

EKTRAKSI FITUR ANGKA JAWA MENGGUNAKAN *DIAGONAL DISTANCE FEATURE* DAN *LONGEST RUN FEATURE*

Muhammad Misbahul Umam¹⁾ Susijanto Tri Rasmana²⁾ Madha Christianto Wibowo³⁾

Program Studi/Jurusan Sistem Komputer

Institut Bisnis dan Informatika Stikom Surabaya

Jl. Raya Kedung Baruk 98 Surabaya, 60298

Email : 1) missbbah@gmail.com, 2) susyanto@stikom.edu, 3) Madha@stikom.edu

Abstract: Java language must be always preserved so that Indonesia does not lose its cultural value. Write the number of Javanese in modern times can be done using a variety of media including computers, tablets, and Smartphones. write Java number using tools that generate an image file. As a result the number of Java could not be recognized as text by computer. One of the methods of pattern recognition, namely multi layer perceptron (MLP). In this study, the input of the MLP is 40 features in extraction using 2 methods i.e. diagonal distance feature and the longest run is featured. The Diagonal Distance feature extraction is an extraction method to calculate the distance between a right angle diagonal image until it finds the black pixels of the characters in the image. While the extraction of longest run featured is the extraction methods to calculate the amount of pixel longest black as well as in a row along the row, column, and diagonal of the two areas of the image. Sample training and sample testing each totaled 25 sets. Each set contains 10 pieces of handwritten patterns figures sample. The success of the introduction of the MLP MLP results weighting using learning with input features on each of the 40 samples of 66,8% in a sample of nearly 63% in training and sample testing. The success of the introduction of the MLP MLP results weighting using learning with a 36 input features on each of the extraction of samples from the longest run reached 100% in a sample of 94,4% in training and sample testing. The success of the introduction of the MLP MLP results weighting using learning with 4 input features on each of the samples from the extraction of the diagonal distance by 10% in a sample of training as well as sample testing.

Keywords : Javanese Character, Diagonal Distance, Method Of Extraction, Longest Run, MLP

Menulis aksara jawa di zaman modern ini dapat dilakukan dengan menggunakan berbagai macam media diantaranya komputer, tablet hingga *handphone*. Hasil tulisan aksara jawa menggunakan komputer, tablet dan *handphone* adalah berupa gambar. Akibatnya tulisan aksara jawa didalam gambar tersebut belum bisa terbaca sebagai teks yang dapat dikenali oleh komputer.

Agar tulisan yang ada di dalam gambar dapat dikenali sebagai teks oleh komputer, maka diperlukan sebuah aplikasi yang mampu mengenali tulisan pada gambar. Aplikasi dapat dibuat dengan menggunakan berbagai macam metode pengenalan pola. Salah satu metode pengenalan pola yang dapat digunakan yaitu *MLP (Multi Layer Perceptron)*.

Pada tahun 2013 telah dilakukan penelitian pengenalan pola pada Aksara Jawa menggunakan metode *multi layer perceptron (MLP)*. Aksara Jawa yang diteliti hanya aksara "Ha Na Ca Ra Ka". Inputan yang digunakan dalam pelatihan MLP adalah berupa piksel-piksel citra yang di ubah

rosulusnya menjadi lebih kecil saja. Data yang digunakan dalam pelatihan MLP adalah data mentah piksel-piksel citra yang hanya dikecilkan saja. presentase keberhasilan pengenalan mencapai 100% pada sampel pelatihan dan 56% pada sampel yang belum pernah dilatihkan (Wibowo dan Wirakusuma, 2013).

Pada tahun 2006 telah dilakukan penelitian mengenai pengenalan pola angka arab *Handwritten Arabic Numeral Recognition using a Multi Layer Perceptron*. Inputan yang digunakan dalam pelatihan MLP adalah *output* dari proses ekstraksi fitur. Hasil dari penelitian menunjukkan keberhasilan mencapai 95% tanpa dilakukan pelatihan sebelumnya (Nibarandkk, 2006).

Berdasarkan penelitian-penelitian telah dilakukan tersebut, pada Tugas Akhir ini dibuat proses ekstraksi fitur pada citra Angka Jawa hasil tulisan tangan sebelum dilakukan *training* pada MLP. Ekstraksi fitur angka jawa dapat dilakukan menggunakan metode *diagonal distance feature* dan *longest run feature*. *Diagonal Distance feature*

adalah metode yang berfungsi mengenali pola dengan menghitung jarak diagonal antara sudut siku gambar hingga di temukannya piksel hitam dari karakter gambar. Sedangkan *longest run feature* adalah metode untuk mengenali pola dengan menghitung jumlah piksel hitam yang terpanjang serta berturut-turut di sepanjang baris dan kolom, serta dua diagonal dari area kotak kecil. Hasil ekstraksi *diagonal distance* dan *longest run Feature* nantinya menjadi *input* pelatihan MLP.

Proses pelatihan akan dilakukan sebanyak satu kali. Setelah dilatihkan dan diperoleh bobot yang optimal maka aplikasi siap untuk melakukan pengenalan pola. Pola yang digunakan untuk pengenalan adalah 10 jenis Angka Jawa. Data yang digunakan berjumlah 25 set pola tulisan tangan Angka Jawa sebagai *training*, dan 25 set pola tulisan tangan Angka Jawa yang berbeda untuk *testing*.

Hasil pengujian dari setiap sampel akan dikomparasikan untuk mengetahui kehandalan sistem yang menggunakan ekstraksi *diagonal distance feature* dan *longest run feature* serta MLP untuk *classifier*-nya

ANGKA JAWA

Carakan (Abjad/Aksara Jawa) adalah huruf yang digunakan dalam ejaan bahasa Jawa yang pada dasarnya terdiri atas dua puluh aksara pokok (*nglegena*), yang ditunjukkan pada Gambar 1, yang bersifat silabik (*kesukukataan*). (Darusuprta, dkk., 2002). Bentuk kontemporer Aksara Jawa terbentuk sejak masa Kerajaan Mataram pada abad ke-17.

Carakan (Aksara Jawa) adalah huruf yang digunakan dalam ejaan bahasa Jawa. Aksara Jawa pada dasarnya terdiri atas dua puluh aksara pokok (*nglegena*). Aksara juga memiliki kelompok huruf vokal (*swara*), huruf kapital atau aksara kapital (*murda*), rekaan (*rekaan*), tanda baca, penutup konsonan (*pasangan*), angka dan lambang bilangan. Angka Jawa terdiri dari 9 bentuk pola.

| | | | | |
|----|-----|----|-----|-----|
| ꦲ | ꦤ | ꦕ | ꦫ | ꦏ |
| ha | na | ca | ra | ka |
| ꦢ | ꦠ | ꦱ | ꦮ | ꦭ |
| da | ta | sa | wa | la |
| ꦥ | ꦢꦲ | ꦗ | ꦪ | ꦚꦏ |
| pa | dha | ja | ya | nya |
| ꦩ | ꦒ | ꦧ | ꦠ | ꦚ |
| ma | ga | ba | tha | nga |

Gambar 1. Angka Jawa

DIAGONAL DISTANCE FEATURE

Diagonal distance feature adalah metode ekstraksi fitur yang digunakan untuk mengenali suatu pola tulisan, angka, atau simbol pada citra dengan cara menghitung jarak sudut diagonal hingga bertemunya piksel hitam pada citra (Nibaran das dkk, 2006). Citra yang dapat digunakan dalam metode ini hanya citra biner, Oleh karena perlu dilakukan konversi citra menjadi biner apabila tidak sesuai. , Seperti diilustrasikan pada gambar 2.



Gambar 2. Diagonal Distance Feature

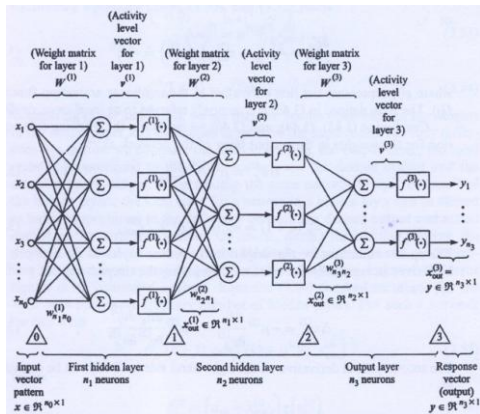
Empat nilai jarak dari sisi-sisi diagonal itulah yang akan menjadi nilai inputan untuk jaringan MLP.

LONGEST RUN FEATURE

Longest run feature yaitu Suatu metode ekstraksi fitur yang digunakan untuk pengenalan angka, tulisan, polabahkan simbol pada sebuah gambar/citra dengan teknik menghitung jumlah piksel berwarna hitam yang saling berurutan pada tiap sub-sub citra. Jumlah piksel hitam yang terpanjang di seluruh baris dan kolom, serta dua diagonal dari area sub citra (Nibaran das dkk, 2006). Citra yang dapat digunakan dalam metode ini hanya citra biner, Oleh karena perlu dilakukan konversi citra menjadi biner apabila tidak sesuai.

MULTI LAYER PERCEPTRONE (MLP)

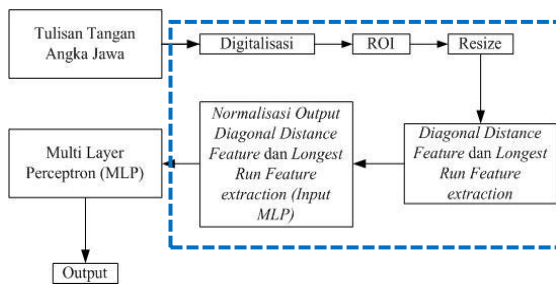
Multi layer perceptron (MLP) merupakan inovasi dari algoritma *perceptron* yang mempunyai *layer* tidak hanya sekedar *input layer* dan *output layer*. Pada MLP, terdapat bobot-bobot yang menghubungkan sejumlah *neuron*. Seluruh *neuron* tersusun dalam sebuah *layer input*, sebuah *layer output* dan satu hingga lebih dari satu lapisan tersembunyi (*hidden layer*). *Input MLP* menerima sinyal dari luar lalu dilewatkan pada lapisan *hidden layer* pertama, kemudian diteruskan hingga mencapai lapisan *output*. Dengan menggunakan *hidden layer* ini MLP memungkinkan untuk melakukan pengenalan pola yang kompleks yang tidak mampu dikenali oleh struktur jaringan *perceptron* biasa. Bentuk struktur jaringan MLP terdapat pada Gambar 3.



Gambar 3. Struktur Jaringan MLP (Ham, Kostanic, 2001)

METODE

Pengerjaan Tugas Akhir ini ditunjukkan dalam bentuk blok diagram pada gambar 4. Pada blok diagram ini menggambarkan proses dari sampel citra hingga output berupa pengenalan angka jawa.



Gambar 4. Blok Diagram Sistem

Pengerjaan tugas akhir ini hanya fokus pada bagian dalam dari area garis biru putus-putus diatas. Sampel Citra tulisan tangan angka jawa diambil melalui *canvas virtual* pada aplikasi. Setiap *participant* akan menggambar 1 digit angka jawa satu demi satu dari angka 0 sampai 9. Data citra yang dihasilkan dapat memiliki resolusi serta motif yang bervariasi.

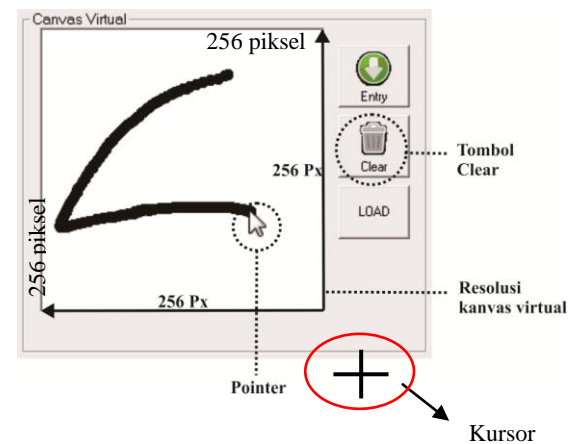
1. Tulisan tangan Angka Jawa diperoleh dari participant dengan menggunakan Tablet .
2. Digitalisasi merupakan proses membaca setiap piksel citra pada gambar, jika lebih condong ke hitam akan bernilai "0", dan jika lebih condong ke putih akan bernilai "255".
3. *ROI* adalah proses yang dilakukan untuk membuang bagian yang tidak diperlukan dari citra, agar lebih fokus pada area pola angka jawa yang akan di ekstraksi.
4. *Resize* adalah proses mengubah ukuran citra yang telah di *ROI* menjadi berukuran 64x64 piksel.
5. *Diagonal distance feature extraction* merupakan metode ekstraksi fitur dengan cara menghitung jarak pola dengan sudut siku citra.

6. *Longest run feature extraction* merupakan metode ekstraksi fitur dengan cara menghitung jumlah piksel hitam terpanjang serta berurutan pada baris, kolom serta dua diagonal dari sub area citra.
7. Normalisasi *output* berfungsi untuk menghindari terjadinya *overflows* saat proses *training* MLP berlangsung.

Perancangan Perangkat Lunak Desain Kanvas Virtual

Kanvas virtual digunakan sebagai media masukan oleh *participant* untuk menggambar pola Angka Jawa. *Device* yang bisa digunakan diantaranya mouse ,tablet pen dan *smartphone*. Kanvas virtual akan menangkap pola tulisan tangan yang telah digambarkan. Kanvas virtual ini dibuat menggunakan *tool PictureBox* yang disediakan pada Visual basic 6.0 dengan resolusi 256x256 piksel.

Untuk membuat pola angka jawa, pertama-tama *participant* harus mengetahui bentuk dari angka jawa yang akan dibuat. Selanjutnya *participant* dapat memulai menggambar pola angka jawa dengan cara *drag* mouse pada area kanvas.



Gambar 5. Kanvas Virtual

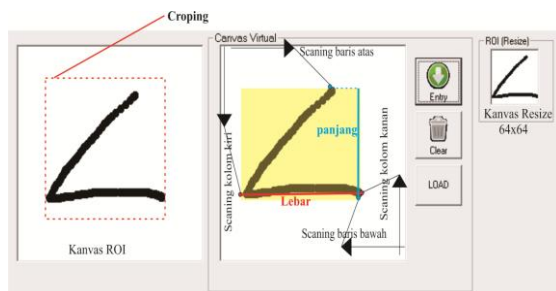
Gambar pola angka jawa *participant* akan disimpan pada menggunakan format *.*bmp(bitmap)*. format bitmap umum digunakan pada sistem operasi windows dan juga dapat dibuka oleh hampir semua *software* pengolah gambar.

Digitalisasi

Digitalisasi adalah proses pembacaan dan pengkodean warna per pixel pada kanvas. Piksel berwarna putih memiliki nilai 255 dan piksel berwarna hitam memiliki nilai 0. Gambar pola Angka Jawa akan di *scanning* per baris untuk membaca warna tiap piksel untuk di klasifikasikan. Saat *scanning* berlangsung kordinat piksel akan disimpan dalam array 2 dimensi dengan format kolom , baris (x,y).

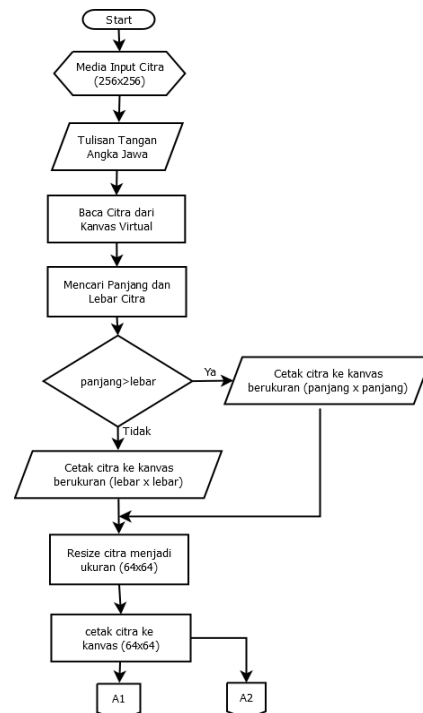
ROI (Region Of Interest) dan Resize

Proses ROI digunakan untuk membuang bagian dari gambar yang tidak diperlukan untuk diolah pada proses ekstraksi, agar lebih fokus pada area pola Angka Jawa yang akan di ekstraksi. Karena hasil tulisan tangan Angka Jawa dari *participant* memiliki pola dan resolusi yang bervariasi. Pertama-tama dilakukan *scanning* secara perbaris dimulai dr kordinat $(x,y)=(0,0)$ atas hingga bertemu piksel hitam, jika sudah bertemu piksel hitam kemudian break dan simpan kordinat untuk menandai batas atas(ba). selanjutnya dilakukan *scanning* secarabaris juga namun dimulai dr kordinat $(x,y)=(255,255)$ atas hingga bertemu piksel hitam, jika sudah bertemu piksel hitam kemudian break dan simpan kordinat untuk menandai batas bawah(bb). Langkah yang sama di ulangi namun *scanning* dijalankan dari secara per kolom. Nilai panjang sama dengan batas bawah – batas atas. Nilai lebar sama dengan batas kanan – batas kiri. Setelah itu dibandingkan nilai yang lebih panjang antara panjang dan lebar, yang nilainya lebih panjang dikurangkan yang terpendek. Selisih dari pengurangan kemudian dibagi dua. Setelah itu nilainya ditambahkan untuk kedua sisi yang terpendek untuk menghasilkan perpotongan yang *square* (persegi). Dalam Proses *cropping*, citra harus di crop secara persegi dan posisi citra berada tengah-tengah persegi (*center*). Setelah *cropping* selesai maka selanjutnya citra di resize menjadi resolusi 64x64 piksel. Penentuan ukuran hasil resize akan berpengaruh pada hasil fitur yang akan didapatkan. Jika citra di *resize* dengan ukuran terlalu kecil maka akan menyebabkan perubahan bentuk karakter yang sangat jauh perbedaannya dari bentuk karakter yang semestinya. Hal ini juga dapat menyebabkan kemiripan fitur antara satu Angka Jawa dengan Angka Jawa yang lainnya menjadi lebih banyak. Namun jika *resize* terlalu besar maka akan menyebabkan proses ekstraksi akan berjalan semakin lama karena terlalu banyak piksel yang harus di *scanning*. Pemilihan hasil *resize* citra menjadi berukuran 64x64 piksel ditentukan berdasarkan penelitian yang pernah dilakukan (Nibaran das dkk, 2006).



Gambar 6. Citra Yang Akan Melewati Proses ROI dan Scalling / Resize

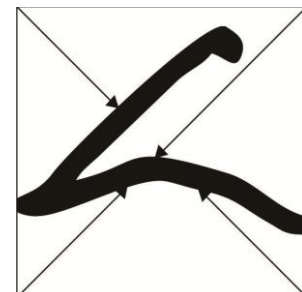
Proses ROI didalam program ditampilkan dalam Gambar 7.



Gambar 7. Flowchart proses ROI dan Resize

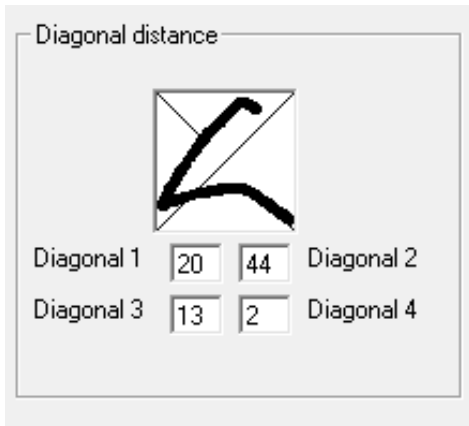
Ekstraksi Diagonal Distance Feature

Diagonal Distance Feature adalah metode untuk mengekstraksi fitur dengan cara menghitung jarak diagonal dari sudut siku citra hingga bertemunya piksel hitam. Citra yang sudah melalui proses ROI dan *resize* akan menghasilkan gambar beresolusi 64x64 piksel. Dari gambar ini akan dilakukan perhitungan jarak di mulai dari sudut siku bujur sangkar citra terus berjalan secara diagonal hingga bertemu piksel hitam. Pada ekstraksi ini satuan jaraknya adalah piksel. Citra yang di ekstrak akan menghasilkan total fitur *Diagonal Distance* berjumlah 4 fitur seperti gambar 9 dibawah ini.



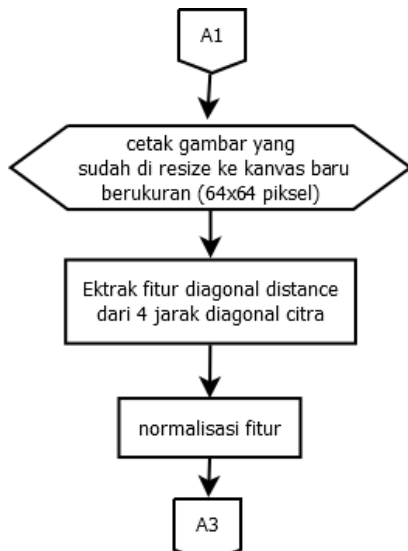
Gambar 8. Ekstraksi Diagonal Distance Feature

Keempat nilai diagonal itulah yang akan digunakan sebagai bagi jaringan MLP, sehingga dapat diklasifikasikan lebih lanjut.



Gambar 9. Tampilan Pada Aplikasi Ekstraksi Diagonal Distance Feature

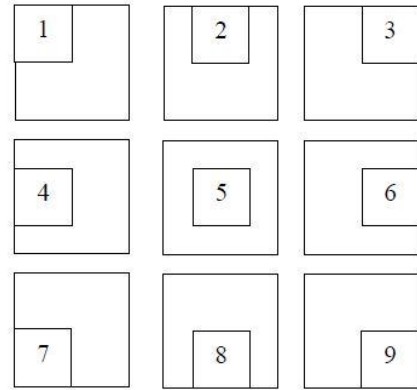
Langkah-langkah ekstraksi dijelaskan dengan *flowchart* pada Gambar 10.



Gambar 10. Flowchart *Diagonal Distance Feature Extraction*

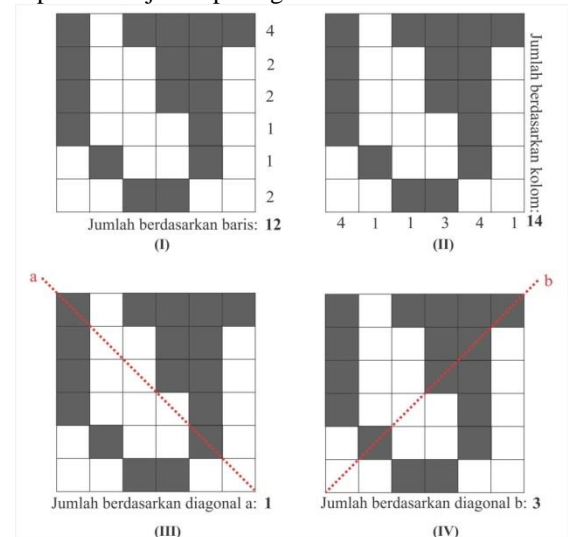
Ekstraksi *Longest Run Feature*

Citra tulisan tangan angka jawa dari *participant*, perlu dilakukan ROI (Region Of Interest), kemudian baru dilakukan *Resize* ukuran citra. Pada ekstraksi *Longest Run*, skala citra sebesar 64 x 64 akan dibagi menjadi 9 area kotak kecil yang saling tumpang tindih, seperti ditunjukkan pada gambar 11.



Gambar 11. Area kotak kecil

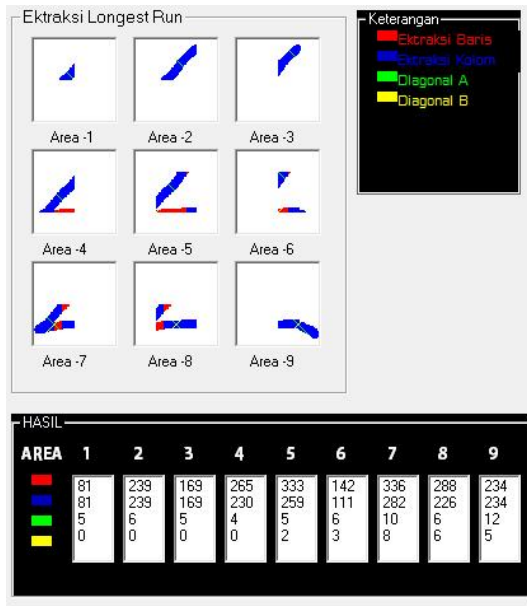
Longest run feature extraction adalah pengenalan pola dari suatu citra dengan cara menghitung jumlah piksel hitam yang terpanjang serta berturut-turut di sepanjang seluruh baris dan kolom, serta dua diagonal dari area kotak kecil, seperti ditunjukkan pada gambar 12.



Gambar 12. Ilustrasi Ekstraksi *Longest Run*

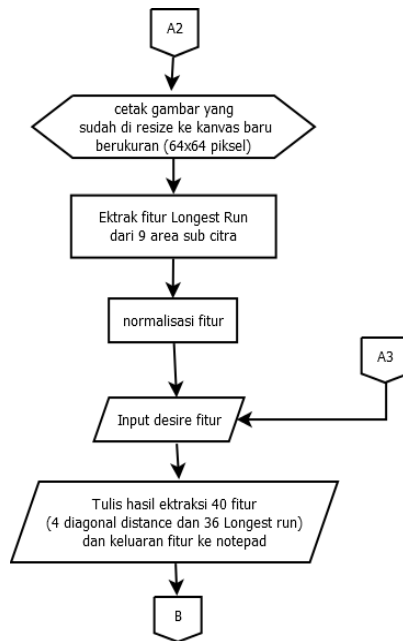
- (I). *Longest run* fitur dari baris
- (II). *Longest run* fitur dari kolom
- (III). *Longest run* fitur dari diagonal a
- (IV). *Longest run* fitur dari diagonal b

Dari tiap area kotak kecil akan menghasilkan 4 fitur. Maka dari seluruhnya akan didapatkan $4 \times 9 = 36$ *longest run feature* dari karakter gambar yang diinputkan.



Gambar 13. Tampilan Pada Aplikasi Ekstraksi Longest Run Feature

Langkah-langkah ekstraksi dijelaskan dengan flowchart pada gambar 14.

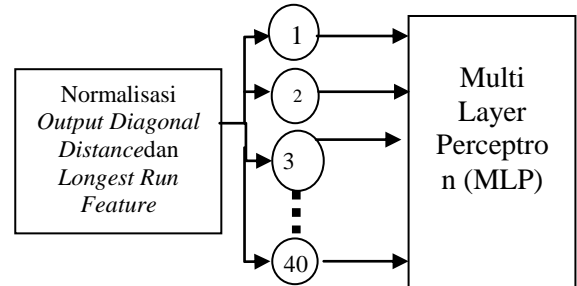


Gambar 14. Flowchart Longest Run Feature Extraction

Input Multi Layer Perceptron (MLP)

Skala citra sebesar 64 piksel x 64 piksel, jika citra dibagi menjadi 4 bagian diagonal, maka panjang diagonal maksimal masing-masing diagonal adalah 32 piksel, maka Nilai fitur *diagonal distance* dari tiap diagonal kemudian di normalisasi dengan cara dibagi dengan 32. Tiap area kotak kecil pada *longest run feature* akan

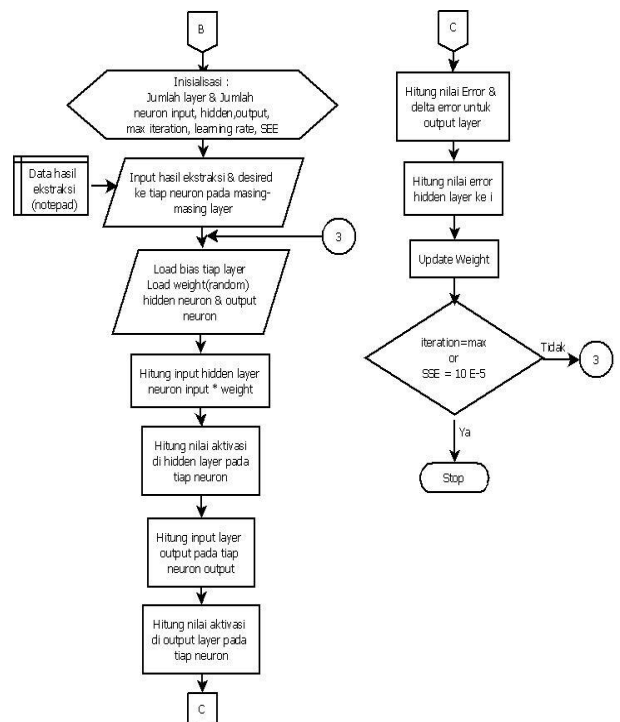
menghasilkan nilai maksimal baris dan kolom sebesar 1024 piksel saat semua piksel berwarna hitam, maka nilai yang didapatkan akan dinormalisasi dengan cara dibagi dengan 1024. Maka nilai dinormalisasi dengan cara yang sama seperti hasil perhitungan pada baris dan kolom ekstraksi *Longest Run Feature* yaitu dibagi dengan 1024.



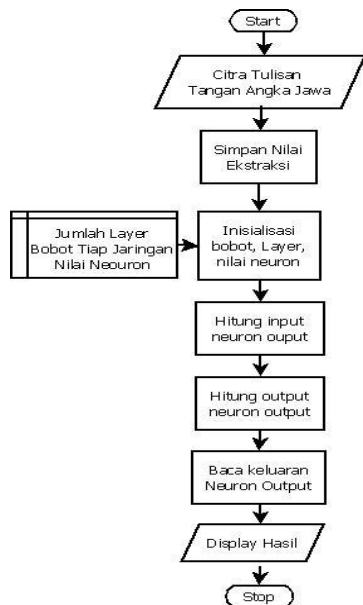
Gambar 15. Output Diagonal Distance Features dan Longest Run Feature Menjadi Input MLP

Total fitur yang akan di hasilkan berjumlah 40 fitur, yaitu 4 fitur dari ekstraksi *diagonal distance feature* dan 36 fitur dari ekstraksi *longest run feature*., begitu juga dengan *neuron input* dari MLP seperti pada gambar 16. Nilai normalisasi tersebut dimasukkan ke MLP untuk dilakukan pembelajaran dan pengenalan angka jawa.

Langkah meng-inputkan nilai output ekstraksi dan pengujian MLP serta pengolahan data ekstraksi diperlihatkan dalam flowchart pada gambar 17.



Gambar 16. Flowchart Training MLP



Gambar 17. Flowchart Testing MLP

Teknik Pengumpulan & Analisis Data Pengumpulan Data

Data sampel primer setiap 1 set (10 pola Angka Jawa) diperoleh dari mengumpulkan pola hasil tulisan tangan Angka Jawa dari *participant* yang berbeda. Diperlukan 25 *participant* untuk menghasilkan 25 set pola Angka Jawa untuk *training MLP* dan 25 *participant* untuk menghasilkan 25 set pola Angka Jawa. Setiap set terdiri dari Angka Jawa 0 sampai 9.

Pengumpulan data sekunder didapatkan dari percobaan yang dilakukan di laboratorium setelah project tugas akhir selesai dibuat, seperti yang dijelaskan pada sub bab sebelumnya. Setelah proses *training MLP* selesai dan didapatkan bobot yang optimal, kemudian MLP dengan menggunakan bobot optimal tersebut akan diuji untuk mengenaliberbagai macam pola Angka Jawa, baik dari sampel *training* maupun pada sampel *testing*.

Pelatihan MLP

Proses *learning* ini menghasilkan 2 buah file log untuk bobot yang dihasilkan. Bobot hasil akhir *learning* tercapai jika error minimum atau iterasi maksimum telah tercapai, selanjutnya disimpan pada file *weight.dat*. Selain itu, dihasilkan juga file bernama *weight2.dat* berfungsi untuk menyimpan bobot per 100 iterasi. Hal ini bermanfaat untuk menghindari resiko proses *learning* berhenti sebelum tercapai akhir proses *learning*. Resiko tersebut mungkin terjadidikarena oleh komputer *hang* atau putusnya aliran listrik dan sebagainya.

Proses *learning* dilakukan dengan konfigurasi sebagai berikut:

A. Konfigurasi *learning MLP* dengan input 40 Fitur

40 fitur adalah nilai-nilai hasil ekstraksi menggunakan *longest run feature* dan *diagonal distance*. 36 fitur dari ekstraksi menggunakan *longest run* dan 4 fitur dari ekstraksi *diagonal distance*.

- Jumlah hidden layer : 2
- Jumlah neuron *input* : 40
- Jumlah neuron layer 1 : 20
- Jumlah neuron layer 2 : 40
- Jumlah neuron output : 10
- Laju *learning* (μ) : 0.1
- Error minimum : 10E-4
- Iterasi maksimum : 10E+5

Proses *learning* dilakukan sebanyak 1000000 iterasi dengan menggunakan 25 *set* sampel *input* dengan nilai error akhir sebesar 3.06665790754129E-04. Ini berarti proses *learning* selesai hingga iterasi maksimal terpenuhi. Bobot yang dihasilkan akan disimpan pada file *weight.dat*.

B. Konfigurasi *learning MLP* dengan input 36 Fitur

36 fitur adalah nilai-nilai hasil ekstraksi menggunakan *longest run feature*. Sehingga hanya memerlukan 36 neuron untuk layer *input*.

- Jumlah hidden layer : 2
- Jumlah neuron *input* : 36
- Jumlah neuron layer 1 : 20
- Jumlah neuron layer 2 : 40
- Jumlah neuron output : 10
- Laju *learning* (μ) : 0.1
- Error minimum : 10E-4
- Iterasi maksimum : 10E+5

Proses *learning* dilakukan sebanyak 1000000 iterasi dengan menggunakan 25 *set* sampel *input* dengan nilai error akhir sebesar 1.22653997160217E-04. Ini berarti proses *learning* selesai hingga iterasi maksimal terpenuhi. Bobot yang dihasilkan akan disimpan pada file *weight.dat*.

C. Konfigurasi *learning MLP* dengan input 4 Fitur

4 fitur adalah nilai-nilai hasil ekstraksi menggunakan *diagonal distance*. Sehingga hanya memerlukan 4 neuron untuk layer *input*.

- Jumlah hidden layer : 2
- Jumlah neuron *input* : 4
- Jumlah neuron layer 1 : 20
- Jumlah neuron layer 2 : 40
- Jumlah neuron output : 10
- Laju *learning* (μ) : 0.1
- Error minimum : 10E-4
- Iterasi maksimum : 10E+5

Proses *learning* dilakukan sebanyak 1000000 iterasi dengan menggunakan 25 *set* sampel *input* dengan nilai error akhir sebesar 27.0015495662298. Ini berarti proses *learning* selesai hingga iterasi maksimal terpenuhi. Bobot yang dihasilkan akan disimpan pada file *weight.dat*.

Analisis Data

Semua data pada proses ekstraksi dan *learning*, baik citra maupun hasil normalisasi ekstraksi kemudian dilakukan *training* terhadap MLP. Kemudian dilakukan pengujian pengenalan pada sampel yang telah *ditrainingkan* pada aplikasi.

Setelah selesai melakukan pengujian pada sampel *training*, kemudian dilakukan pula pengujian pada sampel *testing*. Sampel *testing* adalah sampel yang belum pernah *ditrainingkan* pada MLP. Terdapat 3 parameter yang diperhatikan pada pengujian pengenalan MLP, yaitu pengenalan benar, atau salah (dikenali sebagai Angka Jawa lain).

HASIL DAN PENGUJIAN

Pengujian ini memiliki beberapa tahap untuk mendapatkan presentase keberhasilan pengenalan, dimana pengujian MLP dikategorikan berdasarkan jumlah fitur hasil ekstraksi yang digunakan sebagai *input* MLP. Pengujian ini bertujuan untuk mengetahui seberapa banyak pola angka yang dapat dikenali MLP dari sampel *training* maupun sampel *testing*. Sampel *training* dan sampel *testing* masing-masing berjumlah 25 set pola Angka Jawa. Setiap set berisi Citra Angka Jawa 0 sampai 9.

Pengujian Pengenalan MLP Menggunakan Input 4 Fitur Diagonal Distance

Pada pengujian ini, *input* MLP hanya menggunakan 4 fitur yang dihasilkan oleh ekstraksi *diagonal distance*.

Hasil Pengenalan MLP Menggunakan Input 4 Fitur Pada Sampel Training

Tabel 1. Keberhasilan pengenalan MLP dengan *input* 4 fitur pada sampel *training*

| Angka | Jumlah Data (buah) | Data Yang Berhasil Dikenali (buah) | Kesalahan Pengenalan (buah) | Presentase Keberhasilan (%) |
|-------------------------|--------------------|------------------------------------|-----------------------------|-----------------------------|
| 0 | 25 | 25 | 0 | 100% |
| 1 | 25 | 0 | 25 | 0% |
| 2 | 25 | 0 | 25 | 0% |
| 3 | 25 | 0 | 25 | 0% |
| 4 | 25 | 0 | 25 | 0% |
| 5 | 25 | 0 | 25 | 0% |
| 6 | 25 | 0 | 25 | 0% |
| 7 | 25 | 0 | 25 | 0% |
| 8 | 25 | 0 | 25 | 0% |
| 9 | 25 | 0 | 25 | 0% |
| Jumlah Persentase Total | | | | 10 % |

Tabel 1 menunjukkan tingkat keberhasilan pengenalan MLP dengan *input* fitur *diagonal distance* mencapai 10%.

Tabel 2. Kesalahan pengenalan MLP dengan *input* 4 fitur pada sampel *training*

| Angka | Salah Dikenali Sebagai Angka | | | | | | | | | Presentase Error | |
|----------------------------|------------------------------|---|---|---|---|---|---|---|---|------------------|------|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | 9 |
| 0 | - | - | - | - | - | - | - | - | - | - | 0% |
| 1 | 23 | - | - | - | 2 | - | - | - | - | - | 100% |
| 2 | 25 | - | - | - | - | - | - | - | - | - | 100% |
| 3 | 23 | - | - | - | 2 | - | - | - | - | - | 100% |
| 4 | 25 | - | - | - | - | - | - | - | - | - | 100% |
| 5 | 25 | - | - | - | - | - | - | - | - | - | 100% |
| 6 | 25 | - | - | - | - | - | - | - | - | - | 100% |
| 7 | 25 | - | - | - | - | - | - | - | - | - | 100% |
| 8 | 25 | - | - | - | - | - | - | - | - | - | 100% |
| 9 | 25 | - | - | - | - | - | - | - | - | - | 100% |
| Rata-Rata Persentase Error | | | | | | | | | | | 90% |

Tabel 2 menunjukkan selain angka "0" semuanya salah dikenali sebagai angka lain sebesar 100%. Angka "1" dan "3" salah dikenali sebagai angka "0" sebanyak 23 kali dan salah dikenali sebagai angka "4" sebanyak 2 kali.

Hasil Pengenalan Menggunakan *Input 4* Fitur MLP Pada Sampel *Testing*

Tabel 3. Keberhasilan pengenalan MLP dengan *input 4* fitur pada sampel *testing*

| Angka | Jumlah Data (buah) | Data Yang Berhasil Dikenali (buah) | Kesalahan Pengenalan (buah) | Presentase Keberhasilan (%) |
|--------------------------------|--------------------|------------------------------------|-----------------------------|-----------------------------|
| 0 | 25 | 25 | 0 | 100% |
| 1 | 25 | 0 | 25 | 0% |
| 2 | 25 | 0 | 25 | 0% |
| 3 | 25 | 0 | 25 | 0% |
| 4 | 25 | 0 | 25 | 0% |
| 5 | 25 | 0 | 25 | 0% |
| 6 | 25 | 0 | 25 | 0% |
| 7 | 25 | 0 | 25 | 0% |
| 8 | 25 | 0 | 25 | 0% |
| 9 | 25 | 0 | 25 | 0% |
| Jumlah Persentase Total | | | | 10 % |

Sama seperti pada sampel *training* pada table 3 menunjukkan keberhasilan pengenalan pada sampel *testing* hanya mencapai 10%. Untuk mengetahui kesalan-kesalahan pengenalan pada tiap angka ditunjukkan pada tabel 4.

Tabel 4. Kesalahan Pengenalan MLP Dengan *Input 4* Fitur Pada Sampel *Training*

| Angka | Salah Dikenali Sebagai Angka | | | | | | | | | Presentase Error | |
|-----------------------------------|------------------------------|---|---|---|---|---|---|---|---|------------------|------|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | 9 |
| 0 | - | - | - | - | - | - | - | - | - | - | 0% |
| 1 | 22 | - | - | 2 | 1 | - | - | - | - | - | 100% |
| 2 | 24 | - | - | - | 1 | - | - | - | - | - | 100% |
| 3 | 23 | - | - | - | 2 | - | - | - | - | - | 100% |
| 4 | 25 | - | - | - | - | - | - | - | - | - | 100% |
| 5 | 24 | - | - | - | 1 | - | - | - | - | - | 100% |
| 6 | 25 | - | - | - | - | - | - | - | - | - | 100% |
| 7 | 23 | - | - | - | 2 | - | - | - | - | - | 100% |
| 8 | 25 | - | - | - | - | - | - | - | - | - | 100% |
| 9 | 25 | - | - | - | - | - | - | - | - | - | 100% |
| Rata-Rata Persentase Error | | | | | | | | | | | 90% |

Tabel 4 menunjukkan selain angka “0” semuanya salah dikenali sebagai angka lain sebesar 100% . kesalahan pengenalan paling bervariasi yaitu angka ”1” yang dikenali sebagai angka “0” sebanyak 22 kali, salah dikenali sebagai angka 3 sebanyak 2 kali dan salah dikenali sebagai angka “4” sebanyak 1 kali.

Pengujian Pengenalan MLP Menggunakan *Input 36* Fitur *Longest Run*

Pada pengujian ini, *input* MLP hanya menggunakan 36 fitur yang dihasilkan oleh ekstraksi *longest run*.

Hasil Pengenalan MLP Menggunakan *Input 36* Fitur Pada Sampel *Training*

Tabel 5. Keberhasilan pengenalan MLP dengan *input 36* fitur pada sampel *training*

| Angka | Jumlah Data (buah) | Data Yang Berhasil Dikenali (buah) | Kesalahan Pengenalan (buah) | Presentase Keberhasilan (%) |
|--------------------------------|--------------------|------------------------------------|-----------------------------|-----------------------------|
| 0 | 25 | 25 | 0 | 100% |
| 1 | 25 | 25 | 0 | 100% |
| 2 | 25 | 25 | 0 | 100% |
| 3 | 25 | 25 | 0 | 100% |
| 4 | 25 | 25 | 0 | 100% |
| 5 | 25 | 25 | 0 | 100% |
| 6 | 25 | 25 | 0 | 100% |
| 7 | 25 | 25 | 0 | 100% |
| 8 | 25 | 25 | 0 | 100% |
| 9 | 25 | 25 | 0 | 100% |
| Jumlah Persentase Total | | | | 100% |

Tabel 5 menunjukkan tingkat keberhasilan pengenalan mencapai 100% pada seluruh angka. tidak ada kesalahan satupun pada pengenalan Angka Jawa pada sampel *training*.

Hasil Pengenalan Menggunakan *Input 36* Fitur MLP Pada Sampel *Testing*

Tabel 6. Keberhasilan pengenalan MLP dengan *input 36* fitur pada sampel *testing*

| Angka | Jumlah Data (buah) | Data Yang Berhasil Dikenali (buah) | Kesalahan Pengenalan (buah) | Presentase Keberhasilan (%) |
|--------------------------------|--------------------|------------------------------------|-----------------------------|-----------------------------|
| 0 | 25 | 25 | 0 | 100% |
| 1 | 25 | 25 | 0 | 100% |
| 2 | 25 | 22 | 3 | 88% |
| 3 | 25 | 21 | 4 | 84% |
| 4 | 25 | 25 | 0 | 100% |
| 5 | 25 | 24 | 1 | 96% |
| 6 | 25 | 24 | 1 | 96% |
| 7 | 25 | 23 | 2 | 92% |
| 8 | 25 | 23 | 2 | 92% |
| 9 | 25 | 24 | 1 | 96% |
| Jumlah Persentase Total | | | | 94,4% |

Pada table 6 keberhasilan pengenalan paling besar terjadi pada pola Angka Jawa “0”, ”1”, dan ”4” dengan tingkat keberhasilan pengenalan mencapai 100%.

Tabel 7. Kesalahan Pengenalan MLP Dengan Input 36 Fitur Pada Sampel Testing

| Angka | Salah Dikenali Sebagai Angka | | | | | | | | | | Presentase Error |
|----------------------------|------------------------------|---|---|---|---|---|---|---|---|---|------------------|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| 0 | - | - | - | - | - | - | - | - | - | - | 0% |
| 1 | - | - | - | - | - | - | - | - | - | - | 0% |
| 2 | 2 | | | 1 | - | - | - | - | - | - | 12% |
| 3 | 1 | 2 | 1 | | - | - | - | - | - | - | 16% |
| 4 | - | - | - | - | - | - | - | - | - | - | 0% |
| 5 | - | - | 1 | - | - | - | - | - | - | - | 4% |
| 6 | 1 | - | - | - | - | - | - | - | - | - | 4% |
| 7 | 1 | - | - | - | - | - | - | - | - | 1 | 8% |
| 8 | - | - | - | - | - | - | - | 1 | - | 1 | 8% |
| 9 | - | - | - | - | - | - | - | 1 | - | - | 4% |
| Rata-Rata Persentase Error | | | | | | | | | | | 5,6% |

Tabel 7 menunjukkan kesalahan pengenalan paling besar terjadi pada angka “3” yang salah dikenali sebagai angka lain sebanyak 4 kali. Kesalahan pengenalan terkecil terjadi pada angka “0” dan “1” karena tidak terjadi kesalahan pengenalan sama sekali.

Pengujian Pengenalan MLP Menggunakan Input 40 Fitur

Pada pengujian ini, input MLP menggunakan 40 fitur yang dihasilkan oleh 2 metode ekstraksi. 36 fitur berasal dari ekstraksi *longest run* dan 4 fitur dari ekstraksi *diagonal distance*.

Hasil Pengenalan MLP Menggunakan Input 40 Fitur Pada Sampel Training

Tabel 8. Keberhasilan pengenalan MLP dengan input 40 fitur pada sampel training

| Angka | Jumlah Data (buah) | Data Yang Berhasil Dikenali (buah) | Kesalahan Pengenalan (buah) | Presentase Keberhasilan (%) |
|-------------------------|--------------------|------------------------------------|-----------------------------|-----------------------------|
| 0 | 25 | 23 | 2 | 92 % |
| 1 | 25 | 1 | 24 | 4 % |
| 2 | 25 | 25 | 0 | 100 % |
| 3 | 25 | 25 | 0 | 100 % |
| 4 | 25 | 24 | 1 | 96 % |
| 5 | 25 | 21 | 4 | 84 % |
| 6 | 25 | 16 | 9 | 64 % |
| 7 | 25 | 22 | 3 | 88 % |
| 8 | 25 | 0 | 25 | 0 % |
| 9 | 25 | 10 | 15 | 40 % |
| Jumlah Persentase Total | | | | 66,8 % |

Tabel 8 menunjukkan tingkat keberhasilan pengenalan MLP dengan input 40 fitur mencapai 66,8%. Untuk mengetahui kesalahan pengenalan pada tiap angka ditunjukkan pada tabel 9.

Tabel 9. Kesalahan pengenalan MLP dengan input 40 fitur pada sampel training

| Angka | Salah Dikenali Sebagai Angka | | | | | | | | | | Presentase Error |
|----------------------------|------------------------------|---|---|---|---|---|---|----|---|---|------------------|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| 0 | - | - | - | - | 2 | - | - | - | - | - | 8 % |
| 1 | 8 | - | 1 | - | - | - | - | 15 | - | - | 96 % |
| 2 | - | - | - | - | - | - | - | - | - | - | 0 % |
| 3 | - | - | - | - | - | - | - | - | - | - | 0 % |
| 4 | - | - | 1 | - | - | - | - | - | - | - | 4 % |
| 5 | 3 | - | 1 | - | - | - | - | - | - | - | 16 % |
| 6 | - | - | 8 | 1 | - | - | - | - | - | - | 36 % |
| 7 | 2 | - | - | - | - | - | - | - | - | 1 | 12 % |
| 8 | 12 | - | - | - | - | - | - | 13 | - | - | 100 % |
| 9 | 4 | - | - | - | - | - | - | 11 | - | - | 60 % |
| Rata-Rata Persentase Error | | | | | | | | | | | 33,2% |

Tabel 9 menunjukkan kesalahan pengenalan terkecil terjadi pada angka “2” dan “3” yang semua sampelnya berhasil dikenali dengan benar. Kesalahan pengenalan terbanyak terjadi pada angka “8” yang sama sekali tidak berhasil dikenali dengan benar, melainkan salah dikenali sebagai angka “0” sebanyak 12 kali dan dikenali sebagai angka “7” sebanyak 13 kali.

Hasil Pengenalan MLP Menggunakan Input 40 Fitur Pada Sampel Testing

Tabel 10. Keberhasilan pengenalan MLP dengan input 40 fitur pada sampel training

| Angka | Jumlah Data (buah) | Data Yang Berhasil Dikenali (buah) | Kesalahan Pengenalan (buah) | Presentase Keberhasilan (%) |
|-------------------------|--------------------|------------------------------------|-----------------------------|-----------------------------|
| 0 | 25 | 21 | 4 | 84 % |
| 1 | 25 | 5 | 20 | 20 % |
| 2 | 25 | 22 | 3 | 88 % |
| 3 | 25 | 20 | 5 | 80 % |
| 4 | 25 | 24 | 1 | 96 % |
| 5 | 25 | 18 | 7 | 72 % |
| 6 | 25 | 18 | 7 | 72 % |
| 7 | 25 | 18 | 7 | 72 % |
| 8 | 25 | 0 | 25 | 0 % |
| 9 | 25 | 11 | 14 | 44 % |
| Jumlah Persentase Total | | | | 62,8 % |

Tabel 10 menunjukkan tingkat keberhasilan pengenalan MLP dengan input 40 fitur mencapai 62,8%. Untuk mengetahui kesalahan pengenalan pada tiap angka ditunjukkan pada tabel 11.

Tabel 11. Kesalahan pengenalan MLP dengan input 40 fitur pada sampel training

| Angka | Salah Dikenali Sebagai Angka | | | | | | | | | Presentase Error | |
|----------------------------|------------------------------|---|---|---|---|---|---|----|---|------------------|-------|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | 9 |
| 0 | - | - | - | - | 4 | - | - | - | - | - | 16 % |
| 1 | 9 | - | - | - | - | - | - | 11 | - | - | 80 % |
| 2 | 2 | - | - | 1 | - | - | - | - | - | - | 12 % |
| 3 | - | - | 1 | - | - | - | - | 4 | - | - | 20 % |
| 4 | 1 | - | - | - | - | - | - | - | - | - | 4 % |
| 5 | 4 | - | 3 | - | - | - | - | - | - | - | 28 % |
| 6 | - | - | 5 | - | 2 | - | - | - | - | - | 28 % |
| 7 | 4 | - | - | 2 | - | - | - | - | - | 1 | 28 % |
| 8 | 11 | - | - | - | - | - | - | 14 | - | - | 100 % |
| 9 | 1 | - | - | - | - | - | - | 13 | - | - | 56 % |
| Rata-Rata Persentase Error | | | | | | | | | | | 37,2% |

Tabel 11 menunjukkan kesalahan pengenalan paling besar terjadi pada angka "8" yang sama sekali tidak berhasil dikenali dengan benar, melainkan salah dikenali sebagai angka "0" sebanyak 11 kali dan salah dikenali sebagai angka 7 sebanyak 14 kali.

Dari penelitian ini menunjukkan bahwa input MLP menggunakan ekstraksi *longest run* menghasilkan pengenalan yang lebih baik

dibandingkan dengan ekstraksi *diagonal distance*. Presentase keberhasilan pengenalan MLP dengan input ekstraksi *longest run* (36 fitur) mencapai 100% pada sampel *training* sedangkan keberhasilan pengenalan MLP dengan input ekstraksi *diagonal distance* (4 fitur) hanya mencapai 10%. Presentase keberhasilan pengenalan MLP pada sampel *testing* dengan input ekstraksi *longest run* (36 fitur) mencapai 94,4% sedangkan keberhasilan pengenalan MLP dengan input ekstraksi *diagonal distance* (4 fitur) hanya mencapai 10%.

Kecilnya pengenalan MLP dengan input 40 fitur disebabkan karena hasil ekstraksi *diagonal distance* (4 fitur) yang buruk. Presentase keberhasilan pengenalan MLP dengan input ekstraksi 40 fitur atau hasil 2 metode ekstraksi yaitu *longest run* dan *diagonal distance* mencapai 66,8% pada sampel *training* dan 62,8% pada sampel *testing*. Sehingga ketika MLP hanya menggunakan input dari ekstraksi *longest run* (36 fitur) saja akan menghasilkan pengenalan yang lebih optimal.

KESIMPULAN

Ekstraksi *diagonal distance* kurang cocok bila digabungkan dengan *longest run* sebagai input MLP untuk pengenalan Angka Jawa, karena menghasilkan kesalahan pengenalan mencapai 33,2% dari 25 set pada sampel *training* dan 37,2% dari 25 set sampel *testing*.

Hasil ekstraksi fitur *longest run* sangat baik jika digunakan sebagai input MLP untuk pengenalan Angka Jawa, karena menghasilkan keberhasilan pengenalan mencapai 100% dari 25 set pada sampel *training* dan 94,4% dari 25 set sampel *testing*

DAFTAR PUSTAKA

- Wibowo, M.C. dan Wirakusuma, S., 2013, Pengenalan Pola Tulisan Tangan Aksara Jawa "Ha Na Ca Ra Ka" Menggunakan Multi Layer Perceptron, Prosiding Seminar Nasional Teknologi Informasi (SNASTI) 2013, Oktober 2013, Surabaya, hal. ICCS-27 – ICCS-32
- Das, N., dkk., 2006, Handwritten Arabic Numeral Recognition using a Multi layer Perceptron, Proceeding National Conference on Recent Trends in Information Systems, hal. 200 – 203.
- Ham, F.M. dan Kostanic, I., 2001, Principles of Neurocomputing for Science & Engineering. McGraw-Hill, New York.
- Darusuprpta, dkk., 2002, Pedoman Penulisan Aksara Jawa, Yayasan Pustaka Nusantara, Yogyakarta.