

# Optimasi Klasifikasi Fonem Menggunakan *Backpropagation Neural Network* dan *Principal Component Analysis*

Clara Maria Livia Suitela<sup>1</sup>, Aripin<sup>2</sup>, Erika Dina Permata<sup>3</sup>, Muzalfa Nakiatun Niza<sup>4</sup>, Naeli Laelal Khiaroh<sup>5</sup>

<sup>1,2,3,4,5</sup> Jurusan Teknik Biomedis Fakultas Teknik Universitas Dian Nuswantoro, Jl. Nakula No. 5-11 Semarang 50131 (telp: 024-3517261; fax: 024-3520165; e-mail: <sup>1</sup>513201900085@mhs.dinus.ac.id; <sup>2</sup>arifin@dsn.dinus.ac.id; <sup>3</sup>513201900086@mhs.dinus.ac.id; <sup>4</sup>513201900088@mhs.dinus.ac.id; <sup>5</sup>513201900080@mhs.dinus.ac.id)

[Diterima; 28 Oktober 2022, Revisi: 6 Januari 2023]

Corresponding Author: Clara Maria Livia Suitela

**INTISARI** — Fonem merupakan bunyi terkecil dari suatu ucapan yang tidak memiliki pengertian, tetapi memiliki peranan terpenting untuk membentuk suatu arti. Identifikasi fonem dari sebuah video tentang seorang aktor yang sedang mengucapkan kalimat-kalimat berbahasa Indonesia merupakan bagian penting dalam pengembangan aplikasi *visual-to-text*. Aplikasi ini dapat menerjemahkan gerakan mulut dari sebuah video menjadi rangkaian teks berbahasa Indonesia, sehingga dapat membantu mempermudah komunikasi bagi para penyandang tunarungu. Penelitian ini bertujuan untuk mengoptimasi kinerja proses klasifikasi terhadap data citra sebanyak 32 fonem dari hasil ekstraksi video, sehingga dapat digunakan untuk mendukung proses identifikasi fonem untuk mewujudkan aplikasi *visual-to-text* dalam bahasa Indonesia. Algoritme klasifikasi yang digunakan dalam penelitian ini adalah *backpropagation neural network*. Beberapa upaya yang diusulkan untuk mengoptimasi kinerja proses klasifikasi di antaranya adalah dengan menggunakan perbandingan proporsi *dataset*, estimasi jumlah lapisan tersembunyi, dan reduksi dimensi *dataset* menggunakan metode *principal component analysis* untuk mengurangi jumlah data yang dianggap kurang penting tanpa mengurangi tingkat informasinya. Dimensi data sebelum direduksi sebesar  $1.280 \times 7.100$  matriks data dan setelah direduksi menjadi  $1.280 \times 50$  matriks data. Hasil akurasi yang diperoleh pada optimalisasi data menggunakan *principal component analysis* yaitu sebesar 87,16% dengan proporsi data 8:2 dan digunakan sebanyak 50 data penting dalam proses optimalisasi data menggunakan *principal component analysis*.

**KATA KUNCI** — Klasifikasi Fonem Bahasa Indonesia, *Backpropagation Neural Network*, Optimasi, *Principal Component Analysis*, *Visual-to-Text*.

## I. PENDAHULUAN

Secara etimologi, fonologi berasal dari bahasa Yunani, yaitu *phone* yang berarti bunyi dan *logos* yang berarti ilmu. Maka, fonologi dapat diartikan sebagai ilmu tentang sistem atau unsur bunyi bahasa yang membentuk suatu ujaran bahasa, mulai dari cara pelafalan hingga cara penyampaian suara kepada pendengar [1]. Bunyi yang dipelajari dalam fonologi inilah yang disebut dengan fonem.

Menurut KBBI, fonem merupakan satuan bunyi terkecil yang mampu menunjukkan kontras makna atau suprasegmental satuan bunyi yang berupa tekanan, nada, atau jeda yang fonemis [2]. Fonem suprasegmental merupakan kombinasi pengajaran linguistik yang memiliki komunikasi lisan dan pemahaman pendengaran, seperti sumber tekanan, aksen, intonasi dan tinggi nada, suku kata, serta penyimpangan bunyi [3]. Fonem dapat diartikan juga sebagai bunyi terkecil dari suatu ucapan yang tidak memiliki pengertian, tetapi mempunyai peranan terpenting untuk membentuk suatu arti. Fonem merupakan objek kajian dari fonemis, yang merupakan cabang kajian fonologi yang mengkaji bunyi bahasa dengan memperhatikan fungsinya sebagai pembeda makna atau kata [4].

Suatu kata terbentuk dari beberapa fonem yang membentuk satu arti yang dapat dipahami. Misalnya, kata “akurasi” terbentuk dari fonem /a/, /k/, /u/, /t/, /s/, dan /i/. Pada penelitian ini, klasifikasi fonem memilih data dalam bentuk citra yang akan digunakan sebagai *dataset* klasifikasi. Penggunaan data citra sebagai klasifikasi dan analisis, khususnya fonem, sebelumnya telah dilakukan menggunakan jenis algoritme yang berbeda dan menghasilkan akurasi yang belum optimal [5]. Oleh karena itu, dalam penelitian ini digunakan jenis algoritme berbeda dan berbagai metode untuk menghasilkan akurasi yang

lebih optimal. Pemilihan citra didapatkan dari pengambilan video orang yang sedang mengucapkan kalimat dalam satu waktu. Kemudian, video yang didapatkan diekstraksi menjadi bentuk data citra yang diinginkan. Analisis klasifikasi fonem pada penelitian ini dilakukan menggunakan *machine learning*.

*Machine learning* adalah salah satu cabang ilmu kecerdasan buatan yang memungkinkan komputer untuk mempelajari data, kemudian membuat sistem, model, atau algoritme tanpa harus menuliskan program secara spesifik [6]. Salah satu cara yang digunakan adalah mengekstrak data yang relevan agar menghasilkan informasi yang dapat diinterpretasikan [7]. Penggunaan *machine learning* untuk menyelesaikan masalah data tergantung pada jenis masalah data yang akan diselesaikan, jumlah variabel, dan jenis model algoritme yang paling cocok [8]. Alur kerja *machine learning* dimulai dari *dataset* yang berisi variabel untuk diprediksi atau diklasifikasikan. *Dataset* masukan dibagi menjadi data latih (*training*) dan data uji (*testing*). Data latih berfungsi sebagai model pelatihan, sedangkan data uji sebagai model evaluasi. Lalu, algoritme model berjalan dengan cara menyebarkan data yang akan diolah dengan menuliskan instruksi dalam bentuk kode atau program yang akan digunakan untuk menyelesaikan masalah dan membagi data menjadi bagian masukan (*input*), proses, dan keluaran (*output*). Hasil akhir dari pemrosesan *dataset* adalah akurasi data yang dihasilkan.

Perhitungan akurasi data pada penelitian ini dilakukan secara berulang untuk mendapatkan hasil akurasi yang paling baik. Optimasi klasifikasi data fonem untuk mengetahui tingkat akurasi data dilakukan menggunakan jaringan saraf (*neural network*). Jaringan saraf merupakan serangkaian algoritme untuk mengenali hubungan yang mendasar dalam suatu *dataset* dengan meniru cara otak manusia beroperasi [7]. Jaringan saraf

adalah sebuah model komputasi yang meniru prinsip kinerja otak manusia yang dapat mempelajari data sampel dan memetakan masukan dan keluaran data. Jaringan saraf biasanya dimanfaatkan untuk masalah klasifikasi, prediksi, pendekatan, pengenalan, dan asosiasi [9]. Jenis jaringan yang digunakan pada optimasi klasifikasi fonem adalah salah satu jaringan yang umum digunakan, yaitu *multilayer perceptron neural network*. Jenis jaringan saraf ini memiliki lapisan yang disebut dengan lapisan tersembunyi (*hidden layer*), yang berada di tengah lapisan masukan (*input layer*) dan lapisan keluaran (*output layer*). Lapisan tersembunyi memiliki sifat variabel yang dapat digunakan oleh lebih dari satu lapisan tersembunyi. Hal ini berarti lebih dari satu lapisan tersembunyi dapat digunakan dalam satu arsitektur, sesuai dengan masalah yang akan diselesaikan. Jaringan saraf digunakan dalam pengenalan citra (*image recognition*) dari citra fonem yang akan diklasifikasi untuk memperoleh akurasi yang optimal [10].

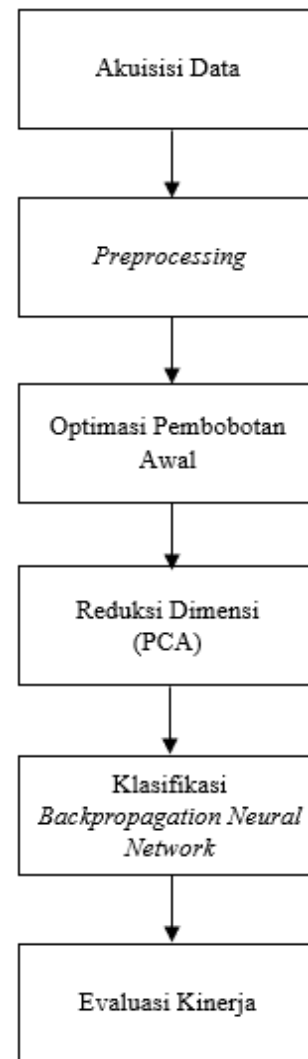
Akurasi data yang dihasilkan dari *machine learning* kemudian memanfaatkan metode *principal component analysis* (PCA) untuk menyederhanakan *dataset*. PCA adalah suatu metode yang sering digunakan dalam proses pengolahan data maupun pengolahan citra, yang bertujuan untuk mereduksi jumlah data yang berukuran besar [11]. Analisis menggunakan metode PCA dilakukan dengan proses normalisasi, dengan cara mengurangi jumlah data yang dianggap kurang penting, kemudian menghitung matriks kovariansnya sehingga hubungan antarvariabel dari *dataset* masukan menjadi tampak.

Pada penelitian ini digunakan data citra *file .jpg* yang merepresentasikan 32 jenis fonem yang berbeda. *Dataset* diperoleh dari proses ekstraksi data (RGB ke *grayscale*). Hasil akurasi yang paling optimal dalam *dataset* ditunjukkan pada proporsi jumlah data yang dianggap penting untuk dianalisis. Pada penelitian optimasi data klasifikasi fonem ini dilakukan identifikasi fonem yang nantinya dapat digunakan sebagai proposi data pembuatan aplikasi pembantu tunarungu. Aplikasi tunarungu memiliki fungsi sebagai alat bantu komunikasi untuk menerjemahkan bahasa dengan mendeteksi gerakan bibir saat berbicara sehingga dapat dibaca dalam susunan kata yang memiliki arti. Dari proses optimasi klasifikasi fonem ini diharapkan diperoleh proposisi *dataset* dengan akurasi paling optimal.

## II. METODOLOGI

Tahapan yang dilaksanakan dalam penelitian ini terdiri atas proses akuisisi data, *preprocessing*, optimasi klasifikasi menggunakan *backpropagation neural network*, dan evaluasi kinerja. Tahapan tersebut disajikan pada Gambar 1. Akuisisi data mencakup proses pengambilan video, pengumpulan citra, dan penggolongan citra berdasarkan fonem. Tahap *preprocessing* mencakup *cropping*, *resize*, proses konversi RGB ke *grayscale*, dan tahap *data split* yang merupakan proses pembagian citra sejumlah 50 citra pada tiap fonem dalam beberapa perbandingan. Tahap selanjutnya adalah optimasi klasifikasi menggunakan *backpropagation neural network* yang bertujuan untuk menghasilkan nilai akurasi yang lebih optimal.

Dalam proses untuk meningkatkan nilai akurasi dari pembacaan citra untuk membaca gerakan mulut, diperlukan perbaikan dengan menggunakan metode maupun algoritme yang dapat meningkatkan nilai akurasi. Penelitian ini menggunakan metode PCA untuk menghasilkan nilai akurasi yang lebih tinggi dibandingkan dengan percobaan tanpa



Gambar 1. Tahapan penelitian akurasi fonem.

metode PCA. Proses ekstraksi fitur menggunakan PCA dilakukan setelah tahap *preprocessing* dan nilai akurasi diketahui. Tahap ini dilakukan untuk meningkatkan nilai akurasi yang belum optimal.

### A. AKUISISI DATA

Akuisisi data adalah teknik pengambilan atau pengumpulan data atau citra yang akan diproses dan diolah menggunakan komputer [12]. Proses pengumpulan citra dilakukan dengan pengambilan video secara *real-time* menggunakan orang yang mengucapkan beberapa kalimat. Kemudian, data citra wajah yang dibutuhkan ketika menyebutkan fonem akan diambil dengan cara melakukan tangkapan layar (*screenshot*) dari hasil video. Setelah diperoleh, citra dipisahkan dengan cara membagi atau menggolongkan citra dari mulut orang ke dalam tiap fonem yang ada. Hasil ekstraksi citra tersebut selanjutnya diseleksi untuk memilih citra yang relevan dan mengeliminasi citra yang tidak relevan. Dari hasil seleksi citra didapatkan sebanyak 3.347 data citra seleksi pada 32 fonem. Dari 3.347 data citra tersebut, kemudian diambil sebanyak 1.600 citra, dengan 50 citra pada tiap fonem, yang digunakan dalam proses klasifikasi menggunakan *backpropagation neural network*.

Gambar 2 menunjukkan proses pengambilan video orang dan proses tangkapan layar citra. Dalam tahap ini, pengambilan video hanya difokuskan pada daerah sekitar wajah lalu diubah dalam bentuk data citra.



Gambar 2. Proses mendapatkan citra, (a) pengambilan video, (b) proses tangkapan layar dari video.



Gambar 3. Citra asli fonem "e".



Gambar 4. Citra hasil cropping.

## B. PREPROCESSING

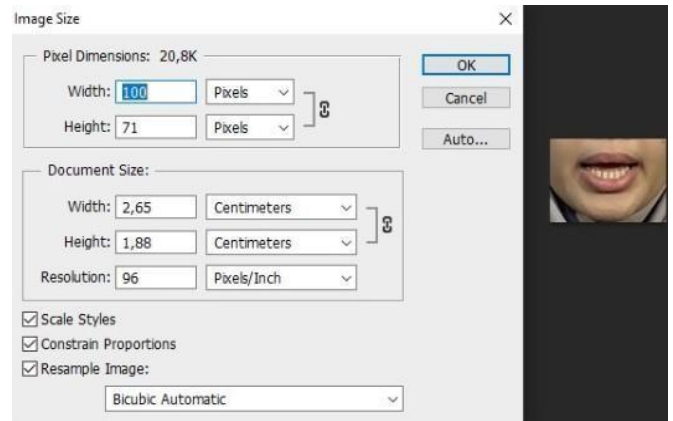
*Preprocessing* merupakan proses yang dilakukan setelah akuisisi data. *Preprocessing* bertujuan melakukan pengukuran citra dalam mendapatkan sampel data yang diperlukan [13]. Konsep dasar *preprocessing* dan metode untuk *preprocessing* data dapat dikategorikan menjadi pembersihan data, integrasi data, dan reduksi data [14]. Tahapan-tahapan dalam *preprocessing* dijelaskan sebagai berikut.

### 1) CROPPING

*Cropping* merupakan proses memotong citra dari hasil tangkapan layar [15]. Pengolahan hasil citra ini dilakukan untuk memperbaiki kualitas citra seperti menghilangkan *noise*. *Noise* dapat diartikan sebagai informasi dalam citra yang tidak diperlukan [15]. *Cropping* dilakukan untuk memfokuskan ekstraksi data. Bagian yang dipotong merupakan bagian sekitar mulut dengan ukuran  $250 \times 350$  piksel. Gambar 3 dan Gambar 4 menunjukkan citra asli fonem dan citra hasil *cropping*.

### 2) RESIZE

Setelah proses *cropping* selesai dilakukan, proses selanjutnya adalah *resize* citra hasil *cropping*. *Resize* merupakan metode untuk mengubah ukuran citra menjadi ukuran tertentu yang diinginkan [16]. Seperti yang terlihat pada Gambar 5, proses *resize* dilakukan dengan pengurangan ukuran piksel pada citra fonem dengan ukuran  $71 \times 100$  piksel. Ukuran piksel pada citra ini adalah acuan yang akan digunakan ketika menjalankan ekstraksi data citra menjadi *dataset* dalam bentuk nominal. Tahap ini penting dilakukan untuk mendapatkan ukuran data yang optimal.



Gambar 5. Proses *resize* citra ukuran  $71 \times 100$ .

### 3) KONVERSI RGB KE GRAYSCALE

Citra merupakan fungsi dua dimensi  $f(x,y)$ , dengan  $x$  dan  $y$  merupakan koordinat spasial dan bidang, serta memiliki sejumlah elemen yang disebut dengan piksel [17]. Pada proses ini, citra yang diubah warnanya memiliki dimensi lebar ( $x$ ) 71 piksel dan panjang ( $y$ ) 100 piksel. Nilai *channel* atau matriks pada RGB terdiri atas dua dimensi pada masing-masing fungsi, yaitu:  $R(x,y)$ ,  $G(x,y)$ , dan  $B(x,y)$ , yang kemudian akan diubah menjadi *grayscale* menggunakan algoritme GS dengan fungsi *rgb2gray*, sehingga hasil konversi citra GS menjadi satu dimensi dan menghasilkan *dataset* yang akan digunakan dalam proses pencarian akurasi data menggunakan *backpropagation neural network* [18]. Sebanyak 1.280 gambar citra yang dikonversi dari RGB ke *grayscale* menghasilkan  $1.280 \times 7.100$  *dataset*, dengan 7.100 *dataset* didapatkan dari dimensi citra  $71 \times 100$  piksel. Proses konversi RGB ke *grayscale* dilakukan bersamaan dengan proses ekstraksi data. Tahap ini dilakukan untuk mendapatkan warna keabuan dari citra yang bertujuan untuk menghapus data berlebihan yang terdapat pada citra.

### C. EKSTRAKSI FITUR MENGGUNAKAN PCA

PCA digunakan untuk mengumpulkan data citra ucapan visual yang diproses melalui ekstraksi fitur. Diberikan  $M$  *dataset* citra dari basis data citra pengucapan fonem  $A_j = [A_1, A_2, \dots, A_M]$ , ( $j = 1, 2, \dots, M$ ). Setiap gambar diubah menjadi matriks dua dimensi dari  $(X_m \times X_n)$ . Lalu, matriks diubah menjadi citra vektor  $T$  dengan ukuran  $(U \times 1)$ , dengan  $U = (X_m \times X_n)$  yang menghasilkan satu set vektor citra dengan ukuran  $(U \times M)$  [19].

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

Lalu, hasil rata-rata secara aritmetika dapat dihitung dari vektor citra yang berukuran  $(U \times 1)$  piksel menggunakan (2).

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

dengan  $M$  adalah jumlah data dan  $\sum_{j=1}^M T_j$  merupakan jumlah setiap baris.

Langkah selanjutnya adalah menghitung  $A_{Train}$  matriks, yang merupakan hasil selisih dari setiap nilai matriks  $T$  dengan matriks  $\bar{A}$ , menggunakan (3):

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

Lalu, nilai  $S_T$  *covariant matrix* dihitung menggunakan (4).

$$S_T = A_{Train} \times A_{Train}'. \quad (4)$$

Dari matriks  $S_T$  *covariant matrix*, nilai *eigen* ( $D$ ) dan vektor *eigen* ( $V$ ) dapat dihitung. Nilai *eigen* merupakan karakteristik

nilai dari matriks persegi, sedangkan vektor *eigen* merupakan nilai yang diambil dari nilai *eigen* yang lebih besar dari 0. Dalam penelitian ini, nilai *eigen* dan vektor *eigen* diperoleh menggunakan fungsi *eig()* yang diimplementasikan menggunakan MATLAB. Langkah berikutnya adalah menghitung nilai *eigenfaces*, yang merupakan karakteristik data citra, menggunakan (5).

$$Eigenfaces = A_{Train} \times V. \quad (5)$$

Fungsi PCA selanjutnya adalah mengurangi karakteristik yang masih disajikan dalam data citra dengan menghapus fitur data yang tidak penting. Persamaan (6) digunakan untuk menghitung matriks proyeksi PCA.

$$PCA\_Projected = Eigenfaces' \times T \quad (6)$$

dengan *PCA\_Projected* merupakan *dataset* yang akan digunakan dalam proses klasifikasi selanjutnya.

#### D. KLASIFIKASI

*Backpropagation neural network* merupakan salah satu model dari jaringan saraf tiruan yang menggunakan arsitektur *multilayer* untuk mencari bobot optimal dan meminimalkan nilai *error* pada hasil jaringan saraf tiruan. Klasifikasi pada *backpropagation neural network* bekerja dengan dua tahap, yaitu perhitungan maju dan perhitungan mundur [20]. Penggunaan *backpropagation neural network* ini bertujuan untuk mengetahui tingkat akurasi dari keberhasilan pengolahan data serta untuk meningkatkan kinerja penelitian sebelumnya [21]. Proses pengolahan data menggunakan *backpropagation neural network* dilakukan dengan memasukkan *dataset* yang akan digunakan sebagai sampel. Sampel *dataset* berjumlah 50 citra pada setiap fonem yang digolongkan menjadi data latih dan data uji dengan perbandingan data 6:4, 7:3, 8:2, dan 9:1. Tahap pertama sebelum melakukan ekstraksi adalah menggabungkan seluruh citra fonem dalam satu *folder* yang nantinya akan dimasukkan ke dalam *source code* pada MATLAB. Hasil ekstraksi citra berupa *dataset* yang kemudian disimpan dalam format Excel yang selanjutnya akan diolah menggunakan *backpropagation neural network* untuk menghasilkan akurasi yang diharapkan.

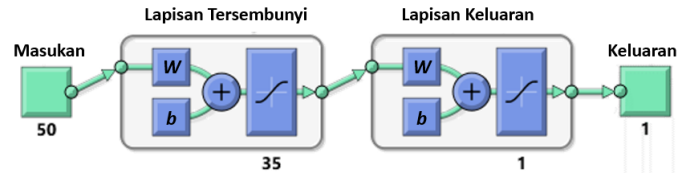
Pada penelitian ini, klasifikasi data dilakukan menggunakan *backpropagation neural network* yang banyak digunakan dalam kasus prediksi data. *Backpropagation neural network* terdiri atas tiga tahap pelatihan. Pada tahap pertama, yaitu tahap maju, lapisan masukan dan keluaran dihitung maju, yang ditentukan berdasarkan fungsi aktivasi [22]. Pada tahap kedua, yaitu tahap mundur, keadaan target keluaran yang diinginkan memiliki selisih dengan keluaran jaringan. Selisih (*error*) ini kemudian dipropagasikan mundur. Pada tahap ketiga, nilai bobot diubah untuk mengurangi kemungkinan kesalahan yang akan terjadi [21], [22].

Pada jaringan saraf terdapat lapisan keluaran yang mempunyai neuron linear, lapisan tersembunyi, dan lapisan masukan, yang masing-masing memiliki neuron nonlinear dan neuron linear, dengan  $[x_1(k) \ x_2(k) \ \dots \ x_n(k)]^T$  sebagai masukan pada vektor [23]. Keluaran dari neuron ke-*j* pada lapisan tersembunyi dihitung menggunakan (7).

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

dengan  $F_j(\cdot)$  adalah fungsi aktivasi *sigmoidal* nonlinear. Keluaran dari jaringan saraf dapat ditulis seperti pada (8).

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



Gambar 6. Arsitektur model optimal.

Kemudian, pendeteksi kesalahan dapat ditulis seperti pada (9).

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

dengan  $d(k)$  adalah sinyal referensi yang diperlukan untuk  $y(k)$ . Titik  $e(k) = 0$  ini adalah titik kesetimbangan dari sistem gangguan dinamis. Bobot jaringan saraf disesuaikan dengan persamaan berikut.

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

$$w_{j,i}^{1,0}(k) = \frac{1}{n x_i(k)} G_j \left( \frac{\beta^{-\frac{k}{2}} e(k-1) + d(k)}{m w_{1,j}^{2,1}(k)} \right) \quad (11)$$

dengan  $G_j = F_j^{-1}(\cdot)$  dan  $\beta$  adalah kecepatan pembelajaran jaringan. Persamaan (10) dan (11) menjamin konvergensi *asymptotic* dari keluaran pendeteksi kesalahan  $e(k)$  sebagai penyimpanan.

Pembobotan pada pengujian ini menggunakan fungsi aktivasi yang dipengaruhi oleh jumlah lapisan neuron, seperti ditunjukkan pada (12).

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

dengan

- $z\_in_j$  = nilai untuk menghitung lapisan tersembunyi
- $x_i$  = nilai lapisan masukan ke-*i*
- $v_{ij}$  = bobot antara lapisan masukan dengan lapisan tersembunyi.

#### E. OPTIMASI PEMBOBOTAN

Proses pengujian dilakukan dengan menggunakan citra dengan jumlah pada masing-masing fonem adalah 50 citra. Akan dilakukan analisis pada hasil akurasi untuk proses latih dengan optimasi pembobotan maupun tanpa optimasi pembobotan. Proses optimasi pembobotan dilakukan dengan menggunakan teknik pembagian proporsi data pada hasil ekstraksi fitur dengan PCA. Proses pencarian parameter terbaik dilakukan dengan melihat hasil nilai akurasi pada masing-masing proporsi data yang digunakan dan dipilih satu nilai akurasi yang paling tinggi. Pembagian proporsi data yang digunakan pada penelitian ini adalah sebagai berikut:

- 60% data latih dan 40% data uji,
- 70% data latih dan 30% data uji,
- 80% data latih dan 20% data uji, dan
- 90% data latih dan 10% data uji.

Pada percobaan ini dilakukan optimasi data dengan mencari model optimal dengan satu lapisan tersembunyi arsitektur jaringan saraf seperti pada Gambar 6. Model optimal ini menggunakan 50 masukan data dan 35 neuron pada lapisan tersembunyi.

### III. HASIL DAN PEMBAHASAN

Proses optimasi data klasifikasi fonem menggunakan data hasil ekstraksi dari konversi RGB ke *grayscale*, yang kemudian dibagi menjadi data latih dan data uji. Dari 1.280 data yang

TABEL I  
 HASIL PENGUJIAN DENGAN PERBANDINGAN PROPORSI DATA

Proporsi Data yang Digunakan	Akurasi yang Diperoleh (%)
6:4	81,42
7:3	77,17
8:2	83,77
9:1	65,60

TABEL II  
 HASIL PERBANDINGAN AKURASI

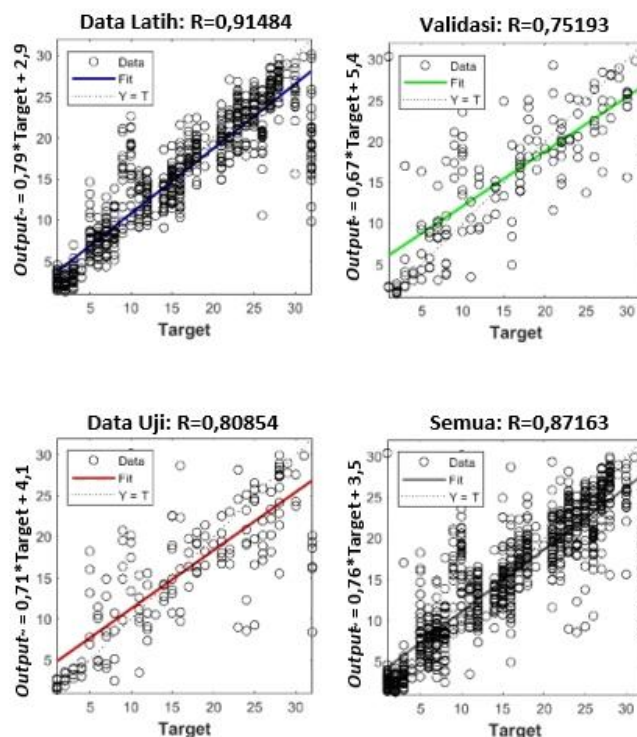
Hasil Penelitian Sebelumnya [21]		Sebelum Menggunakan Fungsi PCA		Setelah Menggunakan Fungsi PCA	
Lapisan Tersembunyi	Akur. %	Lapisan Tersembunyi	Akur. %	Lapisan Tersembunyi	Akur. %
4	20,59	10	76,33	25	85,62
5	11,76	15	83,77	35	87,16
20	52,94	25	76,57	40	84,42
37	32,35	35	71,88	50	82,33

Akur. = Akurasi

telah diekstraksi dihasilkan  $7.100 \times 1.280$  dataset yang selanjutnya akan diolah menggunakan *backpropagation neural network*. *Backpropagation neural network* berperan sebagai alat untuk klasifikasi data sehingga didapatkan besaran ketelitian data atau akurasi data. Penelitian ini menggunakan tiga tahap untuk mengoptimasi kinerja proses klasifikasi, yaitu perbandingan proporsi dataset, estimasi jumlah lapisan tersembunyi, dan reduksi dimensi dataset menggunakan metode PCA untuk mengurangi dimensi data dengan mempertahankan sebanyak mungkin informasi dari dataset asli [24].

Tahap pertama dilakukan untuk mendapatkan akurasi dari dataset fonem dengan melakukan *trial and error* pada proporsi data latih dan data uji yang berbeda-beda agar diperoleh hasil proporsi data bernilai 100%. Proporsi data yang digunakan pada penelitian ini terdiri atas empat proporsi, yaitu 6:4, 7:3, 8:2, dan 9:1. Tabel I menunjukkan hasil akurasi yang diperoleh dari masing-masing perbandingan data latih dan data uji. Hasil akurasi proporsi data dengan perbandingan 6:4 adalah 81,42%, akurasi proporsi data dengan perbandingan 7:3 sebesar 77,17%, akurasi proporsi data dengan perbandingan 8:2 sebesar 83,77%, dan hasil akurasi proporsi data dengan perbandingan 9:1 sebesar 65,60%. Tahap pertama mendapatkan hasil yang optimal pada jumlah proporsi data 8:2, sehingga proporsi data 8:2 dianggap paling optimal pada tingkat akurasi dataset. Namun, pada tahap pertama ini belum didapatkan tingkat akurasi yang aman karena masih di bawah 85%. Sebuah klasifikasi dapat diterima jika akurasinya lebih atau sama dengan 85%.

Tahap kedua untuk mencari optimasi akurasi dataset yaitu dengan mengestimasi jumlah lapisan tersembunyi pada proses klasifikasi menggunakan *backpropagation neural network*. Tabel II menyajikan perbandingan antara hasil akurasi yang diperoleh pada penelitian sebelumnya, hasil akurasi sebelum, dan setelah menggunakan fungsi PCA. Tahap ini menganalisis dengan cara mengamati jumlah lapisan tersembunyi yang berbeda-beda dengan membandingkan penelitian sebelumnya, proporsi data yang berbeda, dan menggunakan fungsi PCA. Berdasarkan hasil penelitian sebelumnya, ketika data pada metode *backpropagation* dibagi menjadi data latih dan data uji, nilai epoch dapat meningkat dan waktu menjadi lebih baik.



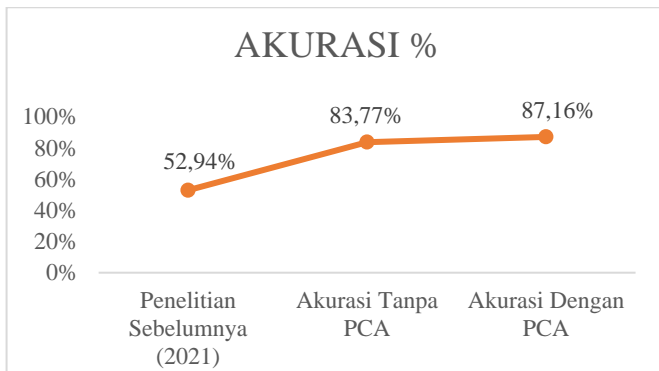
Gambar 7. Grafik hasil optimasi menggunakan PCA.

Akan tetapi, hal tersebut tidak selalu dapat meningkatkan hasil akurasi [21].

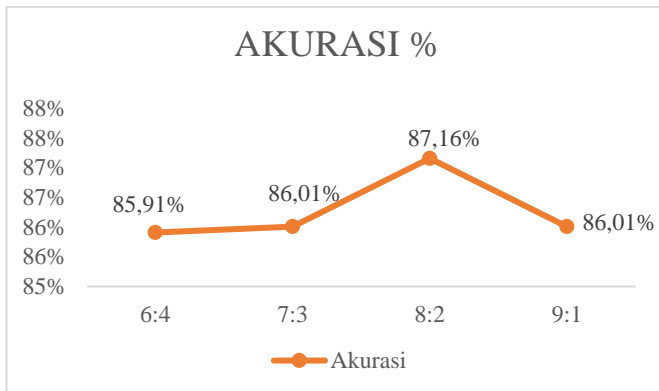
Berdasarkan Tabel II, dapat diketahui bahwa optimasi data tertinggi terjadi pada proses yang menggunakan fungsi PCA. Pada proses PCA, data hasil ekstraksi citra fonem disederhanakan dengan menghilangkan fitur data yang tidak penting, sehingga panjang matriks yang dihasilkan lebih sedikit dibandingkan matriks awal ekstraksi data. Proses optimasi dataset menggunakan pembobotan lapisan tersembunyi menghasilkan nilai akurasi yang lebih baik jika dibandingkan dengan pengujian tanpa pembobotan lapisan tersembunyi. Jumlah lapisan tersembunyi yang optimal dipengaruhi oleh dataset latih yang digunakan.

Dari pengolahan yang telah dilakukan untuk melihat pengaruh lapisan tersembunyi terhadap akurasi, dilakukan pengujian secara *trial and error* terhadap jumlah lapisan tersembunyi yang digunakan untuk mendapatkan hasil akurasi yang diharapkan. Hasil akurasi setelah menggunakan fungsi PCA ditunjukkan pada Gambar 7 dengan lapisan tersembunyi berjumlah 35 dan epoch data sebanyak 200.

Dalam hasil penelitian sebelumnya, proses peningkatan akurasi dengan membagi dataset menjadi data latih dan data uji serta mengubah lapisan tersembunyi yang digunakan masih menghasilkan nilai akurasi yang tergolong rendah, yaitu 52,94%. Pembobotan data latih menghasilkan nilai akurasi lebih tinggi sebesar 83,77% pada proporsi data latih dan data uji 8:2 dengan menggunakan 15 lapisan tersembunyi [21]. Kemudian, hasil tertinggi tersebut dioptimasi menggunakan fungsi PCA dan menghasilkan nilai akurasi 85,62% dengan 25 lapisan tersembunyi; 87,16% dengan 35 lapisan tersembunyi; 84,42% dengan 40 lapisan tersembunyi; dan 82,33% dengan 50 lapisan tersembunyi. Proses optimasi data tertinggi diperlihatkan pada Gambar 8, yang menunjukkan kenaikan akurasi dari hasil pembobotan data sampai proses menggunakan fungsi PCA.



Gambar 8. Grafik perbandingan akurasi data.



Gambar 9. Grafik hasil akurasi proporsi data PCA.

TABEL III  
HASIL PERBANDINGAN AKURASI

Proporsi Data	Lapisan Tersembunyi	Data Penting PCA	Akurasi
6 : 4	80	50	85,91%
7 : 3	40	30	86,01%
8 : 2	35	50	87,16%
9 : 1	60	50	86,01%

Setelah mendapatkan hasil akurasi yang meningkat dengan menggunakan fungsi PCA pada perbandingan proporsi data 8:2, dilakukan pengujian data kembali pada proporsi *dataset* 6:4, 7:3, dan 9:1. Pengujian dilakukan dengan mereduksi jumlah data secara *trial and error* pada data penting dan jumlah lapisan tersembunyi yang digunakan untuk mendapatkan nilai akurasi yang diharapkan. Hasil yang diperoleh diperlihatkan pada Tabel III. Terlihat bahwa nilai akurasi tertinggi masih diperoleh pada proporsi data 8:2 dengan lapisan tersembunyi 35 dan jumlah matriks data penting hasil reduksi PCA  $50 \times 1.280$ . Data penting PCA yang digunakan dalam Tabel III merupakan jumlah ukuran matriks data yang direduksi. Pada konfigurasi PCA yang digunakan dalam penelitian ini, terjadi pengulangan klasifikasi pada jumlah proporsi matriks ekstraksi fitur yang berbeda untuk mendapatkan nilai akurasi paling optimal. Proporsi data 6:4 dengan data hasil reduksi 50 dan lapisan tersembunyi 80 menghasilkan nilai akurasi sebesar 85,91%; proporsi data 7:3 dengan data hasil reduksi 30 dan lapisan tersembunyi 40 menghasilkan nilai akurasi sebesar 86,01%; dan proporsi data 9:1 dengan data hasil reduksi 50 dan lapisan tersembunyi 60 menghasilkan nilai akurasi sebesar 86,91%. Berdasarkan pengujian yang telah dilakukan pada seluruh pembagian proporsi data yang ada, dapat dilihat bahwa nilai akurasi optimal diperoleh ketika data hasil reduksi menggunakan fungsi PCA dan jumlah lapisan tersembunyi

yang digunakan memiliki perbandingan jumlah yang tidak terlalu signifikan. Gambar 9 menunjukkan grafik hasil akurasi menggunakan PCA dari beberapa proporsi data. Persentase akurasi tertinggi diperoleh sebesar 87,16% pada proporsi data 8:2.

#### IV. KESIMPULAN

Dari hasil pengujian dan analisis yang telah dilakukan, dapat disimpulkan bahwa klasifikasi fonem menggunakan fitur PCA dapat meningkatkan akurasi dari hasil pengujian. Hasil akurasi yang diperoleh menggunakan PCA sebesar 87,16%, sedangkan hasil pengujian *backpropagation neural network* sebesar 83,77% dan hasil penelitian sebelumnya sebesar 52,94%. Hal ini menunjukkan bahwa penggunaan fitur PCA dapat berfungsi dalam mengoptimalkan *dataset*.

#### KONFLIK KEPENTINGAN

Penulis menyatakan bahwa tidak terdapat konflik kepentingan.

#### KONTRIBUSI PENULIS

Konseptualisasi, metodologi, perangkat lunak, sumber daya, validasi, penulisan, Clara Maria Livia Suitela, Erika Dina Permata, Muzalfa Nakiatun Niza, dan Naeli Laelal Khiaroh; analisis formal, akurasi data, akuisisi pendanaan, Aripin.

#### REFERENSI

- [1] U. Kulsum, "Masalah Bunyi dalam Bahasa Masyarakat Indonesia," *Caraka*, Vol. 10, No. 1, hal. 299-302, 2021, doi: 10.31980/caraka.v10i1.1316.
- [2] (2022) "KBBI Online," [Online], <http://kbbi.co.id/cari?kata=fonem>, tanggal akses: 11-Jul-2022.
- [3] M. Khaleqyar dan M.Q. Kashaf, "The Importance of Suprasegmental Features in the Structure and Meaning of Pashto Sentence," *Spr. J. Arts. Humanit., Soc. Sci.*, Vol. 1, No. 9, hal. 31-40, Sep. 2022, doi: 10.55559/sjahss.v1i09.54.
- [4] F. Nur, D.E. Aryanto, dan N.I. Islami, "Analisis Segmentasi Silabel dan Fonem dalam Kalimat Perintah (Kajian Fonologi)," *Ghâncaran: J. Pendidik. Bhs. dan Sastra Indones.*, Vol. 3, No. 1, hal. 16-38, Jul. 2021, doi: 10.19105/ghancaran.v3i1.4424.
- [5] N. Ma'muriyah dan Jason, "Penerapan Metode Convolution Neural Network (CNN) Pada Aplikasi Automatic Lip Reading," *J. Inform., Telecommun. Eng.*, Vol. 6, No. 1, hal. 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.*, Vol. 9, No. 1, hal. 381-386, Jan. 2020, doi: 10.21275/ART20203995.
- [8] A. Roihan, P.A. Sunarya, dan A.S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper," *IJCIT (Indonesian J. Comput., Inf. Technol.)*, Vol. 5, No. 1, hal. 75-82, Mei 2020, doi: 10.31294/ijcit.v5i1.7951.
- [9] R.H. Abiyev dan M.K.S. Ma'aitah, "Deep Convolutional Neural Networks for Chest Diseases Detection," *J. Healthc. Eng.*, Vol. 2018, hal. 1-11, Agu. 2018, doi: 10.1155/2018/4168538.
- [10] D. Aprilla C., D.A. Baskoro, L. Ambarwati, dan I.W.S. Wicaksana, *Belajar Data Mining dengan RapidMiner*, R. Sanjaya, Ed., Jakarta, Indonesia: Gramedia Pustaka Utama, 2013.
- [11] O.D. Nurhayati, D. Eridani, dan A. Ulinuha, "Ekstraksi Ciri Orde Pertama dan Metode Principal Component Analysis untuk Mengidentifikasi Jenis Telur Ayam Kampung dan Ayam Arab," *JINIBIS (J. Sist. Inf. Bismis)*, Vol. 9, No. 2, hal. 133-140, Nov. 2019, doi: 10.21456/vol9iss2pp133-140.
- [12] T. Yulianti dkk., "Pengaruh Seleksi Fitur Citra terhadap Klasifikasi Tingkat Kesegaran Daging Sapi Lokal," *J. Tek. Pertan. Lampung*, Vol. 10, No. 1, hal. 85-95, Mar. 2021, doi: <http://dx.doi.org/10.23960/jtep-l.v10i1.85-95>.
- [13] A. Zamfiroiu, R. Boncea, I. Petre, "Data Acquisition Process for Smart Systems," *Int. J. Res., Anal. Rev. (IJRAR)*, Vol. 6, No. 2, hal. 553-556,

- Jun. 2019.
- [14] A. Riadi dan R. Sulaehani, "Analisis Implementasi Preprocessing dengan Otsu-Gaussian pada Pengenalan Wajah," *ILKOM J. Ilm.*, Vol. 11, No. 3, hal. 200–205, Des. 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, hal. 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, hal. 42–50, Okt. 2019, doi: 10.35447/jitekh.v7i02.80.
- [17] M.A. Nasution dkk., "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, hal. 37–48, Apr. 2022, doi: 10.22302/iopri.jur.jpks.v30i1.159.
- [18] H.A. Rabbani, M.A. Rahman, dan B. Rahayudi, "Perbandingan Ruang Warna RGB dan HSV dalam Klasifikasi Kematangan Biji Kopi," *J. Pengemb. Teknol. Inf., Ilmu Komput.*, Vol. 5, No. 6, hal. 2243–2248, Jun. 2021.
- [19] Arifin, Muljono, S. Sumpeno, dan M. Hariadi, "Towards Building Indonesian Viseme: A Clustering-Based Approach," *2013 IEEE Int. Conf. Comput. Intell., Cybern. (CYBERNETICSCOM)*, 2013, hal. 57–61, doi: 10.1109/CyberneticsCom.2013.6865781.
- [20] M.N.D. Sawitri, I.W. Sumarjaya, dan N.K.T. Tastrawati, "Peramalan Menggunakan Metode Backpropagation Neural Network," *E-Jurnal Mat.*, Vol. 7, No. 3, hal. 264–270, Sep. 2018, doi: 10.24843/mtk.2018.v07.i03.p213.
- [21] S.M. Damanik dkk., "Optimasi Data Menggunakan Teknik Backpropagation dalam Meningkatkan Hasil Nilai Akurasi," *Sem. Nas. Sains, Teknol. Inf. (SENSASI)*, 2021, hal. 657–662.
- [22] H. Okprana, M.R. Lubis, dan J.T. Hadinata, "Prediksi Kelulusan TOEFL Menggunakan Metode Resilient Backpropagation," *J. Edukasi, Pen. Inform.*, Vol. 6, No. 2, hal. 275–279, Agu. 2020, doi: 10.26418/jp.v6i2.41224.
- [23] M.S. Aftab dan 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, hal. 1–4, doi: 10.1109/INMIC.2018.8595664.
- [24] U. Nuha dan 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, hal. 138–143, Mar. 2020, doi: 10.26740/jinacs.v1n03.p138-143.