Recursive Journal of Informatics
Vol. XX, No. XX, Month 20XX

p-ISSN 2963-5551 https://journal.unnes.ac.id/sju/index.php/rji e-ISSN 2986-6588

Image classification of Human Face Shapes Using Convolutional Neural Network Xception Architecture with Transfer Learning

Resta Adityatama[1]*, Anggyi Trisnawan Putra[2]
[1,2]Computer Science Department, Faculty of Mathematics and Natural Sciences,
Universitas Negeri Semarang, Indonesia

Abstract. The development of information technology in facial recognition is influenced by a faster and more accurate authentication system. This allows the computer system to identify a person's face.
Purpose: Similar to fingerprints and the retina of the human eye, each person's face has a different shape and contour. Since it is known that the human face provides a lot of information, as well as topics that attract attention make it studied intensively.
Methods/Study design/approach: Several studies examining information from human faces are facial recognition. One of the approaches used to recognize facial imagery is through the use of a Convolutional Neural Network (CNN). CNN is a method in the field of Deep Learning that can be used to recognize and classify objects in digital images. In this study, the method used to implement facial image classification is the Xception architecture CNN algorithm with a transfer learning approach.
Result/Findings: The dataset used in this study was obtained from Kaggle, namely the Face Shape Dataset which contains 5000 data. After testing, an accuracy rate of 96.2% was obtained in the training process and 81.125% in the validation process. This study also uses new data to test the model that has been made, and the results show an accuracy rate of 85.1% in classifying facial imagery.
Novelty/Originality/Value: Therefore, it can be said that the model created in this study has the ability to classify images of facial shapes Human Face Shapes Using Convolutional Neural Network Xception Architecture with Transfer Learning.

Keywords: Image Classfication, Face Shape, Convolutional Neural Network, Xception, Transfer Learning
Received June 02, 2023 / **Revised** June 17, 2023 / **Accepted** September 29, 2023

INTRODUCTION

The development of information technology in facial recognition is influenced by a faster and more accurate authentication system. This allows the computer system to identify a person's face. Similar to fingerprints and the retina of the human eye, each person's face has a different shape and contour. So that the face can be used as a tool or model to identify a person [1]. There are many recognition objects used for recognition studies, one of which is facial shape [2]. Humans have various facial shapes that can be categorized into 5, namely oval, round, oblong, square, and heart [3]. Face recognition is an important first step and is required by many computer applications. Face recognition is the process of identifying a person in an image or video. This identification is carried out through a database that stores many individuals[4], [5].

Face recognition is also a pattern recognition and image processing application that can be used to solve various problems. Previously, various algorithms have been introduced to solve problems regarding facial recognition. Unfortunately, the implementation of the algorithm is not accurate enough. To overcome this problem, there is a new approach in the form of artificial neural networks based on deep learning. Face recognition can be done using deep learning. One of the deep learning methods currently being developed is the Convolutional Neural Network (CNN). CNN has proven to be able to classify images very well [6][7]. The CNN method is a method that takes input data in the form of images. This method has a special layer called the convolution layer. In the convolution layer, the input image will produce patterns from different parts, making it easier to classify later [8][9].

[1]*Corresponding author.
Email addresses: restaadit@students.unnes.ac.id (Adityatama)
DOI: 10.15294/rji.v1i2.70774

CNN is a deep learning algorithm that is currently a trend in classifying images. The CNN algorithm has many types of architectures such as Google-Net, Xception-Net, VGG-Net, DenseNet, Le-Net, and others. One approach that can be applied to the CNN method in classification problems is the transfer learning approach. In this study, the CNN method uses the Xception architecture with a transfer learning approach to classify objects.

METHODS

In this study, classification was carried out using facial image objects and using the Convolutional Neural Network (CNN) method with the Xception architecture transfer learning approach [10][11]. In recent years, the field of computer vision has made significant strides in understanding and interpreting images. One fascinating application of computer vision is the classification of human face shapes. Identifying face shapes is not only important for cosmetic and fashion industries but also plays a crucial role in fields like medical diagnostics and biometrics. In this article, we delve into the world of image classification using deep learning techniques, particularly Convolutional Neural Networks (CNNs). We will explore how to leverage the power of transfer learning and the Xception architecture to create a robust face shape classification model [12]. The model framework that will be carried out can be seen in Figure 1.

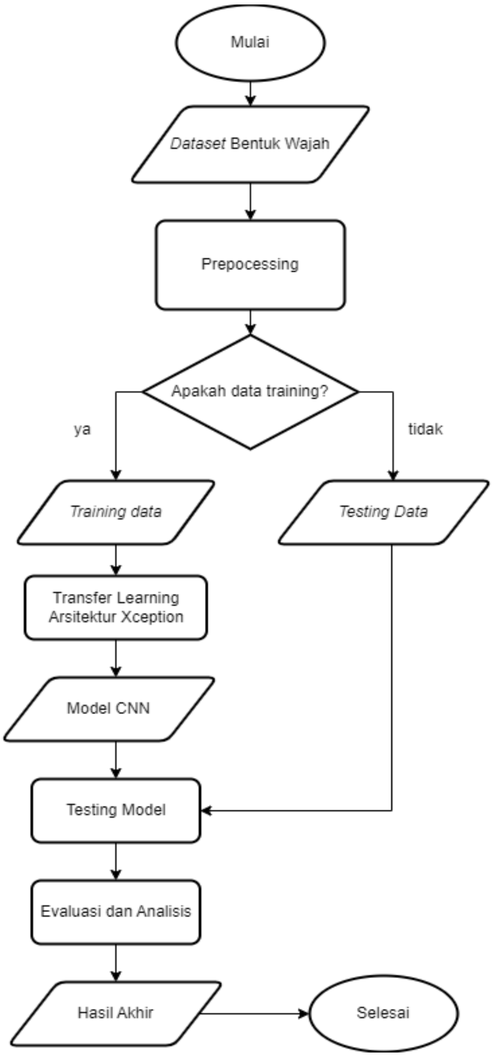

Figure 1. Flowchart Research Design

Convolutional Neural Networks have proven to be exceptionally effective in image classification tasks. They are designed to automatically learn patterns and features from images, making them an ideal choice for our face shape classification problem. Transfer learning is a technique where we take a pre-trained neural network and fine-tune it for a specific task [13][14][15]. In our case, we'll use a pre-trained Xception model, which is known for its excellent performance on image-related tasks. we've explored the exciting world of image classification for human face shapes [16][17]. By employing Convolutional Neural Networks and transfer learning with the Xception architecture, we can create a robust model capable of accurately categorizing faces into oval, round, square, heart, and diamond shapes. The potential applications of such a model are vast, ranging from personalized beauty recommendations to medical diagnostics. As computer vision continues to advance, the accuracy and efficiency of these models will only improve, opening up new possibilities in the world of image analysis and interpretation [18].

RESULT AND DISCUSSION

In this study the method used is the Convolutional Neural Netwrok (CNN) Xception architecture with Transfer learning approach. The research was carried out in several stages, The steps include data collection from Kaggle, preprocessing, data split into training and validation sets, training with the pre-trained Xception model, and evaluating the model using a confusion matrix.

Dataset

In this study, a public dataset called Face Shape Dataset was used which was obtained through Kaggle. This dataset is collected by Niten lama and consists of 5000 images of female celebrities from various countries. Each image is categorized based on face shape into five main classes, namely Heart, Oblong, Oval, Round, and Square. This dataset can be accessed via the link https://www.kaggle.com/datasets/niten19/face-shape-dataset, example dataset as shown in Figure 2.

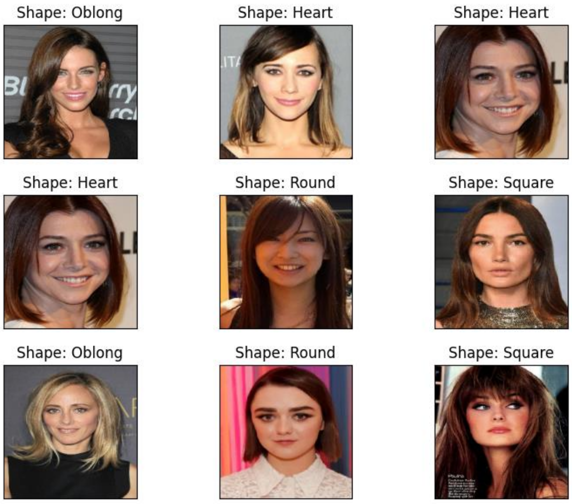

Figure 2. Example Dataset

Preprocessing Data

The augmentation process for facial image data for the training process is carried out to prevent overfitting. Figure 3 shows the results of the data augmentation process. You can see some examples of augmentation results, such as image flipping, shifting, and scaling.

Figure 3. Preprocessing Result

Modeling

Making a model or modeling consists of several processes. To create a classification model for facial imagery, facial imagery is needed as a source of learning data. Image data is loaded into two variables, namely training image data and validation image data. The image size for each variable is changed to 224x224 pixels. Model making in this study was carried out using a transfer learning approach. The part of the model that will be made is the layer for feature extraction using the Xception layer and for the fully connected layer consisting of 3 layers. The final model that has been designed can be seen in Figure 4.

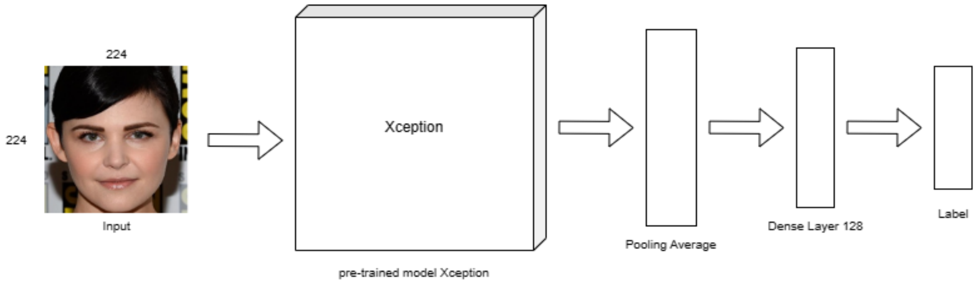

Figure 4 Final Model

Training Process

Table 1 shows the results of the training using 20 iterations (epochs) on the training and validation dataset. In the training dataset, the first epoch produces an accuracy of 0.2702 with a loss of 1.5732. Furthermore, in the second epoch the accuracy increased to 0.4109 with a loss of 1.3908. This process continues until the 20th epoch, where the accuracy reaches 0.9627 with a loss of 0.1107. For the validation dataset, in the first epoch the accuracy was 0.3073 with a loss of 1.6039. In the second epoch, the accuracy increased to 0.3620 with a loss of 2.6849. In the 20th epoch, the accuracy reached 0.8125 with a loss of 0.7738. Thus, the training results show an increase in accuracy and a decrease in loss as the number of epochs increases.

Table 1. Fit Model Result

Epoch	*Data Train*		*Data Validation*	
	Accuracy	Loss	Val Accuracy	Val Loss
1	0,2702	1,5732	0,3073	1,6039
2	0,4109	1,3908	0,3620	2,6849
3	0,4868	1,2285	0,2083	2,8160
4	0,5488	1,1027	0,5729	1,0371
5	0,6079	0,9940	0,5312	1,3159
6	0,6581	0,8849	0,4661	1,9613
7	0,7102	0,7627	0,5938	1,2905
8	0,7436	0,6626	0,6250	1,1270
9	0,7747	0,6019	0,7239	0,7768
10	0,7884	0,5538	0,7135	0,9879
11	0,8237	0,4880	0,6172	1,3503
12	0,8559	0,4049	0,6615	1,0319
13	0,8770	0,3430	0,7214	1,0193
14	0,8994	0,2712	0,7188	0,9928
15	0,9131	0,2384	0,7708	0,8251
16	0,9162	0,2241	0,7604	0,8769
17	0,9294	0,1877	0,7370	1,1662
18	0,9428	0,1671	0,8099	0,7208
19	0,9490	0,1423	0,8099	0,8224
20	0,9627	0,1107	0,8125	0,7738

Figure 5 shows a graph of the movement of accuracy values for the train data and validation data generated at each iteration (epoch). Based on the figure, the red line shows the movement for the data train, while the blue line shows the movement of the accuracy value for data validation, it can be seen that for the data train, the accuracy value obtained at the beginning of the epoch produces 0.2702, continues to rise until the epoch with a value highest accuracy 0.9627. As for data validation, the accuracy value for the last epoch was 0.8125

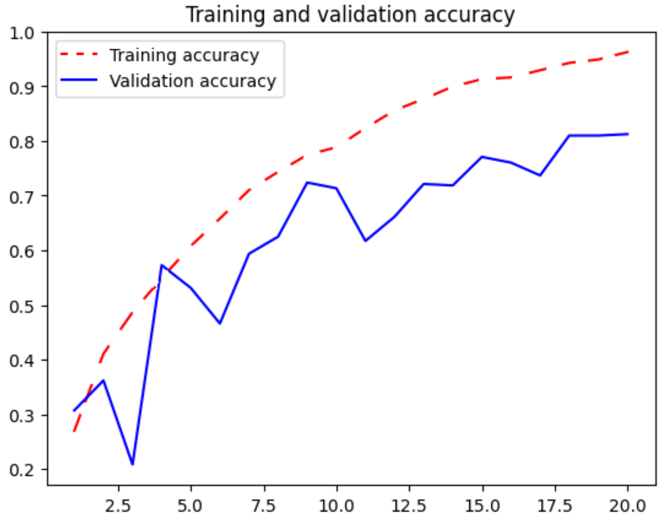

Figure 5. Accuracy per Epoch

The results of the training process on the Face Shape dataset show fluctuating loss values at the validation stage, with the highest loss value of 2.816 and the lowest loss value of 0.7738 at the validation stage. Then at the training stage the highest loss value is 1.5732 and the lowest loss value is 0.7738. The lower the loss value, the better the model performance obtained. The loss graph can be seen in Figure 6.

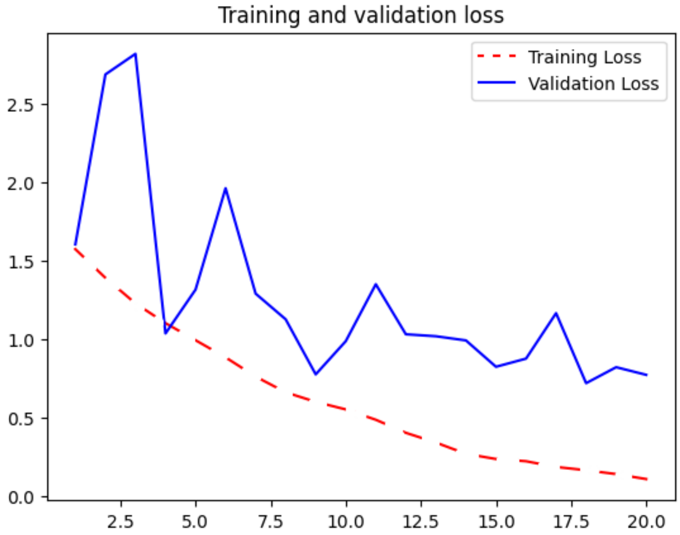

Figure 6. Loss per Epoch

Model Evaluation

Model evaluation in this study uses a confusion matrix. with the confusion matrix is done to determine the performance of the model in predicting the class of data testing. The testing data used amounted to 1000 data which was divided into 200 data for each face shape class. After testing the model on data testing, a confusion matrix will be obtained. From the confusion matrix, several performance metrics will be obtained which can be used to find out how the model performs in predicting the data contained in the data testing. The confusion matrix in this study is shown in Figure 7.

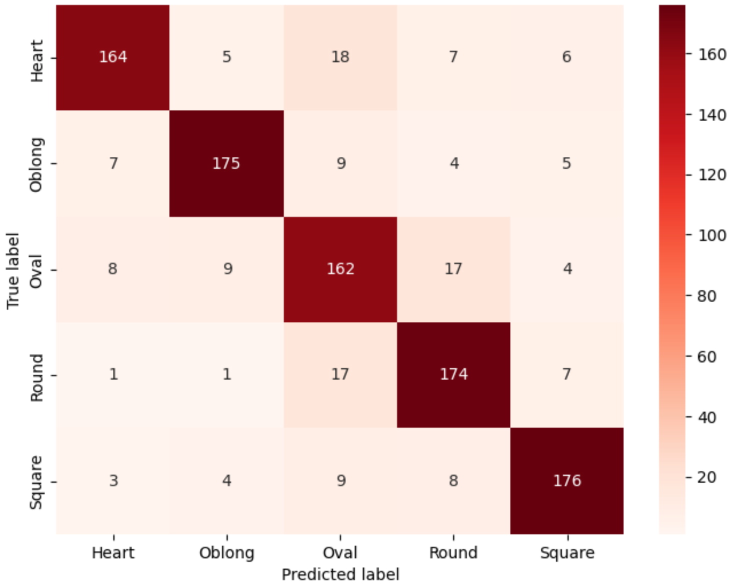

Figure 7. Confussion Matrix

From the confusion matrix shown in Figure 7, several performance metrics will be obtained that can be used to determine the performance of the model created, then the average calculation of the Precision, Recall, and f1-score values will be carried out. The results of the calculations are shown in Figure 8.

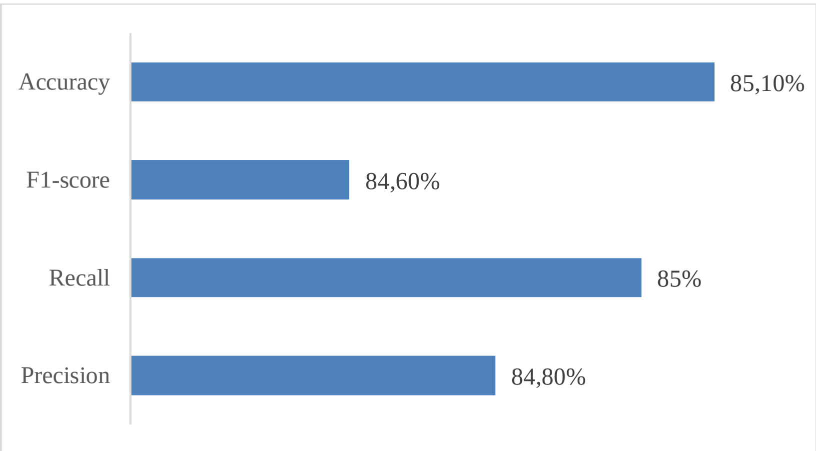

Figure 8. Final Result Performance Metrics

From the results of evaluating the performance of the image classification of human face shapes using convolutional neural network xception architecture with transfer learning approach, the results obtained were 85,1% accuracy, 84,8% precision, 85% recall, and 84,6% f1-score.

CONCLUSION

Based on the results of research that has been done regarding image classification using Convolutional Neural Network (CNN) Xception architecture with Transfer learning approach. The use of the CNN algorithm with transfer learning approach using Xception architecture to classify human facial images has resulted in significant outcomes. In this research, 80% of the total dataset was used as training data to train the model, while the remaining 20% served as testing data to evaluate the performance of the trained model. The findings of this study demonstrate an improvement in the model's performance in classifying human facial shapes. By employing the CNN transfer learning method with Xception architecture, the highest accuracy achieved was 85.1%.

REFERENCES

[1] S. Salamun and F. Wazir, “Rancang Bangun Sistem Pengenalan Wajah Dengan Metode Principal Component Analysis”, rabit, vol. 1, no. 2, pp. 61-77, Jul. 2016.

[2] Muthmainnah Muthmainnah and Rahayu Rahayu, “Pengenalan Bentuk Wajah Manusia Pada Citra Menggunakan Metode Fisherface,” TECHSI - Jurnal Teknik Informatika, vol. 8, no. 1, pp. 215–225, Mar. 2019, doi: https://doi.org/10.29103/techsi.v8i1.128.

[3] Wisuwat Sunhem and Kitsuchart Pasupa, “An approach to face shape classification for hairstyle recommendation,” Feb. 2016, doi: https://doi.org/10.1109/icaci.2016.7449857.

[4] S. Khan, E. Ahmed, M. H. Javed, S. A. A Shah, and S. U. Ali, “Transfer Learning of a Neural Network Using Deep Learning to Perform Face Recognition,” IEEE Xplore, Jul. 01, 2019. https://ieeexplore.ieee.org/document/8940754

[5] A. F. Bobick and J. W. Davis, “The recognition of human movement using temporal templates,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 3, pp. 257–267, Mar. 2001, doi: https://doi.org/10.1109/34.910878.

[6] F. F. Maulana and N. Rochmawati, “Klasifikasi Citra Buah Menggunakan Convolutional Neural Network,” Journal of Informatics and Computer Science (JINACS), vol. 1, no. 02, pp. 104–108, Jan. 2020, doi: https://doi.org/10.26740/jinacs.v1n02.p104-108.

[7] Fathan Arsyadani and Aji Purwinarko, “Implementation of Synthetic Minority Oversampling Technique and Two-phase Mutation Grey Wolf Optimization on Early Diagnosis of Diabetes using K-Nearest Neighbors,” vol. 1, no. 1, pp. 9–17, Mar. 2023, doi: https://doi.org/10.15294/rji.v1i1.64406.

[8] Yahya Nur Ifriza, Trisni Wulandari Veronika, Trisni Suryarini, and Antonius Supriyadi, “The Modeling Of Laboratory Information Systems In Higher Education Based On Enterprise Architecture Planning (EAP) For Optimizing Monitoring And Equipment Maintenance,” Matrix, vol. 13, no. 1, pp. 1–11, Mar. 2023, doi: https://doi.org/10.31940/matrix.v13i1.1-11.

[9] A. B. Susilo and E. Sugiharti, “Accuracy Enhancement in Early Detection of Breast Cancer on Mammogram Images with Convolutional Neural Network (CNN) Methods using Data Augmentation and Transfer Learning,” Journal of Advances in Information Systems and Technology, vol. 3, no. 1, pp. 9–16, Apr. 2021, doi: https://doi.org/10.15294/jaist.v3i1.49012.

[10] A. Adimas and S. Y. Irianto, “Image Sketch Based Criminal Face Recognition Using Content Based Image Retrieval,” Scientific Journal of Informatics, vol. 8, no. 2, pp. 176–182, Nov. 2021, doi: https://doi.org/10.15294/sji.v8i2.27865.

[11] Pulung Nurtantio Andono and Christy Atika Sari, “Remove Blur Image Using Bi-Directional Akamatsu Transform and Discrete Wavelet Transform,” Scientific Journal of Informatics, vol. 9, no. 2, pp. 179–188, Nov. 2022, doi: https://doi.org/10.15294/sji.v9i2.34173.

[12] S. Karimpouli and P. Tahmasebi, “Image-based velocity estimation of rock using Convolutional Neural Networks,” Neural Networks, vol. 111, pp. 89–97, Mar. 2019, doi: https://doi.org/10.1016/j.neunet.2018.12.006.

[13] H. Stone, “Calibrating Rough Volatility Models: A Convolutional Neural Network Approach,” SSRN Electronic Journal, 2019, doi: https://doi.org/10.2139/ssrn.3327135.

[14] A. I. Khan, J. L. Shah, and M. M. Bhat, “CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images,” Computer Methods and Programs in Biomedicine, vol. 196, p. 105581, Nov. 2020, doi: https://doi.org/10.1016/j.cmpb.2020.105581.

[15] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation,” 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 580–587, Jun. 2014, doi: https://doi.org/10.1109/cvpr.2014.81.

[16] M. Jogin, Mohana, M. S. Madhulika, G. D. Divya, R. K. Meghana, and S. Apoorva, “Feature Extraction using Convolution Neural Networks (CNN) and Deep Learning,” IEEE Xplore, 2018. https://ieeexplore.ieee.org/document/9012507.

[17] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature Pyramid Networks for Object Detection,” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jul. 2017, doi: https://doi.org/10.1109/cvpr.2017.106.

[18] A. Wang, W. Wang, J. Liu, and N. Gu, “AIPNet: Image-to-Image Single Image Dehazing With Atmospheric Illumination Prior,” IEEE Transactions on Image Processing, vol. 28, no. 1, pp. 381–393, Jan. 2019, doi: https://doi.org/10.1109/tip.2018.2868567.