs with ELMs has been previously applied by Yusiong J. P. T. to classify tomato ripeness. The ELM model achieved an accuracy score of 96.67%. In waste classification by Yusiong J. P. T., the CNN-ELMs combination model recorded an accuracy of 93.97%. Furthermore, this hybrid method has also been implemented in real-time parking space detection systems by Kaur R. et al., justifying the effectiveness of this model for classification problems. Based on previous research, ELMs will be used to optimize CNNs feature extraction for classifying wayang kulit characters, specifically the Pandava.

2. Method

This section will be explained how this research will be conducted. A dataset containing each character in the Pandava as label data will be required for the method to be developed. The following image will show the classification modelling flow along with evaluation metrics. The research process will consist of three main stages: data preparation, modelling, and evaluation. The work flow of this research can be seen in **Kesalahan! Sumber referensi tidak ditemukan.Kesalahan! Sumber referensi tidak ditemukan.**.

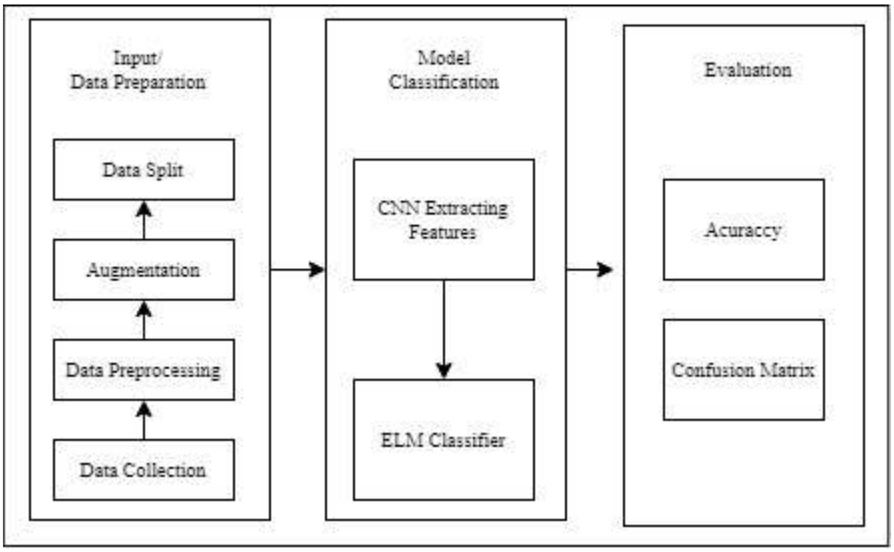

Figure 1. Research model and flow

Data preparation are processes that must be undergone before inputting data into the training model. It ensures that the data inputted into the training model meets the requirements of the model. This process has to be done to ensure that the model is able to function properly with optimal accuracy results. The classification model involves the process of training machine models according to the proposed method, namely CNN-ELMs. The evaluation stage is a review of the model's performance. Evaluation metrics used here include accuracy and confusion matrix consisting of precision, recall, f-1 score, and support. Conclusions regarding the performance of the constructed model will be drawn from the available evaluation metrics.

2.1 Data Collection

The dataset used consists of 400 images divided into 4 classes: Puntadewa, Bima, Arjuna, and Nakula_Sadewa. The combination of the Nakula_Sadewa characters is based on the consideration that they are depicted as twin brothers. The differences between them, such as characteristic traits and behaviours, are not measured in this research. The data was obtained through online scraping using Python tools and the bing_image_downloader library. The images were then stored in four different folders corresponding to the four distinct classes. The sample data excerpts obtained from the scraping process for each Pandava character can be seen in **Kesalahan! Sumber referensi tidak ditemukan.**.

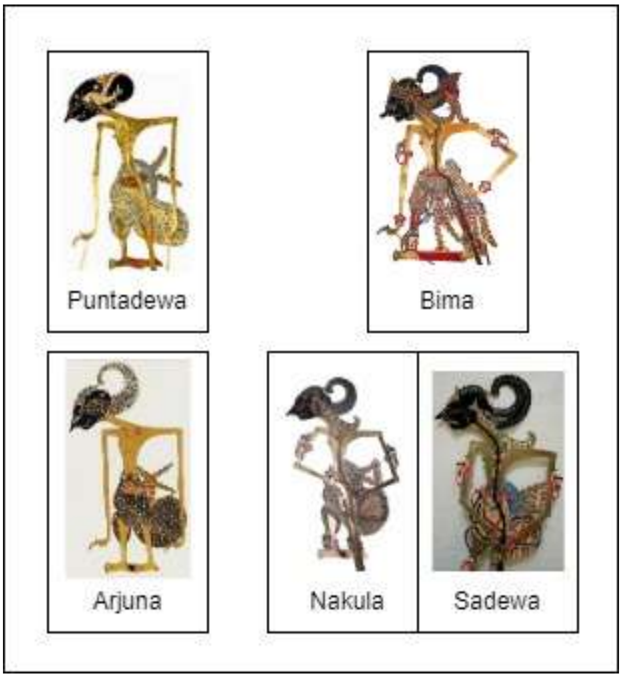

Figure 2. Data sample

2.2 Data Pre-Processing

Data preprocessing is the process of preparing data to ensure that the data format to be inputted meets the requirements of the training model. Data preprocessing aims to improve image quality, normalize images, and reduce noise or artifacts in the images [17]. It includes data filtering, resizing, cantering images, normalization, standardization, and conversion processes. Data filtering involves manually removing unnecessary data after scraping. Subsequently, the data is resized to have a uniform size and meet the model's requirements in this research, the pictures resized into 224 × 224. After resizing, the main training objects in the images are cantered to facilitate model performance. The images are then normalized by normalizing pixel values to the range of 0-1. Standardization reduces the mean and divides by the standard deviation of pixel values. Additionally, images are imported in RGBA format, which stands for Red, Green, Blue, and Alpha, which are the four channels used to define a colour in digital images and graphics. This process does not involve removing, standardizing, or deleting noise backgrounds, as the presence of image noise can help the model learn to capture the shapes of the main characters. The after-preprocessing data samples can be seen in **Kesalahan! Sumber referensi tidak ditemukan.**.

Figure 3. After pre-processing data sample

2.3 Data Augmentation

In this study, image augmentation is used as a technique to increase the quantity and variety of training data by applying random transformations to images. Augmentation aims to make the model more robust and capable of recognizing objects under various conditions. The augmentation techniques used include rotation, which rotates the image at random angles; flipping, which flips the image horizontally or vertically; zooming, which enlarges or reduces the image; shifting, which moves the image horizontally or vertically; brightness adjustment, which changes the brightness of the image; and shearing, which applies geometric distortion to the image. By introducing additional variations in the data, this augmentation helps prevent overfitting and improves the model's generalization ability, enabling the model to perform better on previously unseen data.

2.4 Data Split

Data split is the process of dividing data into training and testing datasets. This study will demonstrate model performance experiments based on different data splits. The first data split uses a ratio of 10% testing and 90% training, followed by 20% testing and 80% training, and finally 30% testing and 70% training data. Diverse data splits are expected to provide better evaluation analysis.

2.5 CNN Features Extraction and ELM Classification

A deep learning model termed Convolutional Neural Network (CNN) is especially made for recognizing and interpreting grid-based data, such pictures [17]. Generally, CNNs consist of three layers of artificial neural networks, including the flattening layer, pooling layer, and convolutional layer. In the convolutional process, a set of kernels is convolved along the input data, extracting local features from the input data [17]. In the pooling layer, the size of the image is reduced while the efficiency of the process is continuously improved [17]. After that, a feature map is created by applying an activation function to the values obtained by the convolutional operation [18].

Convolutional Networks are one of the most compatible methods for extracting features from images, which can be utilized for various machine learning tasks, especially classification and regression problems. ImageNet has provided pre-trained architectures for image classification tasks using Convolutional Network structures. One of the ImageNet architectures suitable for feature extraction is the Visual Geometry Group 16(VGG16) architecture. Visual Geometry Group 16(VGG16) evaluates neural networks by increasing the depth of images using very small convolutional filters, namely 3×3. Constructing such an architecture will enhance the quality of prior art configurations. By increasing the depth to 16-19 layers, approximately 138 trainable parameters can be obtained. VGG16 has become a popular algorithm for image classification and is easily adaptable for transfer learning. The architecture of VGG16 can be seen in **Kesalahan! Sumber referensi tidak ditemukan.**.

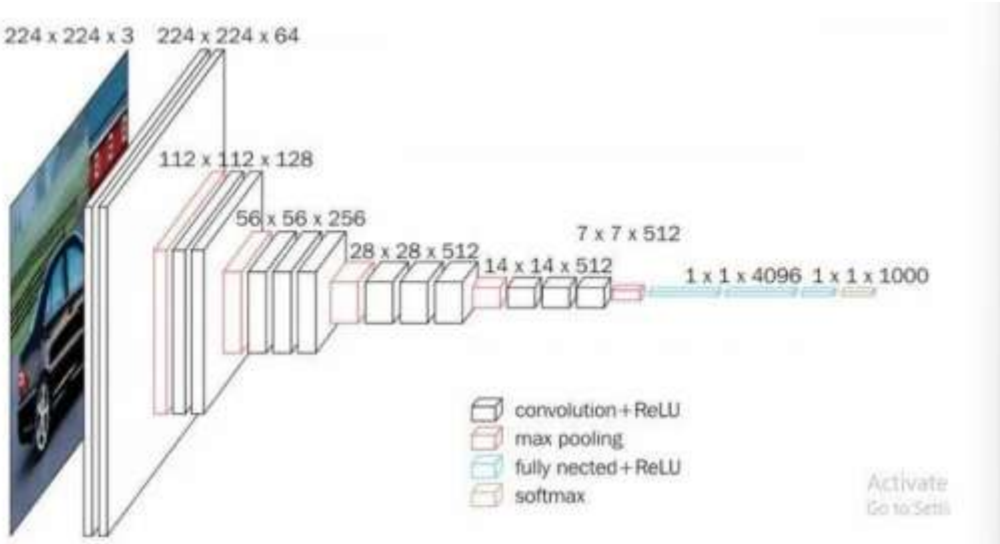

Figure 4. VGG16 features extraction

VGG16 is built using 26 layers with their own weights. This architecture comprises 13 Convolutional Layers, five Max Pooling layers, and three Dense layers. In total, there are 21 layers, but only 16 layers have weights. In this study, the VGG16 used will remove or not involve the Softmax layer, which contains neurons for classification. The Softmax layer will be replaced with hidden layers constructed based on the Extreme Machine Learning (ELM) architecture. The basic architecture of ELMs can be seen in **Kesalahan! Sumber referensi tidak ditemukan.**.

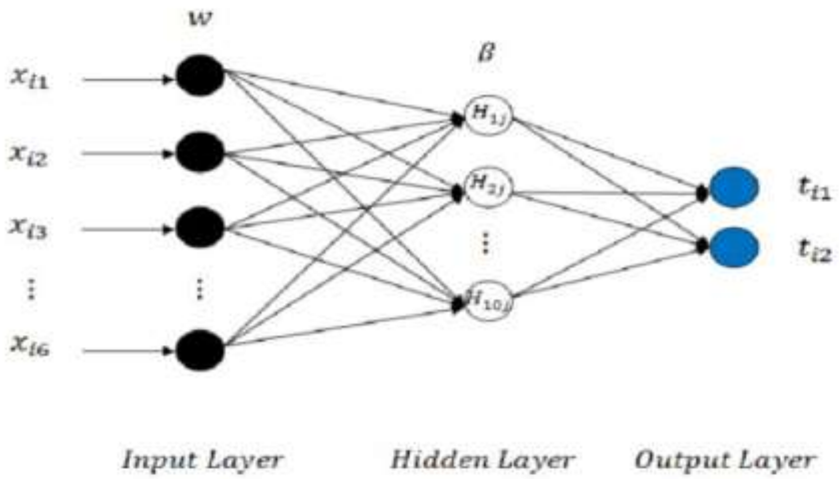

Figure 5. Basic architecture of extreme machine learning algorithm [19]

The ELM architecture consists of three layers: the input layer, hidden layer, and output layer. In this architecture, the input layer and hidden layer are connected by weighted vectors whose values are randomly determined. Biases connecting between nodes in the hidden layer are also randomly determined [19]. The output layer produces weights used to determine the class in the testing process.

2.6 Accuracy

Accuracy is the common evaluation metrics used to assess the performance of a classification model. Accuracy is computed as the ratio of correct predictions to the total number of predictions made by the model. Mathematically, accuracy can be expressed in Equation 1.

$$Accuracy = \frac{Sum\ of\ Right\ Prediction}{Sum\ of\ Prediction} \qquad (1)$$

Accuracy provides a general overview of how effective the model classifies data correctly. However, accuracy alone is not sufficient to understand the model's performance, especially if the data is imbalanced (when some classes are more dominant than others). In such cases, additional evaluation metrics are needed to provide a deeper understanding of the model's performance.

2.7 Confusion Matrix

The confusion matrix, which lists the proportion of accurate and inaccurate predictions for each class, is a crucial tool for evaluating the performance of a classification model. It is set up as a table that shows the numbers of accurate and inaccurate forecasts for every category. Numerous important evaluation measures, including precision, recall, F1-score, and support, can be obtained from this matrix.

Precision is defined as the proportion of correct positive predictions among all the positive predictions generated by the model in Equation 2. High precision signals a low number of false positives. Recall, sometimes referred to as sensitivity or true positive rate, is shown in Equation 3. It measures the proportion of accurate positive predictions among all actual positive events. A high recall means that the model is able to correctly identify most positive examples. The F1-score, which is the harmonic mean of recall and precision and offers a metric that balances both, can be computed using Equation 4. Recall and precision are well-balanced in the model when the F1-score is high.

$$Precisison = \frac{True\ Positives(TP)}{True\ Positives(TP) + True\ Negatives(TN)} \qquad (2)$$

$$Recall = \frac{True\ Positives}{True\ Positive(TP) + False\ Negatives(FN)} \qquad (3)$$

$$F1 - Score = 2 \times \frac{Precisison \times Recall}{Precision + Recall} \qquad (4)$$

Support is the number of actual occurrences of each class in the dataset. Support provides information about the class distribution in the data and helps understand the context of other evaluation metrics.

3. Results and Discussion

The data splitting resulted in three sets with different ratios: 70%:30%, 80%:20%, and 90%:10%. Each split produced a test set containing 232 data points and a training set containing 141 data points, a training set containing 368 data points and a test set containing 95 data points, and a training set containing 414 data points and a test set containing 49 data points. The data splitting was performed per class to ensure a better distribution of data variations.

The classification model's performance on the validation set with 70% data training and 30% data testing is moderately accurate, achieving an overall accuracy of 50%. This indicates the model can identify some validation samples correctly, but there's room for improvement. The confusion matrix gives a detailed information breakdown of true and predicted classes, revealing the model's performance for each class. The classification report can be seen in Table 1.

Table 1. Classification report of 70% and 30% data split classification

Class	Precision	Recall	F1-score	Support
bima	0.36	0.31	0.33	30
arjuna	0.56	0.65	0.60	30
nakula_sadewa	0.36	0.26	0.31	30
puntadewa	0.62	0.73	0.67	45
accuracy	-	-	0.50	133
macro avg	0.47	0.49	0.48	133
weighted avg	0.48	0.50	0.49	133

The model struggles with 'bima,' having a precision of 0.36, recall of 0.31, and F1-score of 0.33. This suggests moderate identification of 'bima' samples with a higher rate of misclassification. The class‘arjuna' performs best, with a precision of 0.56, recall of 0.65, and F1-score of 0.60. The high recall indicates the model successfully identifies most 'arjuna' samples. The nakula_sadewa class underperforms with a precision of 0.36, recall of 0.26, and F1-score of 0.31. The low recall suggests the model misses many 'nakula_sadewa' samples. 'puntadewa' has the best metrics with a precision of 0.62, recall of 0.73, and F1-score of 0.67, demonstrating reliable predictions for this class. Macro-average precision, recall, and F1-score are 0.47, 0.49, and 0.48, respectively. Weighted averages (0.48 precision, 0.50 recall, 0.49 F1-score) account for class imbalance by considering the number of instances in each class. These values better represent the model's overall performance, morever with imbalanced datasets. The confusion matrix highlights classes where the model struggles. 'bima' and 'nakula_sadewa' have a high number of misclassified samples. 'puntadewa' has the highest number of correctly classified samples (32 out of 44), indicating better model performance for this class.

The classification model's performance on the validation set of 80% data training and 20% data testing demonstrates an accuracy of 52%. This suggests that the model correctly classifies just over half of the validation samples. The confusion matrix and classification report provide detailed insights into the model's performance across different classes. The complete classification report for this dataset can be seen in Table 2.

Table 2. Classication report for the 80% and 20% data split classification

Class	Precision	Recall	F1-score	Support
bima	0.37	0.33	0.35	21
arjuna	0.60	0.43	0.50	21
nakula_sadewa	0.52	0.61	0.56	23
puntadewa	0.56	0.63	0.59	30
accuracy	-	-	0.52	95
macro avg	0.51	0.50	0.50	95
weighted avg	0.52	0.52	0.51	95

The model shows a precision of 0.37, recall of 0.33, and F1-score of 0.35 for the class 'bima'. This indicates a moderate ability to correctly identify samples of this class but also highlights frequent misclassifications. The ‘arjuna’ class has a precision of 0.60, recall of 0.43, and F1-score of 0.50. The precision is relatively higher, suggesting the model makes fewer false positive errors for ‘bima’ class. ‘nakula_sadewa’ class perform with a precision of 0.52, recall of 0.61, and F1-score of 0.56, the model shows a better balance between precision and recall for this class, although the recall is notably higher. ‘Puntadewa’ class achieved the highest recall of 0.63 and an F1-score of 0.59, indicating the model's strong performance in identifying 'puntadewa' samples. The macro average precision, recall, and F1-score are 0.51, 0.50, and 0.50 respectively, reflecting the performance of the model across all classes equally. The weighted average precision, recall, and F1-score are 0.52, 0.52, and 0.51 respectively, which account for the support (number of true instances) of each class. These metrics are slightly higher than the macro averages, indicating a slight imbalance in class distribution. The confusion matrix shows that the model has difficulty distinguishing between 'bima' and 'puntadewa', as well as 'arjuna' and 'nakula_sadewa', leading to significant misclassifications. The class 'puntadewa' has the most correctly classified samples with 19 out of 30 samples correctly identified, demonstrating that the model can distinguish this class more effectively than others.

The 90% data training and 10% data test set got overall 0.39 accuracy. The accuracy score of 0.39 suggests that the classifier correctly predicts 39% of the instances in the test dataset. While this indicates some predictive capability, it also highlights the need for improvement. Achieving higher accuracy is crucial for reliable and effective classification tasks. Therefore, it's essential to explore strategies to enhance the classifier's performance. The complete classification report can be seen in Table 3.

Table 3. Classification report for the 90% and 10% data split classification

Class	Precision	Recall	F1-score	Support
bima	0.12	0.09	0.11	11
arjuna	0.50	0.64	0.56	11
nakula_sadewa	0.33	0.08	0.13	12
puntadewa	0.42	0.67	0.51	15
accuracy	-	-	0.39	49
macro avg	0.34	0.37	0.33	49
weighted avg	0.35	0.39	0.34	49

The 'bima' class, the precision stands at a mere 0.12, with recall and F1-score registering at 0.09 and 0.11, respectively, indicative of limited ability to accurately identify instances of this class. Conversely, the 'arjuna' class exhibits more promising metrics, with precision, recall, and F1-score reaching 0.50, 0.64, and 0.56, respectively, suggesting relatively better performance in correctly classifying 'arjuna' samples. 'Nakula_sadewa' class, however, displays lower precision at 0.33, with recall and F1-score at 0.08 and 0.13, respectively, indicating challenges in both accurately identifying and recalling instances of this class. 'Puntadewa' class demonstrates a precision of 0.42, recall of 0.67, and F1-score of 0.51, showed a relatively balanced performance in both precision and recall. The accumulation accuracy recorded at 0.52, indicating correct predictions for 52% of instances in the test dataset. Macro-average metrics, encompassing precision, recall, and F1-score, yield values of 0.34, 0.37, and 0.33, respectively, highlighting the overall model performance across all classes. Similarly, the weighted average metrics, accounting for class support, are marginally higher at 0.35, 0.39, and 0.34 for precision, recall, and F1-score, respectively. These results underscore the necessity for targeted improvements, particularly in classes with lower precision and recall, to enhance overall classifier efficacy and predictive accuracy.

In summary, the performance of the classification model varies across different validation sets. For the 70% data training and 30% data testing validation set, the model achieves a moderate accuracy of 50%, indicating its ability to correctly identify some validation samples but also highlighting areas for improvement. Similarly, with 80% data training and 20% data testing, the model demonstrates an accuracy of 52%, suggesting slightly improved performance in correctly classifying validation samples. However, the validation set with 90% data training and 10% data testing yields a lower accuracy score of 39%, indicating the need for further enhancement to achieve more reliable and effective classification. Therefore, it is imperative to explore strategies aimed at improving the model's performance, such as fine-tuning model parameters, optimizing feature selection, or addressing class imbalances. Additionally, conducting a thorough analysis of the confusion matrix and classification reports across different validation sets can provide valuable insights for refining the model and achieving higher accuracy in future iterations.

4. Conclusion

We have observed the classification model of wayang kulit using CNN-ELM method. The classification model is built by integrating the CNN image features extraction and ELM classification architecture. The model's performance across different validation sets varies, with accuracy scores ranging from 39% to 52%. While the accuracies achieved on the 70% data training and 30% data testing, and 80% data training and 20% data testing sets indicate moderate performance, correctly classifying around half of the validation samples, the lower accuracy of 39% on the 90% data training and 10% data testing set suggests poorer performance. We have explained how the CNN-ELM method used has good performance in classifying images, but this research has not provided evidence of how this method will have significant performance in classifying Wayang Kulit images. This may be due to the quality of the data used, the complexity of the sample images, and the suboptimal model development. This research can be used as a reference for future researchers, so that the application of CNN-ELM in a certain way can be significantly beneficial in broader fields, including Wayang Kulit. We are open to extensive future improvements regarding the method we used. For future research, broader exploration can be conducted in the application of the CNN-ELM method. The shortcomings produced by this method for specific cases such as Wayang Kulit can open up space for greater improvements that will provide exciting possibilities for advances in the field of machine learning, especially image classification. Thus, this method can be applied in broader fields and contribute more to society in general and future researchers.

REFERENCES

[1] Setiawan, E. "Makna Nilai Filosofis Wayang Kulit sebagai Media Dakwah. Al Hikmah", Jurnal Ilmu dakwah dan Pengembangan Masyarkat, vol. 18, no. 1, pp. 33-49, 2020. https://doi.org/10.35719/alhikmah.v18i1.21

[2] Rahmatita, N., Susanto, H., Sriwati. "Menelisik Sejarah dan Nilai Sosial Budaya dalam Pertunjukan Wayang Kulit Banjar", KAGANGA : Jurnal Pendidikan Sejarah dan Riset Sosial Humaniora, vol. 7, no. 1, pp. 103-113, 2024. https://doi.org/10.31539/kaganga.v7i1.8179

[3] M. Z. Alfaqi, "Eksistensi dan Problematika Pelestarian Wayang Kulit Pada Generasi Muda Kec. Ringinrejo Kab. Kediri," J. Praksis dan Dedik., vol. 5, no. 2, pp. 119–128, 2022. http://dx.doi.org/10.17977/um032v5i2p119-128

[4] BPS. (2021). Proyeksi Penduduk Kabupaten Kediri. https://kedirikab.bps.go.id/indicator/-12/62/1/proyeksi-penduduk-kabupaten-kediri-kel-umur-.html

[5] Riyani, D., Ardini, L. "Kontruksi Prinsip Dasar Etika Profesi Akuntan: Internalisasi Karakter Pewayangan Pandava Lima. Wacana Ekonomi". Jurnal Ekonomi, Bisnis dan Akuntansi, vol. 20 no. 2, pp. 84-93, 2021. https://doi.org/10.22225/we.20.2.2021.84-93

[6] A. Herdiansah, R. I. Borman, D. Nurnaningsih, A. A. J. Sinlae, and R. R. Al Hakim, "Klasifikasi Citra Daun Herbal Dengan Menggunakan Backpropagation Neural Networks Berdasarkan Ekstraksi Ciri Bentuk," JURIKOM (Jurnal Ris. Komputer), vol. 9, no. 2, pp. 388–395, 2022, http://dx.doi.org/10.30865/jurikom.v9i2.4066

[7] R. Nuraini, A. Wibowo, B. Warsito, W. A. Syafei, and I. Jaya, "Combination of K-NN and PCA Algorithms on Image Classification of Fish Species", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 7, no. 5, pp. 1026 - 1032, 2023. https://doi.org/10.29207/resti.v7i5.5178

[8] W. Supriyanti, D. Aryo Dimas, "Classification of Pandavas Figure in Shadow Puppet Images using Convolutional Neural Networks", Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika, vol. 7, no. 1, pp. 18-24, 2020. https://doi.org/10.23917/khif.v7i1.12484

[9] F. Fatmayati, M. Nugraheni, R. Nuraini, and F. Rossi, "Classification of Character Types of Wayang Kulit Using Extreme Learning Machine Algorithm", Building of Informatics, Technology and Science (BITS), v. 5, no. 1, pp. 317-326, 2023. https://doi.org/10.47065/bits.v5i1.3568

[10] M. Muhathir, M. H. Santoso, and D. A. Larasati, “Wayang Image Classification Using SVM Method and GLCM Feature Extraction,” Journal of Informatics and Telecomunication Engeneering, vol. 4, no. 2, 2021, https://doi.org/10.31289/jite.v4i2.4524.
[11] J. Wang, S. Lu, S. H. Wang, and Y. D. Zhang, “A review on extreme learning machine,” *Multimedia Tools and Applications*, vol. 81, no. 29, 2022, https://doi.org/10.1007/s11042-021-11007-7.
[12] S. Al-hammouri, M. Fora, and M. Ibbini, “Extreme Learning Machine for Melanoma Classification,” in Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT) 2021, 2021, pp. 114–119. https://doi.org/10.1109/JEEIT53412.2021.9634135
[13] W. Xie, Z. Li, Y. Xu, P. Gardoni, and W. Li, “Evaluation of Different Bearing Fault Classifiers in Utilizing CNN Feature Extraction Ability,” *Sensors*, vol. 22, no. 9, 2022, doi: 10.3390/s22093314.
[14] J. P. T. Yusiong, “A CNN-ELM Classification Model for Automated Tomato Maturity Grading,” *Journal of ICT Research and Applications*, vol. 16, no. 1, 2022, https://doi.org/10.5614/itbj.ict.res.appl.2022.16.1.2.
[15] R. Kaur, R. K. Roul, and S. Batra, “A hybrid deep learning CNN-ELM approach for parking space detection in Smart Cities,” *Neural Computing and Applications*, vol. 35, no. 18, 2023, https://doi.org/10.1007/s00521-023-08426-y.
[16] J. P. T. Yusiong, “An Ensemble Of CNN-ELM Models For Trash Classification,” *ICIC Express Letters*, vol. 16, no. 9, 2022, https://doi.org/10.24507/icicel.16.09.943.
[17] G. Parasa, M. Arulselvi, and S. Razia, “An Enhanced CNN-based ELM Classification for Disease Prediction in the Rice Crop,” *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 7, 2023, https://doi.org/10.17762/ijritcc.v11i7s.7556.
[18] A. Anton, N. F. Nissa, A. Janiati, N. Cahya, and P. Astuti, “Application of Deep Learning Using Convolutional Neural Network (CNN) Method For Women’s Skin Classification,” *Scientific Journal of Informatics*, vol. 8, no. 1, pp. 144–153, 2021, https://doi.org/10.15294/sji.v8i1.26888.
[19] Y. Dinda Ulima Rizky. “Klasifikasi Tingkat Keparahan Non-ProliferativeI Diabetic Retinopathy Bedarsarkan Hard Exudate Menggunakan Extreme Learning Machine,” Jurnal Sains dan Seni ITS, vol. 6, no. 2, pp 88-93, 2017. http://dx.doi.org/10.12962/j23373520.v6i2.27454
[20] Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). "How transferable are features in deep neural networks?" In Advances in neural information processing systems, pp. 3320-3328. https://doi.org/10.48550/arXiv.1411.1792
[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
[22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Commun. ACM, vol. 60, no. 6, pp. 84-90, 2017, https://doi.org/10.1145/3065386.
[23] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme Learning Machine for Regression and Multiclass Classification," IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 42, no. 2, pp. 513-529, 2012, https://doi.org/1109/TSMCB.2011.2168604 .
[24] Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 8, pp. 1798-1828, 2013, https://doi.org/10.1109/TPAMI.2013.50 .
[25] Z. Tang, Y. Li, and J. Hao, "An Improved Extreme Learning Machine Based on Particle Swarm Optimization for Classification," J. Ambient Intell. Humaniz. Comput., vol. 10, no. 6, pp. 2305-2312, 2019, https://doi.org/10.1007/s12652-018-0807-1.
[26] L. Zhang, L. Lin, X. Wang, W. Wang, and Y. Liang, "ResNet and DenseNet Models for Image Classification," IEEE Access, vol. 8, pp. 2304-2311, 2020, doi: https://doi.org/10.1109/ACCESS.2020.2964478