
Pengenalan Tulisan Tangan Huruf Hangul dengan Menggunakan Jaringan Saraf Tiruan Propagasi

Purbarini Sulysthian, Dr. Anastasian Rita Widiarti

Teknik Informatika

Universitas Sanata Dharma Yogyakarta, Kampus 3, Paingan, Maguwoharjo, Sleman,
Yogyakarta

e-mail: rinisulysthian@gmail.com

Abstrak

Perkembangan teknologi semakin memudahkan manusia dalam mengakses dan menerima informasi dari berbagai hal. Berbagai hal dapat diakses dengan mudah, murah, dimana saja, dan kapan saja. Salah satu yang kita dapat akses adalah kebudayaan dari berbagai macam negara, contohnya adalah Korea Selatan. Negara dengan budaya K-Pop nya yang saat ini sedang hangat dibicarakan oleh anak muda Indonesia. Berbicara tentang budaya K-Pop dan kebudayaan Korea maka berbicara juga tentang Bahasa Korea. Bahasa Korea adalah bahasa yang menggunakan aksara dalam penulisannya, aksara ini disebut dengan Hangul.

Dengan banyaknya masyarakat yang belajar bahasa Korea dan banyaknya turis Indonesia yang datang ke Korea, maka dibutuhkan suatu teknologi yang dapat memberikan informasi kepada penggunanya tentang bahasa Korea. Dalam penelitian ini akan membuat sistem dengan menggunakan salah satu penerapan dari Jaringan Saraf Tiruan. Sistem ini dibangun dengan Jaringan Saraf Tiruan Propagasi Balik sebagai pemodelan system dan Intensity of Character dan Mark Direction sebagai metode ekstraksi cirinya. Sistem ini nantinya diharapkan mampu mengenali pola tulisan huruf Hangul secara otomatis dan memberikan informasi terkait dengan pola yang dikenalnya.

Percobaan proses pengenalan dengan menggunakan Jaringan Saraf Tiruan Propagasi Balik dengan melakukan optimalisasi arsitektur jaringan dan menghasilkan akurasi 80,4762%. Percobaan ini menggunakan 2 lapisan tersembunyi dengan jumlah neuron pada lapisan tersembunyi 1 adalah 100 dan pada lapisan tersembunyi 2 adalah 95 neuron.

Kata Kunci: *Pengenalan Pola, Hangul, Intensity of Character, Mark Direction, Jaringan Saraf Tiruan Propagasi Balik*

1. Pendahuluan

Majunya dunia teknologi semakin memudahkan manusia untuk mengakses dan menerima informasi dari berbagai hal. Mendekatkan yang jauh adalah salah satu dampak nyata dari perkembangan dunia teknologi. Berbagai hal dapat kita akses dengan mudah, murah, dimana saja dan kapan saja. Salah satunya adalah kebudayaan, kita dapat mengetahui berbagai macam kebudayaan dari berbagai negara hanya dengan mengakses melalui jaringan internet saja. Berbicara tentang kebudayaan di seluruh dunia, tidak lengkap jika kita tidak berbicara tentang Korea Selatan, negara dengan berbagai kebudayaan, adat

istiadat dan budaya K-Popnya. Budaya dari negara ini tidak henti-hentinya dibahas oleh anak muda Indonesia yang sedang terserang virus *Hallyu wave* (gelombang Korea). Berbicara mengenai budaya K-Pop dan kebudayaan Korea Selatan, maka berbicara juga tentang bahasa Korea. Bahasa Korea berbeda dengan bahasa lain seperti Indonesia, Inggris, Jerman, dan negara-negara lainnya yang menggunakan alfabet ABCD untuk menulis. Bahasa Korea adalah bahasa yang menggunakan aksara dalam penulisannya sama seperti Mandarin dan Jepang. Huruf ini digunakan dalam kehidupan sehari-hari, dan mudah kita temukan dimana saja, seperti lagu-lagu Korea, kebudayaan Korea, drama Korea dan lainnya. Dan hal inilah yang membuat banyak anak muda Indonesia beramai-ramai untuk belajar aksara ini agar dapat mengerti dan dapat memahami kebudayaan Korea. Hal ini diperkuat dengan adanya data dari website *kto.visitkorea.or.kr* yang menunjukkan jumlah wisatawan Indonesia yang berkunjung ke negara Ginseng ini mencapai 25,139 orang per Agustus 2016. Bahkan Indonesia menempati urutan ketujuh wisatawan terbanyak dari wisatawan seluruh dunia yang pernah berkunjung di negara Korea Selatan. Dengan adanya fenomena ini, maka dibutuhkan suatu alat yang dapat membantu wisatawan dan penggemar budaya Korea dalam memahami bahasa Korea, khususnya aksara Hangul.

Penelitian tentang pengenalan aksara sudah pernah dan banyak dilakukan oleh beberapa orang, mulai dari aksara Jepang, Hangul, Thailand, aksara Jawa, dll. Pada penelitian pengenalan aksara Hangul sebelumnya, yang dilakukan oleh Yoon, Wook dan Joon (1999) hasil keakuratan yang dihasilkan cukup tinggi, yaitu mencapai 96,5% pada 17500 data test dan 98,9% pada 19361 data training. Penelitian ini menggunakan metode Jaringan Saraf Tiruan dan Fuzzy Function, Jaringan Saraf Tiruan digunakan untuk mengklasifikasi aksara dan Fuzzy Function digunakan untuk menyesuaikan sistem dengan kebiasaan individu menulis aksara. Pada penelitian ini error disebabkan oleh tulisan yang tidak normal, hubungan posisi tulisan yang ambigu, karakter yang tidak dikenal dan penghubung yang terlalu banyak. Peneliti menggunakan citra dari pena elektronik yang digambarkan pada sebuah tablet dan pada proses preprocessing peneliti menghilangkan noise dan hook pada aksara. Penelitian tentang pengenalan aksara Hangul lainnya dengan menggunakan metode Jaringan Saraf Tiruan Propagasi Balik yang dilakukan oleh Fathia (tanpa tahun), juga menghasilkan tingkat keakuratan yang tinggi, yaitu mencapai 99.7217% dan tingkat error sebesar 0.2783%. Penelitian lainnya yang menggunakan Jaringan Saraf Tiruan Propagasi Balik adalah penelitian yang dilakukan oleh Nugroho (2016), penelitian ini menggunakan Jaringan Saraf Tiruan Propagasi Balik untuk mengenali pola huruf Jepang Hiragana. Dan akurasi tertinggi yang didapat adalah 86,63% dengan menggunakan 2 lapisan tersembunyi.

Jaringan Saraf Tiruan adalah Jaringan Syaraf Tiruan (*Artificial Neural Network* (ANN)), atau juga disebut *Simulated Neural Network* (SNN), atau umumnya hanya disebut *Neural Network*(NN)), adalah jaringan dari sekelompok unit pemroses kecil yang dimodelkan berdasarkan jaringan syaraf manusia. Seperti halnya otak manusia, jaringan syaraf juga terdiri

dari beberapa neuron, dan terdapat hubungan antara neuron-neuron tersebut. Neuron-neuron tersebut akan mentransformasikan informasi yang diterima, melalui sambungan keluarnya menuju ke neuron-neuron yang lain. Pada jaringan syaraf hubungan ini dikenal dengan nama bobot. Informasi tersebut disimpan pada suatu nilai tertentu pada bobot tersebut.

Sedangkan jaringan saraf propagasi balik merupakan jaringan saraf yang menggunakan konsep jaringan berlapis jamak. Lapisan pertama adalah lapisan masukan (input) dan yang terakhir adalah lapisan keluaran (output). Lapisan diantara lapisan masukan dan lapisan keluaran disebut dengan lapisan tersembunyi (hidden).

Dengan adanya berbagai penelitian tentang pengenalan aksara dan penjelasan diatas, metode Jaringan Saraf Tiruan Propagasi Balik relevan dengan topik yang saya ambil. Karena dengan metode ini dapat mengenali aksara dengan baik dan hasil keakuratannya pun tinggi.

2. Tahap-tahap Penelitian

Tahapan dari penelitian ini antara lain:

I. Pengumpulan Data

Data didapatkan melalui form yang di bagikan kepada mahasiswa semester 5 jurusan Bahasa Korea, Universitas Gadjah Mada. Pada form yang dibagikan terdapat 2 kolom kosong disamping setiap 14 konsonan huruf Hangul. Mahasiswa diberi kesempatan untuk menulis tulisan tangan huruf Hangul sebanyak 2×14 data. Sehingga total yang diperoleh ada $25 \times 2 \times 14$ data

II. Pengolahan Data

a. Digitalisasi Data

Setiap form yang didapatkan akan di scan melalui scanner, tujuan dari proses ini adalah mendapatkan data digital dari data tulisan tangan yang sebelumnya didapat pada form kuisioner.

b. Meng-*crop* Data Digital

Kegiatan meng-*crop* data dilakukan dengan cara memotong form yang sudah di-*scan* untuk diambil data perhurufnya.



c. Pemrosesan Citra

1) Mengubah Citra Warna

Kegiatan ini bertujuan untuk mengubah citra yang tadinya berwarna menjadi citra keabuan, setelah mendapat citra keabuan langkah selanjutnya adalah mengubah citra keabuan menjadi citra hitam putih.



2) *Resizing* Citra

Kegiatan ini dilakukan untuk mengubah ukuran dari citra, dari berbagai citra yang tidak sama ukurannya kemudian disamakan ukurannya menjadi 64 x 64 piksel. Hal ini dimaksudkan agar data yang diolah tidak terlalu besar dan dapat dibagi menjadi 4 segmen.



3) *Thinning* Citra

Kegiatan ini dilakukan untuk mengubah ukuran dari citra, dari berbagai citra yang tidak sama ukurannya kemudian disamakan ukurannya menjadi 64 x 64 piksel. Hal ini dimaksudkan agar data yang diolah tidak terlalu besar dan dapat dibagi menjadi 4 segmen.



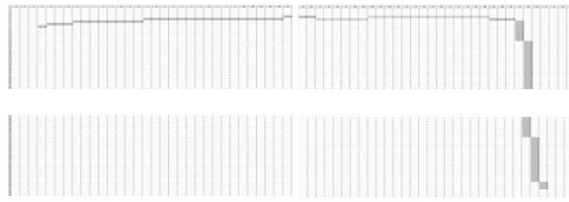
4) Membagi Citra

Kegiatan ini bertujuan untuk membagi citra menjadi beberapa segmen yang lebih kecil. Tujuan dari kegiatan ini adalah lebih mempermudah dalam pencirian citra. Pada penelitian ini citra akan dibagi menjadi 4 segmen.



5) Mengambil Ciri Citra

Pada penelitian ini pengambila ciri citra menggunakan ekstrasi ciri *Intensity of Character* dan *Mark Direction*. *Intensity of Character* digunakan untuk mendapatkan jumlah piksel hitam dalam sebuah citra. Sedangkan *Mark Direction* digunakan untuk menghitung banyaknya piksel yang memiliki tetangga yang berarah horizontal, vertikal, diagonal ke kanan, dandiagonal ke kiri. Pada penelitian ini, *Intensity of Character* dan *Mark Direction* digunakan pada setiap segmen dari citra yang telah dibagi menjadi 4 segmen.



Contoh Citra yang sudah diekstrak Cirinya

Tabel Contoh Hasil Ekstrasi Ciri

Segmen ke	IoC	Mark Direction	Mark Direction	Mark Direction	Mark Direction
		Horizontal	Vertikal	Diagonal 1	Diagonal 2
1	929	29	3	0	6
2	895	26	54	10	2
3	1024	0	0	0	0
4	951	3	54	0	0

d. Mempersiapkan dan membuat alat uji algoritma *Backpropagation*

Langkah pertama dalam kegiatan ini adalah mengambil data ciri menggunakan script `load_dataaaa.m` dan menormalisasi data tersebut. Kemudian membuat target data menggunakan script `target_data.m` dan dimasukkan ke variabel label. Kemudian mengubah vector target menjadi sebuah indeks dengan menggunakan script `Bagi_Target.m` dan dimasukkan ke variabel I1 sampai I14.

e. Pelatihan Alat

```

3 - net = newff(minmax(datanormal),[100, 80, 14],{'logsig','logsig','logsig'}, 'trainscg');
4 - net.trainParam.epochs = 500;
5 - net.divideFcn = 'divideind';
6 - net = init(net);
7 - net.trainParam.lr = 0.01;

```

Contoh Program Jaringan Propagasi Balik

Pada penelitian ini ada 43 data untuk setiap hurufnya dan ada 14 huruf dan total semua huruf ada 602 huruf. Untuk pelatihan alat digunakan data sebanyak 2/3 bagian untuk setiap hurufnya. Jadi total data yang digunakan dalam pelatihan alat 28 data huruf. Dikarenakan ada 602 huruf, maka total data dari semua huruf yang digunakan untuk data pelatihan sebanyak 392 huruf. Fungsi aktivasi yang digunakan dalam penelitian ini adalah `logsig`. Jumlah neuron yang dipakai ada beberapa, mulai dari 50 unit hingga 100 unit. Target yang telah disiapkan sebanyak 14 target karena terdapat 14 huruf Hangul yang berupa matriks.

Percobaan pelatihan alat dalam penelitian ini dilakukan sebanyak 12 kali. Pada setiap percobaan telah dilakukan perubahan jumlah layer, jumlah neuron, atau fungsi aktivasi yang digunakan pada setiap layer untuk melihat peningkatan akurasi.

f. Pengujian Alat

Pengujian alat dilakukan dengan menggunakan 6 kondisi ciri yang berbeda. Setiap percobaan dengan kondisi ciri yang berbeda, jumlah unit pada setiap hidden layer, atau jumlah iterasi yang dipakai telah dirubah dan terlihat ada peningkatan dan penurunan akurasi.

III. Analisis Data

Pengujian dilakukan untuk menguji apakah pengenalan pola yang dilakukan sudah sesuai atau belum. Pengujian dalam penelitian ini adalah dengan menghitung keakurasian sistem yaitu dengan menghitung hasil jumlah data yang sesuai/benar dibagi dengan jumlah seluruh data. Pengujian dilakukan dengan rumus :

$$\text{Akurasi (\%)} = \frac{\text{Jumlah data benar}}{\text{Jumlah semua data}} \times 100\%$$

3. Analisa dan Hasil

Pada penelitian ini, peneliti melakukan dua belas percobaan dengan jumlah hidden layer dan fungsi aktivasi yang berbeda. Untuk setiap percobaan kombinasi ciri yang dipakai ada 6 kombinasi, yaitu:

1. Kombinasi ciri 1 (ciri dari *Intensity of Character*)
2. Kombinasi ciri 2 (ciri dari *Mark Direction* diagonal 1)
3. Kombinasi ciri 1 dan ciri 2
4. Kombinasi ciri 1 dan ciri 3 (ciri dari *Mark Direction* diagonal 2)
5. Kombinasi ciri 2 dan ciri 3
6. Kombinasi ciri 1, ciri 2, ciri 3, ciri 4 (ciri dari *Mark Direction* horisontal), dan ciri 5 (ciri dari *Mark Direction*)

3.1. Hasil Uji Percobaan

Setelah melakukan 12 kali percobaan dengan mengubah metode ekstraksi ciri, banyaknya jumlah *hidden layer*, jumlah neuron pada *hidden layer*, fungsi aktivasi pada *hidden layer* dan fungsi aktivasi pada jaringan terlihat adanya perubahan pada akurasi alat uji.

Hasil terbaik dapat dilihat pada percobaan 12, dengan banyak data benar/data yang dapat dibaca ada 168 data huruf dengan akurasi mencapai 80,4762%. Dapat dilihat juga data huruf paling sedikit yang dapat dibaca ada pada percobaan 2, dimana akurasi yang didapat adalah 33,1633%. Dari hasil tersebut dapat dikatakan bahwa alat uji yang telah dibuat dapat mengenali tulisan tangan huruf Hangul.

Selain itu juga telah dilakukan pengujian data tunggal untuk melihat hasil dari pelatihan jaringan. Pengujian data tunggal ini menggunakan 70 data huruf. Dari 70 data, data yang dikenali dengan baik ada 43 huruf dengan rata-rata akurasi yang dihasilkan mencapai 61,42857%. Akurasi yang rendah disebabkan data yang digunakan kurang jelas dan kurang baik dalam penulisannya.

Hasil percobaan pengujian data tunggal data menunjukkan bahwa hasil yang didapat sangat beragam, total data yang dapat dikenali dengan baik 43 huruf sedangkan yang tidak dapat dikenali ada 24 huruf, dan rata-rata akurasi yang didapat sebesar 61,42857%. Hasil huruf Hangul yang dapat dikenali dengan baik adalah huruf G, N, S, H. Hal ini disebabkan data yang ditulis kurang baik dan kurang jelas, sehingga saat diubah menjadi citra hitam putih bagian tidak jelasnya tidak bisa dikenali.

4. Kesimpulan

Dari hasil penelitian yang telah dilakukan disimpulkan jika alat uji yang dibuat dengan menggunakan algoritma propagasi balik dapat digunakan untuk mengenali tulisan tangan huruf Hangul. Hal ini dapat dilihat dengan adanya beberapa huruf yang terbaca oleh alat uji yang menggunakan algoritma propagasi balik. Tingkat akurasi tertinggi didapat pada percobaan 12 dengan akurasi sebesar 80,4762% menggunakan kombinasi ciri 6 yang merupakan matriks kombinasi dari ciri *Intensity of Character*, *Mark Direction* horz, *Mark Direction* vert, *Mark Direction* dig1, *Mark Direction* dan dig2. Pola jaringan yang digunakan adalah jaringan dengan 2 hidden layer dan 1 lapisan keluaran, dengan jumlah unit neuron pada hidden layer 1 dan hidden layer 2 100, 95, dan pada lapisan keluaran jumlah unit neuronnya ada 14. Fungsi aktivasi yang digunakan pada hidden layer 1 dan hidden layer 2 adalah *logsig*, *logsig*, pada lapisan keluaran *logsig*, dan pada pelatihan jaringan menggunakan fungsi aktivasi *trainscg*. Epoch yang digunakan adalah 750 dengan waktu perhitungan 00:03, dan MSE sebesar 1.00E-06.

5. Saran

Dari hasil penelitian pengenalan tulisan tangan huruf Hangul ini, ada beberapa saran yang untuk memperbaiki kekurangan pada penelitian tersebut, antara lain:

1. Mencari data tulisan tangan huruf Hangul yang lebih banyak.
2. Tentukan Batasan-batasan dalam responden menulis tulisan tangan, jangan campurkan antara tulisan tangan huruf Hangul normal dan tulisan tangan huruf Hangul yang digayakan.
3. Mencari dan menggunakan metode ekstraksi ciri yang berbeda.
4. Mengubah susunan jaringan dengan menambah jumlah hidden layer, jumlah neuron pada hidden layer, fungsi aktivasi pada hidden layer, dan fungsi aktivasi pada jaringan.

Referensi

- [1] Keechul and Hang Joon Kim, 1999, "On-line recognition of cursive Korean characters using graph representation". Taegu, South Korea
- [2] Handoyo, Erico Darmawan, Lydia Wiguna Susanto, 2011. *Penerapan Jaringan Syaraf Tiruan metode Propagasi Balik Dalam Pengenalan Tulisan Tangan Huruf Jepang Jenis Hiragana*

dan Katakana. Universitas Kristen Manarata. Manarata University Press : Bandung. ISSN 0216-4280

- [3] Nugroho, Nicolaus Euclides Wahyu, 2016. "*Pengenalan Pola huruf Jepang Hiragana Menggunakan Algoritma Backpropagation*". Universitas Sanata Dharma, Yogyakarta.
- [4] Handoyo, Kasih, 2017. "*Transliterasi Nama Jalan Beraksara Jawa*". Universitas Sanata Dharma, Yogyakarta.
- [5] Gonzales, Rafael C. & Woods, Richard E., 2008. *Digital Image Processing*(3rd ed.). United States of America: Pearson Prentice Hall.
- [6] Kristanto, Andri, 2004. "*Jaringan Syaraf Tiruan, Konsep Dasar, Algoritma dan Aplikasi*". Gava Media. Yogyakarta.
- [8] Surinta, Olarik, 2010. "*Overview of Handwritten Thai Character Recognition*". Diakses dari <http://www.ai.rug.nl/~molarik/APSMeeting/09-07-2010%20Overview%20of%20Handwritten%20Thai%20Character%20Recognition.pdf> pada tanggal 29 september 2017.
- [9] Fathia, Siska. 2013. "*Penerapan Jaringan Saraf Tiruan Dalam Pengenalan Tulisan Tangan Huruf Korea (Hangul) Menggunakan Metode Propagasi Balik*". Universitas Dian Nuswantoro, Semarang.
- [10] Nasution, A. H., dan Prasetyawan, Y. 2008. *Perencanaan & Pengendalian Produksi. Edisi Pertama*. Graha Ilmu, Yogyakarta.
- [11] Korea, Monthly Statistics. <http://kto.visitkorea.or.kr>. Diakses pada tanggal 11 Oktober 2016
- [12] Hwa, A K, Cho H Y, Rura Ni, Agung Suray, Florian Hutagalung. 2008. "*Ayo Belajar Bahasa Korea dengan Bahasa Korea Terpadu Untuk Orang Indonesia*". Korea Language Department, Korea Foundation. Secho-gu, Seoul.
- [13] Kristanto, Andri, 2004. "*Jaringan Syaraf Tiruan, Konsep Dasar, Algoritma dan Aplikasi*". Gava Media. Yogyakarta.
- [14] Sibagariang, S, 2016. "*Klasifikasi Citra Mammogram dengan Metode Ekstraksi Ciri Zoning Menggunakan Ssvm*". Universitas Sumatera Utara, Medan. Diakses dari <http://repository.usu.ac.id/xmlui/handle/123456789/57829> pada tanggal 5 Desember 2016.
- [15] Sinaga, D Y, 2013. "*Pengenalan Pola Plat Nomor Kendaraan Bermotor Dengan Menggunakan PCA dan Metode Propagasi Balik*". Universitas Sanata Dharma, Yogyakarta.