Jurnal Mantik, 6 (2) (2022) 263-270

Published by:Institute of Computer Science (IOCS)

Jurnal Mantik

Journal homepage: www.iocscience.org/ejournal/index.php/mantik/index

Ear Identification Using Convolution Neural Network

Nadiyah Hidayati[1] , Mawadatul Maulidah[2] , Elin Panca Saputra[3]

[123] Fakultas Teknik dan Informasi, Universitas Bina Sarana Informatika, Jakarta Indonesia

E-mail: [1]nadiyah.nyy@bsi.ac.id, [2]mawadatul.mwm@bsi.ac.id, [3]elin.epa@bsi.ac.id

ARTICLE INFO

Article history:
Received: Mart 28, 2022
Revised: Apr 17 , 2022
Accepted: Apr 28, 2022

Keywords:
Biometrics,
Ear Identification,
CNN.

ABSTRACT

Today's identification system has become a necessity for system security. One method of identification system that has a high level of security and accuracy is biometrics. Biometrics uses parts of the human body that are considered unique and can differentiate between one individual and another. One of the new biometrics that has become a concern in the world of research on biometrics is the ear. Ears have several advantages that other biometrics do not have, one of which is that they are not affected by changes in age. The purpose of this study was to determine the accuracy of the Convolutional Neural Network (CNN) algorithm in identifying ear images. CNN is currently one of the most superior algorithms in the field of object classification and identification. In this study, the ears that will be identified are images taken from the Kaggle dataset of 780 ears from 13 individuals with 60 images for each individual. This study resulted in a training accuracy of 96,3% and a testing accuracy of 79,7%.

1. Introduction

Biometrics is an individual recognition system based on special physical characteristics and individual characteristics (Ratnadewi and Tua 2015). A person's physical or behavioral characteristics that can be used as biometrics must have the following characteristics: (i) universal (everyone has these characteristics), (ii) unique (no two individuals have the same characteristics), (iii) permanent (characteristics are not affected by changes in time), (iv) collectable (characteristics are easily obtained and can be measured quantitatively)(Siregar 2017).

The ear is one of the most interesting biometrics to study because the shape of the human ear will remain the same from birth to old age, the only thing that changes is its size so that the ear feature that can be used as a differentiator is its geometric shape. (Siregar 2017). The difficulty of detecting and locating the ear precisely in an image is the first step in an ear-based biometric recognition system (Raveane et al. 2019) Based on the unique structure of the ear shape, ear images can provide a rich source of biometric information for building a successful recognition system (Alshazly et al. 2020). addition, the ear image has many advantages over other images such as: it is easy to obtain from a distance without the cooperation of individuals, has a uniform color distribution for the ear surface and does not differ from facial expressions. (Hassaballah, Alshazly, and Ali 2019).

Convolutional Neural Network (CNN) become one of the most superior algorithms in the field of classification and identification today (Mubarok 2019). In several studies deep learning or CNN has shown extraordinary performance. Because with CNN, previous research was able to combine feature extraction and ear recognition tasks into one network with the aim of solving problems such as variations in lighting, contrast, rotation, scale, and pose.(Rao, Kumari, and Rao 2019) This is largely due to its ability to study large datasets and techniques for training deeper networks. In addition, with the support of more powerful computing devices, deep learning has now become widely used (Bengio, Aaron, and Goodfellow 2016). CNN is a development of Multilayer Perceptron (MLP) to process two-dimensional data. Multilayer

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Perceptron itself is the development of Artificial Neural Network (ANN) which is intended to cover the shortcomings of ANN with a single layer perceptron in solving complex logic operations (Yuliani, Aini, and Khasanah 2019). Besides that, ANN is a term for a conventional regression analysis in which there is an equation (Saputra, Indriyanti, and Supriatiningsih 2020).

Several related studies regarding ear identification have been carried out including, Sabet et al (Sabet et al. 2015) perform human ear recognition using geometric feature extraction (shape, mean, centroid and Euclidean distance between pixels). First, they made a pre-processing stage by making all images the same size. Then they used the snake model to detect the ear, and they applied a median filter to remove noise, they also converted the image to binary format. After that they used the canny edge and did some improvements to the image, the largest boundary was calculated and a distance matrix was created then they extracted the image features. Finally, the extracted features are classified using the nearest neighbor with absolute error distance. The experimental results show that the proposed approach gives better results and obtains almost 98% accuracy.

Resmi & Raju (Resmi and Raju 2019) detects ears automatically using Banana Wavelets and circular hough transformation. Banana wavelets originating from the stretched and curved edges of Gabor wavelets are used to identify the curved ear structure. The proposed algorithm is compared with three existing algorithms and evaluated on a standard database. In addition to manual detection accuracy, they also calculated the efficiency of the proposed method using an automated classification technique. LBP and Gabor features extracted from segmented ear images are used by different classifiers to determine whether the segmented image portion is Ear class or not ear.Siregar (Siregar 2017) perform ear identification based on geometric features and KNN. The feature used in the research is the comparison value between the long axis length and feature points that are not affected by rotation, scaling, and displacement, while the classifier used as a measure of the accuracy of the ear identification system is KNN. The accuracy value generated from the ear identification system using geometric features and KNN is 97%, this accuracy value is higher when compared to using the MLP Backpropagation ANN classifier which produces an accuracy of 79.67%.

Hassaballah etc (Hassaballah, Alshazly, and Ali 2019) perform ear recognition using the Local Binary Pattern feature extraction. They conducted a comprehensive comparative study on the identification and verification scenarios separately. In addition, a new variant of the traditional LBP operator named averaged local binary patterns (ALBP) was also proposed and its ability to represent ear image texture is compared with other LBP variants. Ear identification and verification experiments were performed extensively on five ear datasets; namely IIT Delhi (I), IIT Delhi (II), AMI, WPUT and AWE. The results obtained for both identification and verification indicate that the current LBP texture descriptor is a successful feature extraction candidate¬ for ear recognition systems in the case of limited imaging conditions and can achieve recognition rates reaching up to 99%. Ratnadewi & Syafril (Ratnadewi and Tua 2015) identify a person by ear biometrics using the Hough transform. The recognition process is carried out by the Euclidean distance method. The database consists of 50 ear images from 10 individuals with each individual represented by 5 images. Tests were carried out using 20 test images of individuals in the database and 50 test images of individuals not in the database. The results of observations in program testing, obtained the percentage of FRR of 15% and the percentage of FAR of 6%.

Based on the research that has been done. In this paper we propose the CNN method for ear identification. Because Studies in Ear biometrics report that using Convolutional Neural Network (CNN) is a better and suitable alternative for machine learning (Raveane et al. 2019). In addition, deep Convolutional Neural Networks (CNNs) have now been shown to achieve superior performance on a number of computer vision tasks such as image recognition, classification, and object detection. (Jamil et al. 2018). In the next section, the method used in this study will be explained, section 3. will explain the results obtained from the test, and section 4. will discuss the conclusions of the research that has been carried out.

2. Method

In this study, the steps taken were entering the ear image then preprocessing and building a CNN architecture for ear identification. The research flow that will be used in this study can be seen in Figure 1

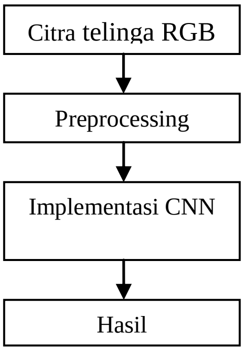
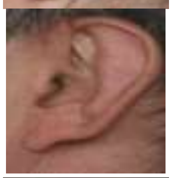
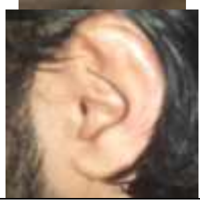


Figure 1. Flow of research method

2.1 Dataset

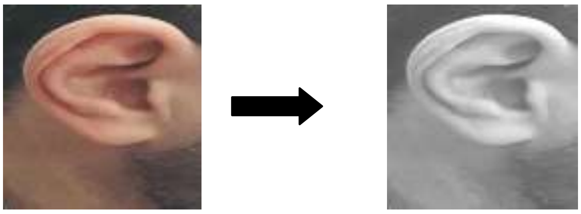
The dataset used in this study is the ear recognition dataset obtained from kaggle (Hatif 2019). The number of ear image datasets used are 780 images with image dimensions of 227 x 227 pixels, consisting of 13 individuals named 000, 001, 002, 003, 004, 005, 006, 007, 008, 009, 010, 011, and 012, where each individual consists of 60 images. The following is an example of an image from each individual can be seen in Table 1.

Tabel 1. Citra Dataset Tiap Individu

Nama Individu	Contoh Citra	Nama Individu	Contoh Citra
001		007	
002		008	
003		009	
004		010	
005		011 012	

2.2 Prepocessing

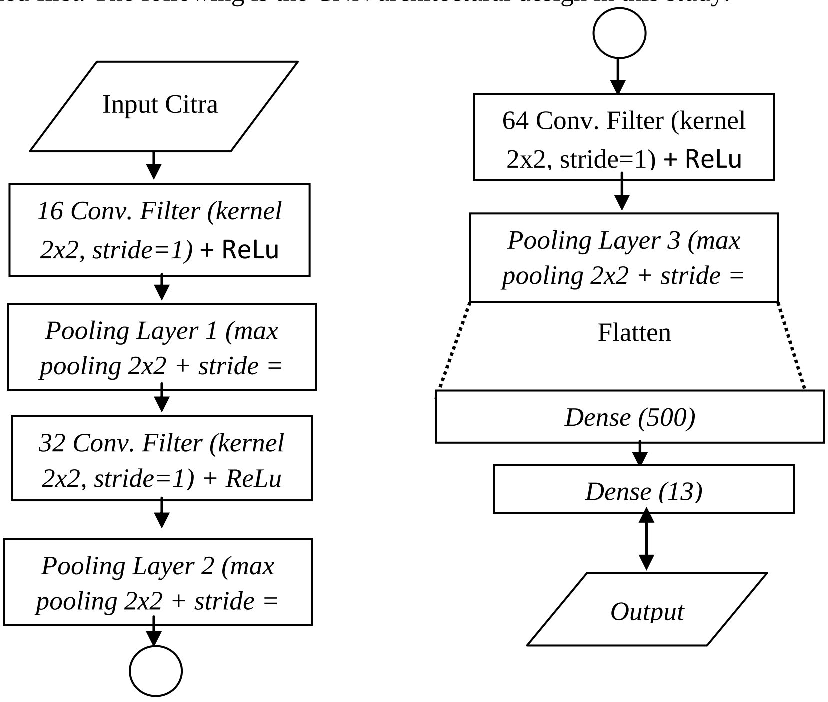
In the preprocessing stage, the ear image resizes to 128 x 128 and converts the RGB image into a grayscale image. Grayscale transformation is done using equation (1).$I(x, y) = \alpha.R + \beta.G + \gamma B$ (1)where I(x,y) is the gray level in a coordinate obtained by adjusting the color composition of R (red), G (green), B (blue) indicated by the parameter values and . The simplest algorithm to perform a grayscale transformation is to find the average value of the components R, G, and B so that the total value is 1. The following is an example of improving a grayscale image which can be seen in Figure 1.



Gambar 2. Transformasi Citra Grayscale

2.3 Implementasi CNN

Before the CNN algorithm is implemented to identify ears, the architecture of this CNN algorithm needs to be designed first. The following is the CNN architectural design in this study.



Figur 3. Arsitektur CNN

To make it easier to understand the CNN architecture in the block diagram above, an explanation of the CNN architecture is as follows:

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

1. Before the image is input, the image will be resized and converted to a grayscale image as described in the preprocessing section above. The results of the preprocessing are then called to be used as input for this CNN model. The input image to be used is 128x128x1. The number 1 in the image size indicates the channel image which is a grayscale image.
2. After that, the inputted image is entered in the first convolution process. In the first convolution, the image will be multiplied by a 2x2 kernel with 16 filters. The multiplication process is done by shifting the kernel by 1 stride.
3. From this first convolution process, the resulting feature map is 64x64x16. After that, the generated feature map will be activated using ReLu.
4. Pooling process, pooling is a subsampling process on the convoluted feature map. Basically this pooling is to reduce the size of the feature map by using a kernel of a certain size which will be shifted over the entire feature map area. In this study using max-pooling with a kernel size of 2x2 and a stride of 2 pixels.
5. After pooling, the pooled feature map will be used as input for the second convolution process. In the second covolution process, the kernel size used is still the same, namely 2x2, but the filter used is increased to 32 filters. The addition of the number of filters is done because in the pooling process more and more information is being discarded, therefore the addition of filters is done so that the variation of information obtained from the available information is more and more. In this second convolution process, a 32x32x32 feature map will be generated. Just like before, this second convolution process also uses the ReLu activation function.
6. The next process goes to the second pooling process, this process is almost the same as the first pooling process, by using a 2x2 kernel with a stride of 2 pixels. From the second pooling process, a 32x32x32 feature map is generated.
7. In the covolution process, the three kernel sizes used are still the same, namely 2x2, but the filter used is increased to 64 filters. In the third convolution process, a 16x16x64 feature map will be generated. Just like before, this second convolution process also uses the ReLu activation function.
8. The next process goes to the third pooling process, this process is almost the same as the second pooling process, by using a 2x2 kernel with a stride of 2 pixels.
9. Before entering the fully connected layer process, the feature map in the form of a 3D array will be converted into a vector 1-D list first, this process is often known as flattening. From the feature map measuring 16x16x64, a vector value of 16384 pixels will be obtained. This result will be used as input in the fully connected layer process.
10. After the flatten stage, it will be forwarded to the fully connected layer (FC Layer) network. This layer is the layer that is usually used in the application of MLP and aims to transform the data dimensions so that the data can be classified linearly. Each neuron in the convolution layer needs to be transformed into one-dimensional data before it can be entered into a fully connected layer. Because it causes the data to lose its spatial information and is not reversible, the fully connected layer can only be implemented at the end of the network (Ilahiyah and Nilogiri 2018).

3. Results And Discussion

The ear identification process was carried out with 780 ear image data from 13 individuals, namely 000, 001, 002, 003, 004, 005, 006, 007, 008, 009, 010, 011, and 012. training uses 706 images, 84 images for testing, and 50 images for validation. After testing using the Convolutional Neural Network algorithm, the following results were obtained:

3.1. Result Proses Training

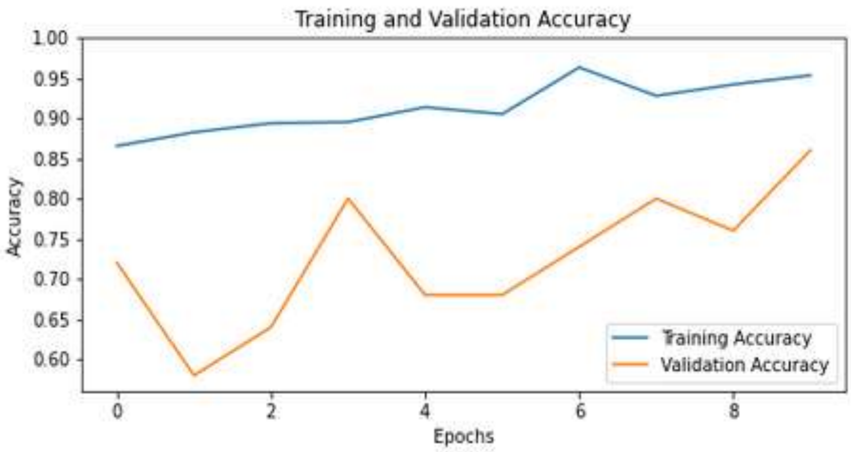
After the architecture is formed and the model fitting process is carried out, the algorithm will immediately conduct training on the previously prepared data. The iteration parameter carried out in this training process is 10 epochs with a batch size value of 32. So the training process will take place and be repeated 10 times to obtain feature extraction from the required features. The following are the results of the training process in this study as shown in the table below.

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Table 2.
Results of the 10 epoch training process

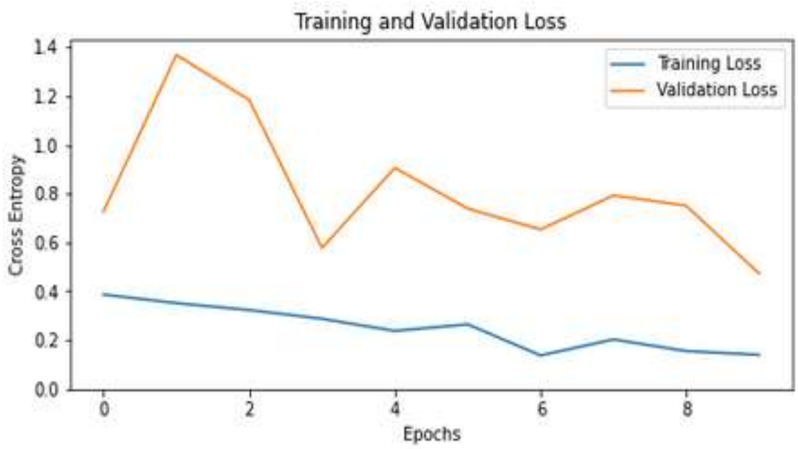
Epoch	Training		Validation	
	Accuracy(%)	Loss	Accuracy(%)	Loss
1	86,54	0,39	72,59	0,72
2	88,24	0,35	58	1,37
3	89,83	0,32	64	1,18
4	89,52	0,29	80	0,58
5	91,36	0,24	68	0,91
6	90,51	0,26	68	0,74
7	96,32	0,14	74	0,65
8	92,78	0,2	80	0,79
9	94,19	0,15	76	0,75
10	95,33	0,14	86	0,47

From the table above, it can be seen that the accuracy of the training process at epoch 10 is 95.3% with a loss value of 0.14. The validation accuracy value is 86% with a loss value of 0.47 at epoch 10. The accuracy and loss values can be expressed on graphs as shown in Figure 4 and Figure 5.



Figurr 4. Training dan Validation Accuracy

Figure 4. Shows that the highest accuracy value in the training process is 96.3% at the 7th epoch and the highest validation accuracy is 86% at the 10th epoch.



Figur 5. Training dan Validation Loss

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Figure 5. Shows the loss value for each epoch, in the training process the largest loss in the first epoch is 0.39 and the highest validation loss is 1.37 in the second epoch.

3.2. Result Proses Testing

The testing process uses test data as many as 84 images, and each individual there are 60 images. The accuracy results obtained in the testing process are 79.7%. From this calculation process, it can be seen that the results of this testing process reached 80% of the data tested and the results were correct. As a reinforcement of the results of the accuracy calculation in the testing process, it is carried out by testing all the testing data one by one. The results of the tests carried out can be seen in the table below. Below are the results of ear identification from the test results.

Table 3.
Ear Identification Trial Results

No	Input	Output	Status Identifikasi
1	000.png	000	Benar
2	001.png	006	Salah
3	002.png	002	Benar
4	003.png	003	Benar
5	004.png	000	Salah
6	005.png	005	Benar
7	006.png	006	Benar
8	007.png	007	Benar
9	008.png	008	Benar
10	009.png	009	Benar
11	010.png	008	Salah
12	011.png	011	Benar
13	012.png	009	Salah

Based on Table 3. it can be seen that there are 4 incorrect image data on individuals 001, 004, 010, and 012. This shows that the prediction results of the model on testing data show good results.

4. Conclusion

From the results of research carried out to identify ears based on images using the Convolutional Neural Network (CNN) method, it can be concluded that the CNN algorithm is quite good at ear identification. By doing training on 10 epochs, the training accuracy is 96.3% and the testing accuracy is 79.7%. This result is quite good considering the quality and amount of data obtained is not so good and the amount is not much.

References

Alshazly, Hammam, Christoph Linse, Erhardt Barth, and Thomas Martinetz. 2020. “Deep Convolutional Neural Networks for Unconstrained Ear Recognition.” IEEE Access: 170295–310.

Bengio, Yoshua, Courville Aaron, and Ian J Goodfellow. 2016. “Deep Learning: Adaptive Computation and Machine Learning.”

Hassaballah, M, Hammam A Alshazly, and Abdelmgeid A Ali. 2019. “Ear Recognition Using Local Binary Patterns : A Comparative Experimental Study.” Expert Systems With Applications 118: 182–200.

Hatif, Omar. 2019. “Datasets for Ear Detection and Recognition.”

Ilahiyah, Sarirotul, and Agung Nilogiri. 2018. “Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network.” : 49–56.

Jamil, Nursuriati et al. 2018. “Can Convolution Neural Network (CNN) Triumph in Ear Recognition of Uniform Illumination Invariant?” Indonesian Journal of Electrical Engineering and Computer Science 11(2): 558–66.

Mubarok, Hamdani. 2019. “IDENTIFIKASI EKSPRESI WAJAH BERBASIS CITRA MENGGUNAKAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK (CNN).”

Rao, G Sasibhushana, G Vimala Kumari, and B Prabhakara Rao. 2019. 2 Network for Biomedical Applications. Springer Singapore. http://dx.doi.org/10.1007/978-981-13-1595-4_9.

Ratnadewi, and Syafril Tua. 2015. “Realisasi Identifikasi Seseorang Dengan Biometrik Telinga Menggunakan Transformasi Hough.” EEJ 6(1): 85–94.

Raveane, William, Pedro Luis Galdámez, María Angélica, and González Arrieta. 2019. “Ear Detection and Localization with Convolutional Neural Networks in Natural Images and Videos.”

Resmi, K R, and G Raju. 2019. “A Novel Approach to Automatic Ear Detection Using Banana Wavelets and Circular

Hough Transform.” 2019 International Conference on Data Science and Communication (IconDSC): 1–5.
Sabet, Asmaa, Kareem Kamal, A Ghany, and Hesham Elmahdy. 2015. “Human Ear Recognition Using Geometrical Features Extraction.” Procedia - Procedia Computer Science 65(Iccmit): 529–37.
Saputra, Elin Panca, Indriyanti, and Supriatiningsih. 2020. “Classification Using Artifical Neural Network Method in Protecting Credit Fitness.” Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM) 3(1): 50–56. http://ejournal.uin-suska.ac.id/index.php/IJAIDM/article/view/9442.
Siregar, Alda Cendekia. 2017. “Identifikasi Telinga Berdasarkan Fitur Geometrik Dan KNN.” CYBERNETICS 01(01): 1–9.
Yuliani, Enie, Afifah Nur Aini, and Chan Uswatun Khasanah. 2019. “Perbandingan Jumlah Epoch Dan Steps per Epoch Pada Convolutional Neural Network Untuk Meningkatkan Akurasi Dalam Klasifikasi Gambar.” 5: 2–6.

Accredited “Rank 4”(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.