Phoneme Classification Optimization Using Backpropagation Neural Network and Principal Component Analysis

Clara Maria Livia Suitela[1], Aripin[2], Erika Dina Permata[3], Muzala Nakiatun Niza[4], Naeli Laelal Khiaroh[5]

[1,2,3,4,5] Biomedical Engineering Faculty of Engineering Universitas Dian Nuswantoro, Jl. Nakula No. 5-11 Semarang 50131 (tel.: 024-3517261; fax: 024-3520165; email: [1]513201900085@mhs.dinus.ac.id, [2]arifin@dsn.dinus.ac.id, [3]513201900086@mhs.dinus.ac.id, [4]513201900088@mhs.dinus.ac.id, [5]513201900080@mhs.dinus.ac.id)

[Received: 28 October 2022, Revised: 6 January 2023]
Corresponding Author: Clara Maria Livia Suitela

ABSTRACT — A phoneme is the smallest sound in a sentence that has no meaning but plays the most important role in meaning formation. Phoneme identification from a video that shows an actor speaking Indonesian sentences is an important part of developing visual-to-text applications. This application can translate mouth movements from a video into a series of Indonesian texts so that it can facilitate communication for the deaf. This study aims to optimize the performance of the classification process on image data, including as many as 32 phonemes from video extraction results so that they can be used to support the phoneme identification process to realize visual-to-text applications in Indonesian. The classification algorithm used in this study was neural network backpropagation. Some of the proposed efforts to optimize the performance of the classification process included using a comparison of the proportion of datasets, estimating the number of hidden layers, and reducing the dimensions of the dataset using the principal component analysis (PCA) method to reduce the amount of data that is considered less important without reducing the level of information. The dimensions of the data before reduction were 1280 × 7100 data matrices and 1280 × 50 data matrices after reduction. The accuracy results obtained in data optimization using the PCA were equal to 87.16% with a data proportion of 8 : 2 and fifty important data points were used in the data optimization process using the PCA.

KEYWORDS — Classification of Indonesian Phonemes, Neural Network Backpropagation, Optimization, Principal Component Analysis, Visual-to-Text.

I. INTRODUCTION

Etymologically, phonology comes from the Greek words “phone,” which means sound, and “logos,” which means science. Thus, it can be interpreted that phonology is the science of the system or element of language sounds that form a language utterance, starting from the way of pronunciation to the way the sound is conveyed to the listener [1]. The sounds studied in phonology are called phonemes.

According to the Kamus Besar Bahasa Indonesia, phonemes are the smallest units of sound capable of showing contrasts in meaning or suprasegmental sound units in the form of phonemic stresses, tones, or pauses [2]. Suprasegmental phonemes are a combination of oral communication and auditory understanding of linguistic teaching, such as sources of stress, accent, intonation, pitch, syllables, and sound deviations [3]. A phoneme can also be defined as the smallest sound in a speech that has no meaning but plays a crucial role in the formation of meaning. Phonemes are the object of study in phonetics, which is a branch of phonological studies that investigates the sound of language by focusing on its unction as a differentiator of meaning or words [4].

A word is formed from several phonemes that have an understandable meaning. For example, the word “accuracy” is made up of the phonemes /a/, /c/, /u/, /r/, and /y/. In this study, phoneme classification selected data in the form of images to be used as a classification dataset. Other researchers have previously used image data for classification and analysis, particularly of phonemes, with a different type of algorithm and with less than optimal accuracy [5]. In this study, different types of algorithms and various methods were used to produce a more optimal level of accuracy. The images were selected by taking videos of people saying one sentence at a time. After that, the video obtained was extracted into the desired form of image data. Phoneme classification analysis in this study used machine learning.

Machine learning is a branch of artificial intelligence that allows computers to study data and then create systems, models, or algorithms without having to write specific programs [6]. One of the methods used is to extract relevant data in order to produce information that can be interpreted [7]. The use of machine learning to solve data problems depends on the type of data problem to be solved, the number of variables, and the most suitable type of algorithmic model [8]. A machine learning workflow began with a dataset of variables to predict or classify. The input dataset was divided into training data and testing data. The training data functions as a training model, while the testing data serves as an evaluation model. Then, the model algorithm ran by spreading the data to be processed, writing instructions in code or program form that was used to solve problems, and dividing the data into input, processing, and output parts. The end result of dataset processing was the accuracy of the resulting data.

The accuracy calculations for the data in this study were repeated to obtain the best accuracy results. Optimization of phoneme data classification to determine the level of data accuracy was carried out using a neural network. Neural network is a series of algorithms that recognize the underlying relationships in a dataset by imitating the way the human brain operates [7]. Neural network is a computational model that mimics the performance principles of the human brain and can study sample data and map data input and output. Neural networks are usually used for classification, prediction, approximation, recognition, and association problems [9]. The network used in the optimization of phoneme classification was

p-ISSN 2301–4156 | e-ISSN 2460–5719

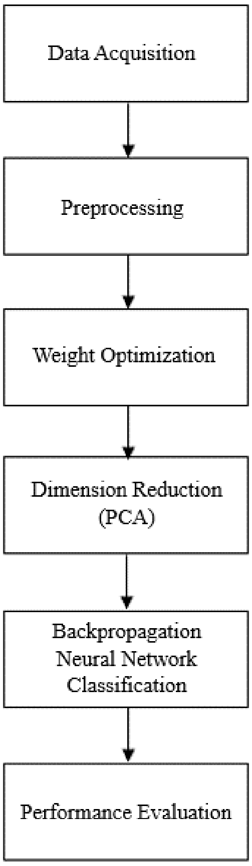

Figure 1. Phoneme accuracy research stages.

(a)

(b)

Figure 2. Process of a) taking videos of test person and b) taking screenshots of the video.

one of the most widely used networks, namely the multilayer perceptron neural network, which has a layer called "hidden" in the middle of the input and output layers. The hidden layer has variable properties that can be used by more than one hidden layer, meaning that more than one hidden layer can be used in one architecture according to the problem to be solved. Neural networks are used for image recognition of phoneme images, which will be classified for optimal accuracy [10].

The accuracy of the data generated from machine learning then utilizes the principal component analysis (PCA) method to simplify the dataset. PCA is a data and image processing methods that aim to reduce the amount of large data [11]. An analysis using the PCA method was carried out by normalizing by reducing the amount of data that was considered less important, then calculating the covariance matrix to reveal the relationship between variables from the input dataset.

Image data from .jpg files were used to represent 32 different types of phonemes in this study. The dataset was obtained from the data extraction process (RGB to grayscale). The results of the most optimal accuracy in the dataset were shown in a proportion of the amount of data that was considered important for analysis. On the optimization of phoneme classification data, phoneme identification was carried out, which could later be used as data propositions for making deaf assistant applications. The deaf application works as a communication tool by detecting lip movements while speaking and reading them as meaningful words. The phoneme classification optimization process is expected to yield the proportion of datasets with the highest accuracy.

II. METHODOLOGY

The stages carried out in this study consists of data acquisition, preprocessing, classification optimization using neural network backpropagation, and performance evaluation. These stages are presented in Figure 1. Data acquisition includes the processes of taking video, collecting images, and classifying images based on phonemes. The preprocessing stage includes cropping, resizing, RGB to grayscale processing, and the data split stage, which is the process of dividing fifty images for each phoneme in several comparisons. The next stage was classification optimization using neural network backpropagation, aiming to produce a more optimal accuracy value.

In order to improve the accuracy value of image reading to read mouth movements, methods and algorithms that can increase the accuracy value must be used. This study used the PCA method to produce a higher accuracy value compared to experiments without the PCA method. The feature extraction process using PCA was carried out after the preprocessing stage and the accuracy value was known. This stage was carried out to increase the accuracy value, which was not optimal.

A. DATA ACQUISITION

Data acquisition is a technique of taking or collecting data or images that will be processed using a computer [12]. The image collection process was carried out by recording a video in real-time with a test person saying a few sentences. Subsequently, a screenshot of the video results was taken to capture the facial image data needed when mentioning the phoneme. Finally, the image was separated by dividing or classifying the image from the mouth of the test person to every existing phoneme. The results of the image extraction were then used to select relevant images and eliminate irrelevant images. From the results of the image selection, the number of images obtained was 3,347 selected image data on 32 phonemes. From the 3,347 images, 1,600 images were taken,

Figure 3. Original image of the phoneme /e/.

Figure 4. Cropped image.

with fifty images for each phoneme, which were used in the classification process using neural network backpropagation.

Figure 2 shows the process of taking videos of the test person and taking screenshots of images. In this stage, the results of the video were carried out by focusing on taking the video only on the area around the test person's face and then converting it into the form of image data.

B. PREPROCESSING

Preprocessing is a process occurring after data acquisition. Preprocessing is conducted to carry out image measurements and obtain the necessary data samples [13]. The basic concepts and methods of preprocessing data can be categorized into three categories: data cleaning, data integration, and data reduction [14]. In this process, the preprocessing stage was carried out.

1) CROPPING

Cropping is the process of cutting an image from the original screenshot [15]. Image processing was done to improve image quality, such as by removing noise. Noise can be interpreted as unneeded information emerging in the image [15]. Cropping was done to focus on data extraction; the part that was cut was the part around the mouth with a size of 250 × 350 pixels. Figure 3 and Figure 4 show the original phoneme image and the cropped image, respectively.

2) RESIZE

After the cropping process was complete, the next step was to resize the cropped image. Resize is a method for resizing an image to a certain desired size [16]. As seen in Figure 5, the resizing process was carried out by reducing the pixel size of the phoneme image to a size of 71 × 100 pixels. The size of the pixels in the image is a reference to be used when extracting image data into datasets in nominal form. This stage is important to complete to get the optimal data size. The following are the resizing steps:

3) RGB TO GRAYSCALE

An image is a two-dimensional function f (x, y), where x and y are spatial coordinates and planes, and has a number of

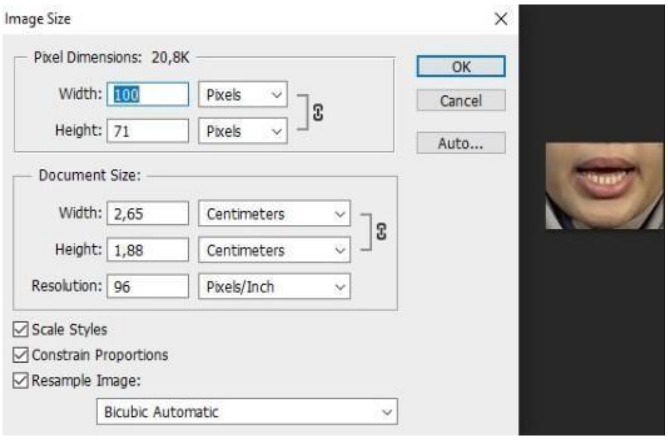

Figure 5. Process of resizing the image size of 71 x 100.

elements called pixels [17]. In this process, the color-changed image had dimensions of width (x) of 71 and length (y) of 100; the channel or matrix values in RGB consisted of two dimensions for each function, namely R (x, y), G (x, y), and B (x, y), which were then converted to grayscale using the GS algorithm with the rgb2gray function so that the conversion results of the GS image became one-dimensional and produced a dataset that was used in the search for data accuracy using neural network backpropagation [18]. A total of 1,280 images converted from RGB to grayscale produced 1280 × 7100 datasets, where 7,100 datasets were obtained from image dimensions of 71 × 100 pixels. The RGB to grayscale conversion was carried out simultaneously with the data extraction process. This stage was carried out to obtain a gray color from the image, aiming to remove redundant data contained in the image.

C. FEATURE EXTRACTION USING PCA

PCA was used to collect visual speech and image data, which was processed through feature extraction. Given the M image dataset from an image database with the image pronunciation of phonemes A_j = [A_1, A_2, ..., A_M], (j = 1, 2, ...,M), each image is converted into a 2-dimensional matrix of ($X_m x X_n$). Then the matrix is converted into a vector image T with size (U x 1), where U = ($X_m x X_n$) which produces a set of vector images with size (U x M) [19]:

$$T = [T_1 T_2 \ldots . T_M]. \quad (1)$$

Then, the arithmetic average results can be calculated from a vector image with size (U x 1) pixels by using (2).

$$\bar{A} = \frac{1}{M} \sum_{j=1}^{M} T_j \quad (2)$$

where M denotes the number of data points and $\sum_{j=1}^{M} T_j$ is the number of each row.

Then, the next step was to calculate the *ATrain* matrix, which was the result of the difference between each value of the T matrix and the $\bar{A}$ matrix, using (3):

$$ATrain = T_j - \bar{A}. \quad (3)$$

Next, the S_T covariant matrix was calculated using (4).

$$S_T = ATrain \; x \; ATrain'. \quad (4)$$

Based on the S_T covariant matrix, the eigenvalues (D) and eigenvectors (V) could be calculated. Eigenvalues are characteristic values of a square matrix. These values are taken from eigenvalues that are greater than 0. In this study, the

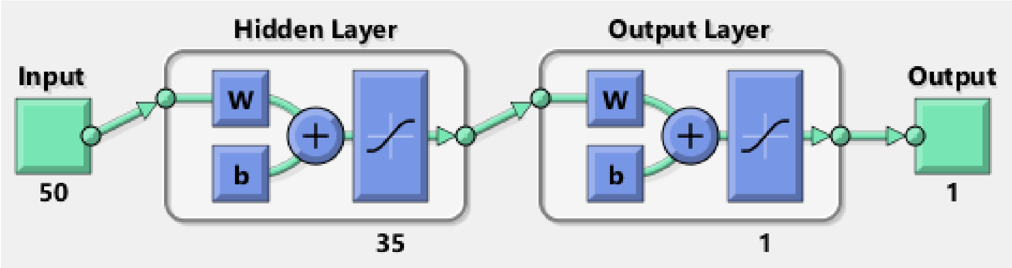

Figure 6. Optimal model architecture.

eigenvalues (D) and eigenvectors (V) were obtained using the eig () function, which was implemented using Matlab. The next step was to calculate the eigenfaces. Eigenfaces are characteristics of image data determined using (5):

$$Eigenfaces = ATrain \; x \; V \quad (5)$$

The next function of PCA was to reduce the characteristics that remained present in the image data by removing unimportant data features. Equation for calculating the PCA projection matrix is (6):

$$PCA\_Projected = Eigenfaces'x \quad (6)$$

PCA_Projected was a dataset that was used in the next classification process.

D. CLASSIFICATION

Neural network backpropagation is a model of artificial neural networks that uses a multilayer architecture to find optimal weights and minimize error values in the results of artificial neural networks. Classification in neural network backpropagation works in two stages, namely forward and backward calculations [20]. The use of backpropagation neural networks aims to determine the level of accuracy of successful data processing and to improve the performance of previous studies [21]. Data processing used neural network backpropagation by entering datasets that was used as samples. The sample dataset consisted of fifty images for each phoneme, which were classified into training data and testing data with a ratio of 6:4, 7:3, 8:2, and 9:1 data proportions. The first stage preceeding extraction process was to combine all the phoneme images in one folder, which would be included in the source code in Matlab. The image extraction results were in the form of datasets. They were then stored in Excel format and were processed using neural network backpropagation to produce the expected accuracy.

In this study, data classification used a backpropagation neural network, which was widely used in data prediction cases. Neural network backpropagation consists of three stages, namely, the advanced stage where the input and output layers are calculated forward and determined from the activation function [22], the backward stage where the desired output target state has a difference with the output network and an error that occurs so that the error is propagated backward, and the third is the stage of changing the weight values to reduce the possibility of errors that will occur [21], [22].

In a neural network there is an output layer that has linear neurons, a hidden layer, and an input layer that has nonlinear neurons and linear neurons, respectively, with $[x_1(k) \quad x_2(k) \quad \dots \quad x_n(k)]^T$ as input to the vector [23]. Output from j^{th} neuron in the hidden layer.

$$S_j(k) = F_j\left(\textstyle\sum_{i=1}^{n} w_{j,i}^{1,0}(k)x_i(k)\right) \quad (7)$$

TABLE I
TEST RESULTS WITH COMPARISON OF DATA PROPORTIONS

Data Proportion	Accuracy (%)
6:4	81.42
7:3	77.17
8:2	83.77
9:1	65.60

TABLE II
ACCURACY COMPARISON RESULTS

Results of Previous Studies [21]		Before Using PCA		After Using PCA	
Hidden Layer	%	Hidden Layer	%	Hidden Layer	%
4	20.59	10	76.33	25	85.62
5	11.76	15	83.77	35	87.16
20	52.94	25	76.57	40	84.42
37	32.35	35	71.88	50	82.33

where F_j (.) denotes the nonlinear sigmoidal activation function. The output of the neural network can be written as (8).

$$y(k) = \textstyle\sum_{i=1}^{n} w_{1,i}^{2,1}(k)S_j(k). \quad (8)$$

Then, the error detector can be written like equation (9).

$$e(k) = y(k) - d(k) \quad (9)$$

where $d(k)$ is the required reference signal for $y(k)$ and the point $e(k) = 0$ is the equilibrium point of the dynamic disturbance system. The weights of the neural network are adjusted according to the following equation:

$$w_{1,j}^{2,1}(k) = \frac{\beta^{-\frac{k}{2}}e(k-1)+d(k)}{nS_j(k-1)} \quad (10)$$

$$w_{j,j}^{1,0}(k) = \frac{1}{nx_i(k)}G_j\left(\frac{\beta^{-\frac{k}{2}}e(k-1)+d(k)}{mw_{1,j}^{2,1}(k)}\right). \quad (11)$$

$G_j - = F_j^{-1}(-)$ and β are the learning speeds of the network. Equations (10) and (11) guarantee the asymptotic convergence of the error detection output $e(k)$ as storage.

The weighting in this test used an activation function that was influenced by the number of layer neurons, as shown in (12).

$$z\_in_j = v_{j0} + \textstyle\sum_{i=1}^{n} x_i \,.\, v_{ij} \quad (12)$$

where

z_in_j = value to calculate the hidden layer
x_i = i's input layer value
v_{ij} = the weight between the input layer and hidden layer.

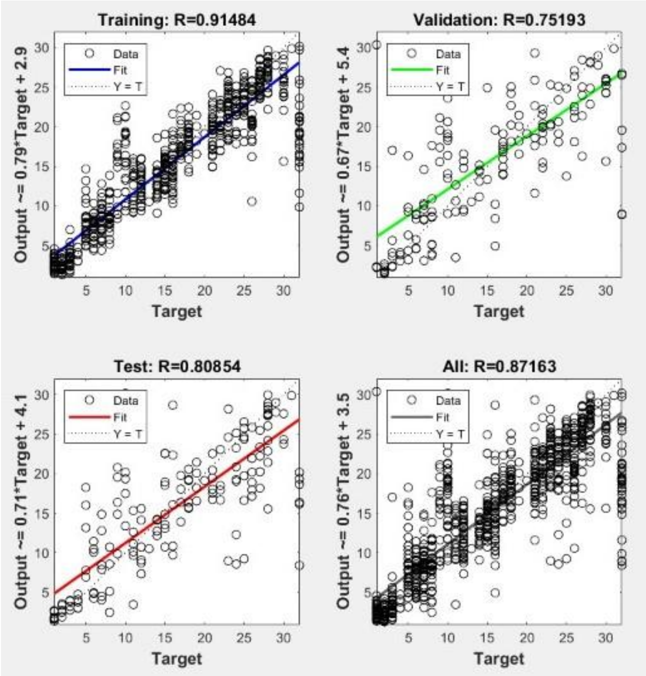

Figure 7. Optimization graph using PCA.

E. WEIGHTING OPTIMIZATION

The testing process was carried out using images, with a total of fifty images for each phoneme. An analysis process was carried out on the accuracy results for the training process with or without weighting optimization. The weighting optimization process was carried out using the technique of dividing the proportion of data on the results of feature extraction with PCA. The process of finding the best parameters was carried out by looking at the results of the accuracy values for each proportion of the data used and selecting the one with the highest accuracy value. The distribution of the proportion of data to be used in this study is as follows:

- 60% training data and 40% testing data.
- 70% training data and 30% testing data.
- 80% training data and 20% testing data.
- 90% training data and 10% testing data

In this study, data optimization was carried out by obtaining an optimal model with one hidden layer of neural network architecture, as shown in Figure 6. This optimal model used fifty inputs and 35 neurons in the hidden layer.

III. RESULT AND DISCUSSION

The extracted data from the RGB to grayscale results was used in the process of optimizing the phoneme classification data, which was then divided into training data and testing data. The data to be processed was the data from the extraction results. The 1,280 pieces of data that had been extracted produced a 7100 x 1280 dataset that would be processed using neural network backpropagation. Backpropagation of neural networks was used as a tool for data classification to increase data accuracy. This study used three stages to optimize the performance of the classification process: comparing the proportion of datasets, estimating the number of hidden layers, and reducing dataset dimensions using the principal component analysis method to reduce data dimensions while retaining as much information as possible from the original dataset [24].

The first stage was to obtain accuracy from the phoneme dataset by conducting trial and error on different proportions of training data and testing data, so that the resulting proportion of data was100%. The data proportions used in this study consisted of four proportions, namely the proportions of 6:4, 7:3, 8:2, and 9:1. Table I shows the accuracy results obtained from each comparison of training and testing data. The results of the accuracy of the proportion of data with a ratio of 6:4 were 81.42%; the results of the accuracy of the proportion of data with a ratio of 7:3 were 77.17%; the results of the accuracy of the proportion of data with a comparison of 8:2 were 83.77%; and the results of the accuracy of the proportion of data with a comparison of 9:1 were 65.60%. The first stage was to get optimal results on the number of data proportions (8:2). The proportion of data (8:2) was considered optimal at the level of dataset accuracy. However, in this first stage, a safe level of accuracy had not been obtained because it was still below 85%. A classification is acceptable if it is more than or equal to 85%.

The second stage was to look for optimization of dataset accuracy, namely using a method for estimating the number of hidden layers in the classification process using neural network backpropagation. Table II shows a comparison between the accuracy results obtained in previous studies and the accuracy results before and after using the PCA function. This stage was conducted by comparing previous studies, different data proportions, and the PCA function to determine the number of hidden layers. Based on the previous research’s results,

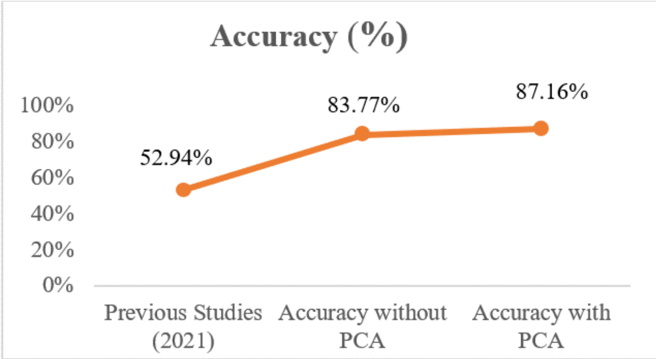

Figure 8. Chart of Data accuracy comparison.

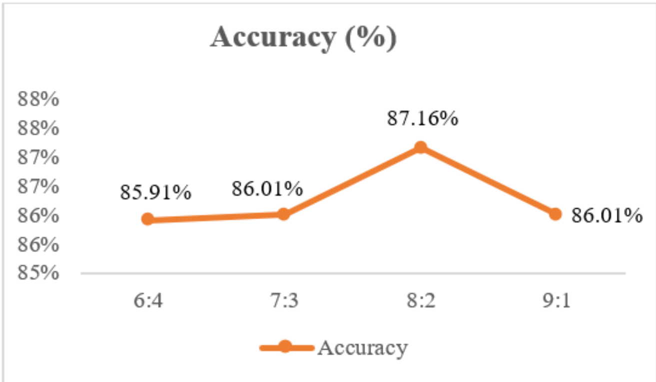

Figure 9. Chart of PCA data composition accuracy.

TABLE III
ACCURACY COMPARISON RESULTS

Data Proportion	Hidden Layer	Important PCA Data	Accuracy (%)
6 : 4	80	50	85.91
7 : 3	40	30	86.01
8 : 2	35	50	87.16
9 : 1	60	50	86.01

dividing the data in the backpropagation method into training and testing data can increase the epoch value and time, but it cannot constantly improve the accuracy results [21].

According to Tabel II, the highest data optimization is in the process of using the PCA function. In the PCA process, the phoneme image extraction data was simplified by eliminating unimportant data features, so that the length of the resulting matrix was less than the initial matrix of data extraction. The dataset optimization process using hidden layer weighting produced a better accuracy value when compared to testing without hidden layer weighting. The training dataset influences the optimal number of hidden layers.

Following the processing to determine the effect of the hidden layer on accuracy, a trial-and-error test was carried out on the number of hidden layers used to obtaine the expected accuracy value. Figure 7 exhibits the accuracy results after using the PCA function with 35 hidden layers and 200 epochs of data.

According to previous research's results, increasing the results of the accuracy value by dividing the dataset into training data and testing data and changing the hidden layer used remained producing a relatively low accuracy value of 52.94%. Training data weighting produced a higher accuracy value of 83.77%, in the proportion of training data to testing data of 8:2 using hidden layer 15 [21]. Then, the highest results were optimized using the PCA function to produce an accuracy value of 85.62% with hidden layer 25, 87.16% with hidden layer 35, 84.42% with hidden layer 40, and 82.33% with hidden layer 50. Therefore, the highest data optimization process can be seen in Figure 8, which shows an increase in accuracy from the results of data weighting to the process using the PCA function.

After obtaining the results of the increased accuracy value using the PCA function at a ratio of 8:2 for the data proportions, the data was tested again on the proportions of the 6:4, 7:3, and 9:1 datasets. The testing was done by reducing the amount of data by performing trial and error on important data and the number of hidden layers used to obtain the expected accuracy value, as shown in Table III. The highest accuracy value is still obtained in the proportion of 8:2 data with hidden layer 35, as shown in Table III, and the number of important data matrices resulting from PCA reduction is 50 x 1280. The important PCA data used in Table III is the total size of the reduced data matrix, where in the PCA configuration used in this study, there was repetition of classification based on the number of different proportions of the feature extraction matrix to obtain the most optimal accuracy value. In the proportion of data 6:4 with reduced data of 50 and hidden layers of 80, the accuracy was 85.91%; in the proportion of data 7:3 with reduced data of 30 and hidden layers of 40, the accuracy was 86.01%; and in the proportion of data 9:1 with reduced data of 50 and hidden layers of 60, the accuracy was 86.91%. Based on the tests that has been carried out on all divisions of the existing data proportions, it can be seen that the optimal accuracy value is obtained when the data is reduced using the PCA function and the number of hidden layers used has a ratio that is not too significant.

Figure 9 shows a graph of the accuracy results using PCA from several data proportions. The highest percentage of accuracy was obtained at 87.16% in the proportion of data (8:2).

IV. CONCLUSION

Based on the results of the discussion and testing that have been done, it can be concluded that phoneme classification using the PCA feature can improve the accuracy of the test results. The accuracy results obtained using PCA were 87.16%. The results of the neural network backpropagation test were 83.77%; meanwhile, the results of previous studies were 52.94%. These findings show that the use of the PCA feature can function as dataset optimization.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTION

Conceptualization, methodology, software, resources, validation, writing, Clara Maria Livia Suitela, Erika Dina Permata, Muzalfa Nakiatun Niza, and Naeli Laelal Khiaroh; formal analysis, data accuracy, funding acquisition, Aripin.

REFERENCES

[1] U. Kulsum, "Masalah Bunyi dalam Bahasa Masyarakat Indonesia," *Caraka*, Vol. 10, No. 1, pp. 299–302, Feb. 2021, doi: 10.1017/s1742058x13000143.

[2] (2022) "KBBI Online," [Online], http://kbbi.co.id/cari?kata=fonem, access date: 11-Jul-2022.

[3] M. Khaleqyar and M. Qaseem Kashaf, "The Importance of Suprasegmental Features in the Structure and Meaning of Pashto Sentence," *Sprin J. Arts, Humanit. Soc. Sci.*, Vol. 1, No. 9, pp. 31–40, Sep. 2022, doi: 10.55559/sjahss.v1i09.54.

[4] F. Nur, D.E. Aryanto, and N.I. Islami, "Analisis Segmentasi Silabel dan Fonem dalam Kalimat Perintah (Kajian Fonologi)," *Ghâncaran: J. Pendidik. Bhs. dan Sastra Indones.*, Vol. 3, No. 1, pp. 16–38, 2021, doi: 10.19105/ghancaran.v3i1.4424.

[5] N. Mamuriyah and J. Sumantri, "Penerapan Metode Convolution Neural Network (CNN) Pada Aplikasi Automatic Lip Reading," *J. Inform. Telecommun. Eng.*, Vol. 6, No. 1, pp. 276–287, Jul. 2022, doi: 10.31289/jite.v6i1.7523.

[6] I.D. Id, *Machine Learning: Teori, Studi Kasus dan Implementasi Menggunakan Python*, 1st ed. Pekanbaru, Indonesia: UR PRESS, 2021, doi: 10.5281/zenodo.5113507.

[7] B. Mahesh, "Machine Learning Algorithms - A Review," *Int. J. Sci., Res. (IJSR)*, Vol. 9, No. 1, pp. 381–386, Jan. 2020, doi: 10.21275/ART20203995.

[8] A. Roihan, P.A. Sunarya, and A.S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review Paper," IJCIT (*Indonesian J. Comput. Inf. Technol.*), Vol. 5, No. 1, pp. 75–82, 2020, doi: 10.31294/ijcit.v5i1.7951.

[9] R.H. Abiyev and M.K.S. Ma'aitah, "Deep Convolutional Neural Networks for Chest Diseases Detection," *J. Healthc. Eng.*, Vol. 2018, pp. 1–11, Aug. 2018, doi: 10.1155/2018/4168538.

[10] D. Aprilla C., D.A. Baskoro, L. Ambarwati, and I.W.S. Wicaksana, "Belajar Data Mining dengan RapidMiner," in *Belajar Data Mining dengan RapidMiner*, R. Sanjaya, Ed. Jakarta, Indonesia: Gramedia Pustaka Utama, 2013, pp. 1–125.

[11] O.D. Nurhayati, D. Eridani, and A. Ulinuha, "Ekstraksi Ciri Orde Pertama dan Metode Principal Component Analysis untuk Mengidentifikasi Jenis Telur Ayam Kampung dan Ayam Arab," *JSINBIS (J. Sist. Inf. Bisnis)*, Vol. 9, No. 2, pp. 133-140, Nov. 2019, doi: 10.21456/vol9iss2pp133-140.

[12] T. Yulianti, M. Telaumbanua, H.D. Septama, and H. Fitriawan, "Pengaruh Seleksi Fitur Citra Terhadap Klasifikasi Tingkat Kesegaran Daging Sapi Lokal," *J. Tek. Pertan. Lampung*, Vol. 10, No. 1, pp. 85–95, Mar. 2021, doi: http://dx.doi.org/10.23960/jtep-l.v10.i1.85-95.

[13] A. Zamfiroiu, R. Boncea, and I. Petre, "Data Acquisition Process for Smart Systems," *Int. J. Res., Anal. Rev.*, Vol. 6, No. 2, pp. 553–556, Jun. 2019.

[14] A. Riadi and R. Sulaehani, "Analisis Implementasi Preprocessing dengan Otsu-Gaussian pada Pengenalan Wajah," *ILKOM J. Ilm.*, Vol. 11, No. 3, pp. 200–205, Dec. 2019, doi: 10.33096/ilkom.v11i3.457.200-205.

[15] Yuhandri, "Perbandingan Metode Cropping pada Sebuah Citra untuk Pengambilan Motif Tertentu pada Kain Songket Sumatera Barat," *J.*

KomtekInfo (Komput. Teknol. Inf.), Vol. 6, No. 1, pp. 96–105, Jun. 2019, doi: 10.35134/komtekinfo.v6i1.45.

[16] M. Zen, "Perbandingan Metode Dimensi Fraktal dan Jaringan Syaraf Tiruan Backpropagation dalam Sistem Identifikasi Sidik Jari pada Citra Digital," *JiTEKH (Jurnal Ilm. Teknol. Harapan)*, Vol. 7, No. 2, pp. 42–50, Sep. 2019, doi: https://doi.org/10.35447/jitekh.v7i02.80.

[17] M.A. Nasution *et al.*, "Determinasi Nilai RGB dan Grayscale pada Citra Tandan Buah Segar (TBS Kelapa Sawit (Elaeis Guineensis Jacq.) Menggunakan Matlab," *J. Pen. Kelapa Sawit*, Vol. 30, No. 1, pp. 37–48, Apr. 2022, doi: https://doi.org/10.22302/iopri.jur.jpks.v30i1.159.

[18] H.A. Rabbani, M.A. Rahman, and B. Rahayudi, "Perbandingan Ruang Warna RGB dan HSV dalam Klasifikasi Kematangan Biji Kopi," *J. Pengemb. Teknol. Inf., Ilm. Komput.*, Vol. 5, No. 6, pp. 2243–2248, Jun. 2021.

[19] Arifin, Muljono, S. Sumpeno, and M. Hariadi, "Towards Building Indonesian Viseme: A Clustering-Based Approach," *2013 IEEE Int. Conf. Comput. Intell., Cybernetics (CYBERNETICSCOM)*, 2013, pp. 57–61, doi: 10.1109/CyberneticsCom.2013.6865781.

[20] M.N.D. Sawitri, I.W. Sumarjaya, and N.K.T. Tastrawati, "Peramalan Menggunakan Metode Backpropagation Neural Network," *E-J. Mat.*, Vol. 7, No. 3, pp. 264–270, Sep. 2018, doi: 10.24843/mtk.2018.v07.i03.p213.

[21] S.M. Damanik, *et al.*, "Optimasi Data Menggunakan Teknik Backpropagation dalam Meningkatkan Hasil Nilai Akurasi," *Punyaku sementara Sem. Nas. Sains, Teknol. Inf. (SENSASI)*, 2021, pp. 657–662.

[22] H. Okprana, M.R. Lubis, and J.T. Hadinata, "Prediksi Kelulusan TOEFL Menggunakan Metode Resilient Backpropagation," *J. Edukasi, Pen. Inform.*, Vol. 6, No. 2, pp. 275–279, Aug. 2020, doi: 10.26418/jp.v6i2.41224.

[23] M.S. Aftab and F. Aftab, "A Study on Lyapunov Function Backpropagation Algorithm for Suitability as Neuro-Adaptive Inverse Controller," *2018 IEEE 21st Int. Multi-Topic Conf. (INMIC)*, 2018, pp. 1–4, doi: 10.1109/INMIC.2018.8595664.

[24] U. Nuha and N. Rochmawati, "Klasifikasi Kesahihan Hadits Berdasarkan Perawi Hadits Menggunakan Principal Component Analysis (PCA) dan Backpropagation Neural Network (BPNN)," J. Inform., Comput. Sci., Vol. 1, No. 3, pp. 138–143, Mar. 2020, doi: 10.26740/jinacs.v1n03.p138-143.