

Pengenalan Aksara Lampung Menggunakan Jaringan Syaraf Tiruan

Adhika Aryantio

School of Electrical Engineering and Informatics
Institute Technology of Bandung
10th Ganeca Street
Bandung, Indonesia.
Adhikaaryantio.x6@gmail.com

Rinaldi Munir

School of Electrical Engineering and Informatics
Institute Technology of Bandung
10th Ganeca Street
Bandung, Indonesia.
rinaldi@informatika.org

Abstract—Aksara Lampung merupakan warisan budaya yang perlu dilestarikan. Salah satu cara melestarikannya yaitu melalui pendekatan teknologi. Aksara Lampung memiliki jumlah aksara induk sebanyak 20 aksara dan setiap aksara dapat dilekati 23 jenis anak aksara. Hal tersebut mengharuskan komputer untuk menggunakan teknik pengenalan pola dalam mengenali Aksara Lampung. Struktur dari Aksara Lampung yang unik juga harus memerlukan algoritma khusus untuk mensegmentasi aksara Lampung.

Pada penelitian ini, akan dibuat aplikasi untuk mengenali aksara Lampung. Pengenalan pola terdiri dari pemindaian citra, pengolahan awal citra, ekstraksi fitur dan klasifikasi. Proses klasifikasi dilakukan dengan menggunakan jaringan syaraf tiruan propagasi balik.

Proses pemindaian citra dilakukan dengan mengambil citra suatu Aksara Lampung kemudian disimpan dalam bentuk PNG. Pengolahan awal akan dilakukan *cropping* dan *resize* setelah itu dilanjutkan teknik *zoning*. Ekstraksi fitur dilakukan dengan mengambil 64 fitur aksara dan disimpan pada *array of fitur*. Proses klasifikasi dilakukan oleh jaringan syaraf tiruan. Jaringan syaraf tiruan yang dibentuk memiliki 2 jaringan syaraf tiruan yaitu jaringan syaraf tiruan utama dan jaringan syaraf tiruan karakter.

Dalam pengujian sistem yang dibuat menggunakan data latih dan data uji. Data latih yang digunakan sebanyak 5040 data dengan komposisi 3 jenis *font* yang masing-masing memiliki ukuran *font* 24,48,dan 72. Data uji yang digunakan sebanyak 3360 data dengan *font* yang sama dengan data latih tetapi memiliki ukuran *font* 36 dan 60. Hasil akurasi yang didapatkan jaringan syaraf tiruan terhadap data latih adalah 100% dan akurasi terhadap data uji adalah 90,8%.

Kata Kunci—Aksara Lampung; Jaringan Syaraf Tiruan; Pengenalan Pola

I. PENDAHULUAN

Indonesia merupakan negara kepulauan yang memiliki beragam kebudayaan. Kebudayaan tersebut mencakup kesenian, tatanan bahasa hingga tingkah laku masyarakat. Salah satu bentuk tatanan bahasa yaitu berupa tulisan atau yang biasa disebut dengan aksara. Indonesia memiliki banyak jenis

aksara. Salah satu aksara yang dimiliki Indonesia merupakan Aksara Lampung.

Aksara Lampung memiliki 20 bentuk grafis yang memiliki ciri khas pada setiap aksaranya. Aksara Lampung juga dapat merangkai sebuah kalimat dengan menggabungkan induk aksara beserta anak aksara. Anak aksara pada Aksara Lampung terdapat 12 buah yang dapat diletakkan di atas, di bawah, maupun dikiri dari induk aksara. Perkembangan penelitian Aksara Lampung masih sebatas dalam pengembangan *font* Aksara Lampung yang dilakukan oleh Adi Yuza. Sedangkan, penelitian untuk pengenalan Aksara Lampung oleh komputer belum pernah dilakukan.

Keberagaman Aksara Lampung yang memiliki 20 bentuk grafis dengan kombinasinya menggunakan anak aksara membutuhkan teknik pengenalan pola agar komputer dapat mengenali suatu aksara. Menurut Khawaja, dkk [3], pengenalan pola suatu karakter memiliki beberapa tahapan yaitu pengambilan karakter, pengolahan awal karakter, ekstraksi fitur karakter dan pengenalan atau proses klasifikasi karakter.

Penelitian terkait tentang pengenalan pola yang dilakukan oleh [2][4] menggunakan jaringan syaraf tiruan propagasi balik sebagai sarana pelatihan. Akurasi yang didapatkan dengan menggunakan model jaringan syaraf tiruan propagasi balik pada [2] dan [4] memiliki performansi yang baik.

Pada makalah ini akan dibahas tentang pengenalan Aksara Lampung huruf cetak menggunakan jaringan syaraf tiruan. Model yang dihasilkan diharapkan dapat mengenali berbagai jenis Aksara Lampung.

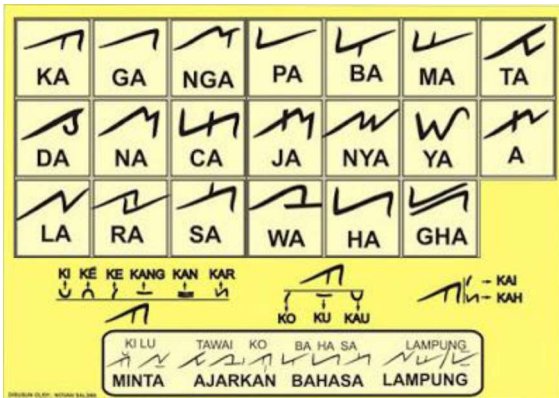
II. DASAR TEORI

Pada bab ini akan dijelaskan tentang beberapa teori yang digunakan dalam pembuatan makalah. Teori yang digunakan yaitu tentang Aksara Lampung, pengenalan pola, jaringan syaraf tiruan propagasi balik.

A. Aksara Lampung

Aksara Lampung merupakan tulisan aksara perkembangan dari aksara *Devanagari* yang berasal dari India Selatan. Aksara Lampung memiliki 20 jenis aksara dan 12 anak aksara yang dapat dilihat pada Gambar II-1. Anak aksara melekat

pada induk aksara akan menghasilkan jenis cara membaca yang berbeda-beda. Total jenis kombinasi cara membaca aksara Lampung terdapat 560 suku kata.



Gambar II-1. Aksara Lampung

B. Pengenalan Pola

Pengenalan pola terbagi menjadi 3 tahapan yaitu pemrosesan awal, ekstraksi fitur dan klasifikasi atau pengenalan.

Pemrosesan awal merupakan tahapan pertama yang dilakukan yang bertujuan untuk menyeragamkan citra masukan agak siap dilakukan ekstraksi fitur. Pada tahap pengolahan awal biasanya dilakukan *cropping*, penghilangan *noise* dan *resize* citra.

Ekstraksi fitur dilakukan bertujuan untuk mendapatkan ciri khas dari suatu citra. Ekstraksi fitur dapat dilakukan dengan mentransformasikan citra menjadi sebuah *array* atau memberikan tanda-tanda khusus yang mencerminkan ciri dari suatu citra.

Pengenalan atau klasifikasi bertujuan untuk mengenali atau mengklasifikasi citra tertentu ke dalam suatu kelas. Kelas yang dimaksud dalam pengenalan aksara merupakan cara membacanya. Pengenalan dilakukan dengan memasukkan hasil ekstraksi fitur dalam sebuah model. Hasil keluaran model merupakan kelas dari suatu karakter.

C. Jaringan Syaraf Tiruan Propagasi Balik

Propagasi balik merupakan pembelajaran *supervised*. Vektor masukan diberikan pada layer masukan dan dilakukan *feedforward* yang akan menghasilkan suatu keluaran. Kemudian, perhitungan MSE (Mean Square Error) dengan cara mengurangi target yang diinginkan dengan keluaran yang didapatkan. Hasil eror yang didapatkan kemudian di propagasikan balik hingga ke *layer* masukan untuk menyesuaikan bobot-bobot pada setiap *epoch*.

Algoritma propagasi balik merupakan prosedur iteratif. Pembelajaran dilakukan terus-menerus setiap 1 set pasangan masukan dan keluaran sampai kesalahan dari keluaran mencapai nilai yang mendekati nilai yang diinginkan. Ketika nilai tersebut sudah memenuhi nilai minimal eror maka pembelajaran disebut konvergen.

Berikut merupakan algoritma propagasi balik

1. *Inisialisasi* berat dan *threshold* dengan nilai random yang kecil
2. Kenalkan data masukan dan tentukan keluaran yang diinginkan dari himpunan latihan.
3. Hitung keluaran aslinya dengan menggunakan persamaan (II.1) dan (II.2)

$$Sum_i^n = \sum_j w_{ij}^n o_j^{n-1} \tag{II.1}$$

$$o_i^n = f_s(sum_i^n) \tag{II.2}$$

4. Hasil dari keluaran pada jaringan propagasi balik, selanjutnya hitung perbedaan beban berdasarkan persamaan II.3

$$\Delta w_{ij}(t) = \eta \delta_i a_j + \alpha \Delta w_{ij}(t - 1) \tag{II.3}$$

5. Pada awal dari *layer* keluaran, modifikasi berat dari setiap *layer* berdasarkan persamaan (II.4)

$$w_{ij}^n(t + 1) = w_{ij}^n(t) + \Delta w_{ij}^n \tag{II.4}$$

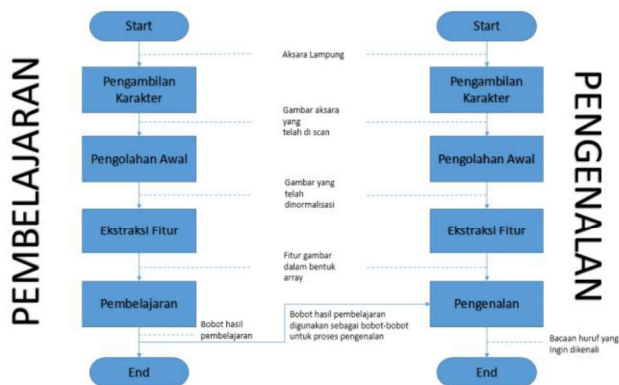
6. Kembali pada tahapan 2 dan lakukan terus hingga pasangan masukan dan keluaran cocok dengan himpunan latihan.

Keterangan :

- i_i^n : masukan *node i* pada *layer n*
- o_i^n : keluaran *node i* pada *layer n*
- w_{ij}^n : nilai dari beban dari *node j* pada (*n-1*) *layer* dengan *node i* pada (*n*) *layer*
- Sum_i^n : penjumlahan berdasarkan persamaan (1) dari *node i* pada *layer n*
- t_i : nilai yang diinginkan atau nilai target *node i* pada *layer* keluaran
- Δw_{ij} : Hasil modifikasi beban setiap w_{ij} untuk meminimalkan kesalahan
- δ_i : $a_i(1 - a_i)(t_i - a_i)$ untuk unit keluaran
- δ_i : $a_i(1 - a_i) \sum_k \delta_k w_{ik}$ untuk unit *hidden*
- a_i : level dari unit yang aktif
- t_i : keluaran dari unit target
- α : konstanta momentum antara 0 sampai 1
- η : *learning rate*.

III. SISTEM PENGENALAN AKSARA LAMPUNG

Sistem pengenalan Aksara Lampung yang dibuat memiliki 2 proses utama yaitu proses pembelajaran dan proses pengenalan. Pada setiap proses memiliki tahapan pengambilan citra, pengolahan awal citra, ekstraksi fitur citra, pembelajaran pada proses pembelajaran dan pengenalan pada proses pengenalan. Sistem keseluruhan dapat dilihat pada Gambar III-1.



Gambar III-1. Sistem Keseluruhan Pengenalan Aksara Lampung

A. Pengambilan Citra

Pada tahapan pengambilan citra dilakukan pengambilan citra berupa Aksara Lampung dengan cara memindai suatu objek Aksara Lampung atau dengan cara melakukan *snipping* pada tulisan Aksara Lampung yang terdapat pada *text editor*. Hasil pemindaian disimpan dalam bentuk PNG *format image*.

B. Pengolahan Awal Citra

Pada tahapan pengolahan awal citra, citra hasil pemindaian dinormalisasikan agar semua citra dari aksara mempunyai bentuk yang standar. Tahapan yang dilakukan dalam pengolahan awal citra yaitu :

1. Menghilangkan daerah putih pada suatu Aksara Lampung sehingga didapatkan Aksara Lampung yang utuh.
2. Melakukan *resize* ukuran Aksara Lampung menjadi berukuran 64x64 pixel.

C. Ekstraksi Fitur

Pada tahapan ekstraksi fitur dilakukan pengambilan fitur dari setiap aksara yang berukuran 64x64 pixel. Fitur yang diambil pada pengenalan Aksara Lampung merupakan komposisi Aksara Lampung pada citra. Tahapan ekstraksi fitur yang dilakukan adalah sebagai berikut :

1. Mentransformasikan pixel-pixel tersebut ke dalam sebuah nilai 0 dan 1. Nilai 0 diberikan terhadap pixel yang memiliki citra berwarna putih. Sedangkan, nilai 1 diberikan terhadap pixel yang memiliki citra berwarna hitam.
2. Melakukan teknik *zoning* dengan cara mengubah setiap 8x8 pixel menjadi 1 pixel. Nilai dari pixel yang

baru dihitung dengan persamaan III-1. Teknik zoning dapat dilihat pada Gambar III-2.

$$f_v = (1/pv) \sum C_{ij} \tag{III-1}$$

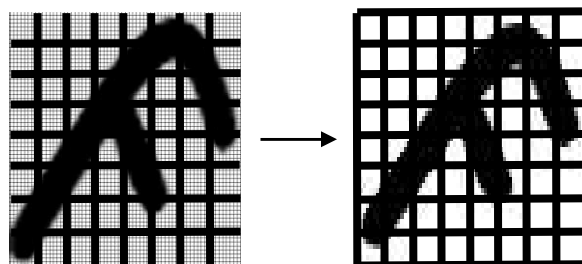
Keterangan :

f_v = *feature value* pada submatrix yang ke v th

pv = jumlah sel dalam submatrix v

C_{ij} = nilai dari sel (i,j) th pada submatrix

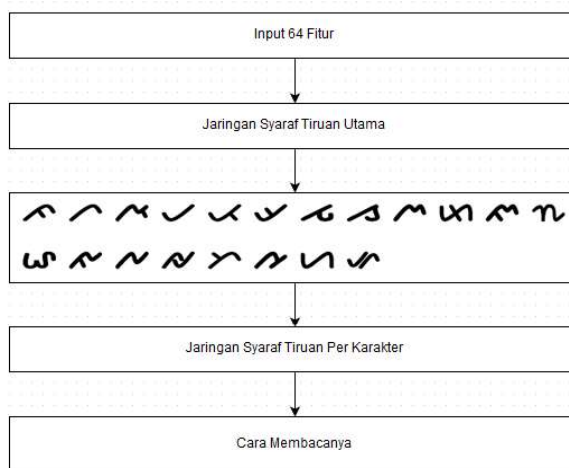
3. Hasil dari teknik *zoning* merupakan *array of real* yang berisi nilai dari 0-1 dengan jumlah 64 fitur.



Gambar III-2. Zoning

D. Arsitektur Jaringan Syaraf Tiruan

Pada proses pelatihan dan pengenalan diperlukan arsitektur jaringan syaraf tiruan yang telah diimplementasikan. Jaringan syaraf tiruan yang dibentuk dalam sistem pengenalan Aksara Lampung memiliki 2 jaringan syaraf tiruan. Jaringan syaraf tiruan 1 atau jaringan syaraf tiruan utama bertujuan untuk mengenali aksara induk. Sedangkan, jaringan syaraf tiruan 2 atau jaringan syaraf tiruan karakter untuk mengenali anak aksara. Struktur jaringan syaraf tiruan yang dibentuk dapat dilihat pada Gambar III-3.



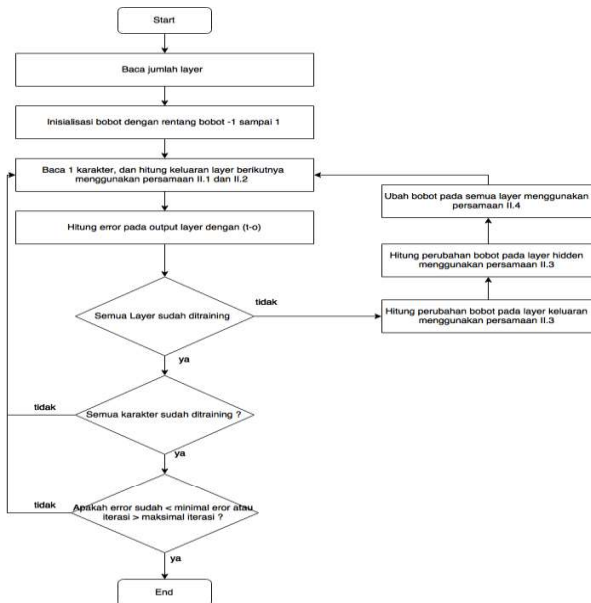
Gambar III-3. Arsitektur Jaringan Syaraf Tiruan

Jaringan syaraf tiruan utama dan jaringan syaraf tiruan karakter dibentuk dengan 64 *node* masukan, 46 *hidden node* dan 5 *node* keluaran. 64 *node* masukan berasal dari 64 fitur yang menjadi masukan jaringan syaraf tiruan, 5 *node* keluaran

berasal dari jumlah aksara yang ingin dikenali yaitu 20 pada jaringan syaraf utama dan 23 pada jaringan syaraf karakter sehingga jumlah nilai biner dari 2^5 bisa menampung jumlah aksara yang ingin dikenali.

E. Pembelajaran

Pembelajaran dilakukan dengan pembelajaran *supervised* menggunakan jaringan syaraf tiruan propagasi balik. Pembelajaran dilakukan pada jaringan syaraf tiruan utama dan jaringan syaraf tiruan karakter. Pembelajaran memerlukan target dari setiap aksara untuk mendapatkan bobot-bobot jaringan syaraf tiruan. Proses pembelajaran dapat dilihat pada Gambar III-4.



Gambar III-4 Proses Pembelajaran

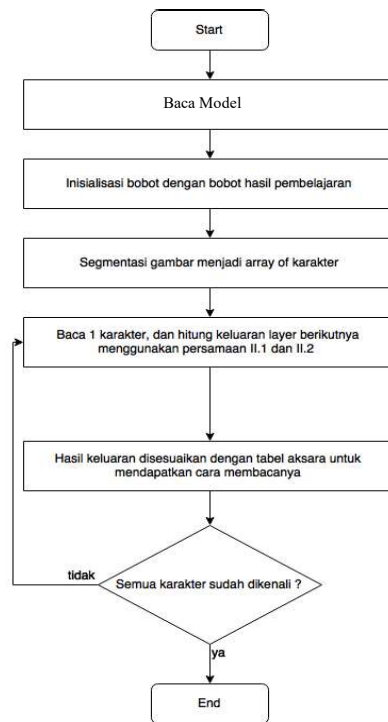
F. Pengenalan

Pada proses pengenalan masukan merupakan 1 citra utuh yang membentuk suatu kalimat. Pengenalan pada jaringan syaraf tiruan dilakukan 1 per 1 aksara. Hal tersebut mengharuskan melakukan segmentasi terlebih dahulu sebelum dimasukkan ke dalam jaringan syaraf tiruan.

Segmentasi Aksara Lampung dilakukan dengan tahapan berikut :

1. Ambil 1 citra kalimat Aksara Lampung
2. Potong setiap bagian Aksara Lampung yang dipisahkan bagian putih secara horizontal.
3. Hitung jarak setiap potongan , bila potongan memiliki jarak yang kecil dengan aksara induk maka digabungkan, bila tidak maka dipisahkan.
4. Potong setiap bagian Aksara Lampung yang dipisahkan bagian putih secara vertikal.
5. Hitung jarak setiap potongan, bila potongan memiliki jarak yang kecil dengan aksara induk maka digabungkan, bila tidak maka dipisahkan.
6. Hasil dari pemotongan merupakan *array of* Aksara Lampung.

Proses pengenalan dengan masukan *array of* aksara akan menghasilkan cara baca dari setiap karakter. Proses pengenalan dapat dilihat pada Gambar III-5.



Gambar III-5. Proses Pengenalan

IV. EKSPERIMEN

Pada bab ini akan dijelaskan tentang eksperimen pembentukan jaringan syaraf tiruan. Bab ini terdiri dari pembentukan data latih dan data uji, menentukan optimum *hidden node*, pembentukan jaringan syaraf tiruan.

A. Pembentukan Data Latih dan Data Uji

Data latih untuk pembelajaran jaringan syaraf tiruan terdiri dari 3 *font* Aksara Lampung yaitu Aksara Lampung Rounded, Aksara Lampung YZ dan Aksara Lampung Liuk dan setiap jenis aksara dengan ukuran *font* 24,48 dan 72. Data latih yang terbentuk sebanyak 5040 aksara.

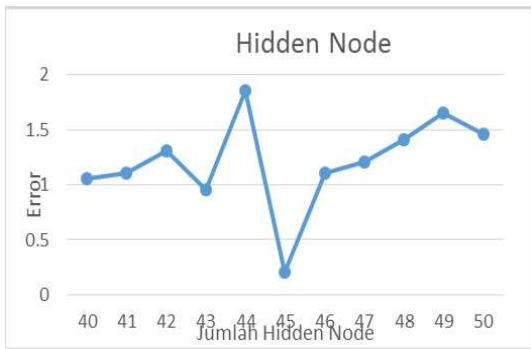
Data uji yang dibentuk menggunakan 3 *font* Aksara Lampung yang sama dengan data latih dan setiap jenis aksara dengan ukuran 36 dan 60. Data uji yang terbentuk sebanyak 3360 aksara.

B. Jumlah Hidden Node Optimum

Hidden node berdasarkan “rules of thumbs” memiliki nilai optimum pada $(2/3 * (input\ node + output\ node))$, pada makalah ini yaitu berjumlah 46. Oleh karena itu dilakukan eksperimen menggunakan *hidden node* disekitar 46 yaitu pada rentang 40-50 untuk mencari *hidden node* optimum.

Setting eksperimen yaitu diambil 1 *setting* dengan *learning rate* 0.4, momentum 0,3 dan *epoch* 5000 kali dilakukan 10 kali pengulangan. Hasil *node* yang paling optimum yaitu pada

hidden node berjumlah 45. Hasil dapat dilihat pada Gambar IV-1.



Gambar IV-1. Grafik Eksperimen Hidden Node

C. Pembentukan Jaringan Syaraf Tiruan

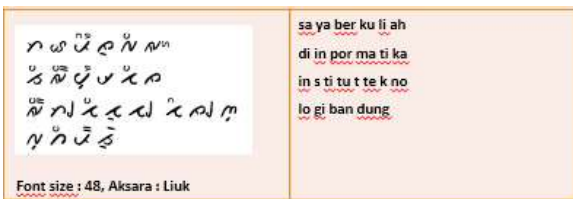
Jaringan syaraf tiruan dibentuk dengan menggunakan 46 node masukan, 45 node hidden dan 5 node keluaran. Pembelajaran dilakukan berkali-kali dengan learning rate dan momentum yang adaptif. Hasil model jaringan syaraf tiruan utama yang terbaik yaitu dengan memiliki eror keseluruhan 0.5 dan hasil model jaringan syaraf tiruan karakter yang terbaik yaitu dengan memiliki eror keseluruhan 0.1.

V. PENGUJIAN

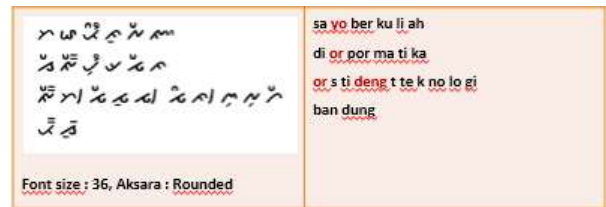
Pengujian model dibentuk dilakukan terhadap data latih dan data uji. Pengujian dilakukan dengan membandingkan hasil keluaran dengan hasil target yang diinginkan dengan persamaan V-1.

$$Akurasi = \left(\frac{Jumlah\ Benar}{Jumlah\ Data} * 100\% \right) \quad (V-1)$$

Hasil akurasi pengujian terhadap data latih yaitu 100% dan hasil akurasi pengujian terhadap data uji yaitu 90.8%. Akurasi yang dimiliki data uji cukup besar karena memiliki tingkat kemiripan yang tinggi dengan data latih. Data uji dan data latih hanya berbeda pada ukuran font-nya saja. Hasil pengujian pengenalan terhadap data latih dan data uji dapat dilihat pada Gambar V-1 dan Gambar V-2.



Gambar V-1. Pengujian Pengenalan Data Latih



Gambar V-2. Pengujian Pengenalan Data Uji

VI. KESIMPULAN

Berdasarkan hasil pengujian dan eksperimen dapat diambil kesimpulan yaitu :

1. Tahapan yang dilakukan untuk mendapatkan masukan yang tepat untuk jaringan syaraf tiruan yaitu melakukan pengolahan awal berupa crop dan resize dan ekstraksi fitur dengan menggunakan zoning.
2. Jaringan Syaraf Tiruan dapat mengenali Aksara Lampung dengan baik terhadap data latih dan data uji.
3. Akurasi terhadap data latih tinggi yaitu 100% dan akurasi terhadap data uji yaitu 90.8%.

ACKNOWLEDGMENT

Penulis berterima kasih kepada Dr. Ir. Rinaldi Munir, M.T selaku dosen pembimbing yang telah memberikan bimbingan dan masukan dalam pengerjaan makalah.

REFERENCES

- [1] Haykin,S (1999), Neural Networks A Comprehensive Foundation Second Edition. Prentice Hall, Inc : New Jersey.
- [2] Handoyo, Erico Darmawan, dkk (2011), Penerapan Jaringan Syaraf Tiruan metode Propagasi Balik Dalam Pengenalan Tulisan Tangan Huruf Jepang Jenis Hiragana dan Katakana. Jurnal Informatika Universitas Kristen Maranata 2011.
- [3] Khawaja, Attaullah, dkk (2006), Recognition of printed Chinese characters by using Neural Network, IEEE Conference Publication.
- [4] Budiwati,dkk (2011), Japanese Character (Kana) Pattern Recognition Application Using Neural Network, IEEE Conference Publication.
- [5] Sanossian, Hermineh Y.Y. (1996), An Arabic Character Recognition System Using Neural Networks, IEEE Conference Publication.
- [6] Introduction to Artificial Neural Networks (1995), hal 36-62, IEEE Conference Publication.
- [7] Abed, Majida Ali dan Hamid Ali Abed Alasad (2015), High Accuracy Arabic Handwritten Characters Recognition Using Error Back Propagation Artificial Neural Networks, (IJACSA) International Journal of Advanced Computer Science and Applications.
- [8] <http://www.cse.wustl.edu/~bayazit/courses/cs527a/> Tanggal Akses : 10 November 2014
- [9] <http://sosbud.kompasiana.com/2013/02/03/font-aksara-lampung-untuk-penulisan-di-komputer-525252.html>. Tanggal Akses : 1 November 2014
- [10] <http://www.apakabardunia.com/2013/09/yuk-mengenal-abjad-bahasa-lampung-dan.html>. Tanggal Akses : 20 Desember 2014
- [11] <http://aksaralampung.blogspot.com>. Tanggal Akses : 1 November 2014