

# IDENTIFIKASI HURUF HIJAIYAH BERBASIS GLCM MENGGUNAKAN JARINGAN SYARAF TIRUAN BACKPROPAGATION

Katya Blinda Putri <sup>1)</sup>, Noveri Lysbetti Marpaung, <sup>2)</sup>

<sup>1)</sup>Mahasiswa Jurusan Teknik Informatika, <sup>2)</sup>Dosen Jurusan Teknik Informatika  
Laboratorium Teknik Elektro Universitas Riau  
Program Studi Teknik Informatika S1, Fakultas Teknik, Universitas Riau  
Kampus Bina Widya Km 12,5 Simpang Baru, Panam, Pekanbaru 28293

## ABSTRACT

*Hijaiyah Letter or also called the Arabic alphabet can be found in the Holy Book of Islam, the Qur'an. Every Muslim is obliged to read the Qur'an and recognize each of Hijaiyah Letters as its constituent letters in order to read and write in Arabic correctly. Humans are essentially intelligent because they are able to distinguish between one object and another through the object's characteristics or patterns. Certainly it is not difficult to identify a handwriting of Hijaiyah Letters even though written by different people. But it's hard if a machine is trying to identify it. The research aims to establish an Identification System of Hijaiyah Letters based on Gray Level Cooccurrence Matrix (GLCM) using the Backpropagation Artificial Neural Network to determine system accuracy results and produce sound output in according to how to read it. Image Data of Hijaiyah Letters obtained from hand-written scanning of five different people with total of Hijaiyah Letters as much as 600 data measuring by 300x300 pixels with extension. JPEG for each letter. GLCM-based systems as the extraction of texture features produce values of Energy, Contrast, Homogeneity, Entropy and Correlation, used as input for the training process of Backpropagation Artificial Neural Network. Image Data of Hijaiyah letters used for the training process amounted to 420 letters and 180 letters of testing process. The result of the Hijaiyah Letters Identification System identifies 30 letters from Alif to Ya. The system is able to identify Hijaiyah Letters with an accuracy rate of 96.11% with error 3.89%.*

**Keywords:** *Hijaiyah Letters, GLCM, Backpropagation Artificial Neural Network, Matlab R2018a.*

## 1. Pendahuluan

Bahasa Arab tersusun dari Huruf Hijaiyah atau disebut juga Huruf Arab yang dapat dijumpai dalam kitab suci agama Islam yaitu Al-Qur'an. Setiap muslim wajib bisa membaca Al-Qur'an serta mengenali setiap Huruf Hijaiyah sebagai penyusunnya agar dapat membaca dan menulis Bahasa Arab dengan benar. Seperti huruf alphabet lainnya dalam menyusun kata dan kalimat untuk berbahasa, Huruf Hijaiyah memiliki peran yang sama untuk Bahasa Arab.

Proses Identifikasi Huruf Hijaiyah yang paling sederhana dan langsung yaitu menggunakan penglihatan. Manusia pada dasarnya cerdas karena mampu membedakan antara satu objek dengan lainnya melalui ciri atau pola objek tersebut. Tentunya tidaklah sulit untuk mengidentifikasi sebuah tulisan tangan Huruf Hijaiyah walaupun ditulis oleh orang yang berbeda-beda. Namun hal itu sulit jika mesin yang berusaha untuk

mengidentifikasinya, dikarenakan miripnya bentuk penulisan Huruf Hijaiyah dan cara pencapaiannya.

Oleh karena itu solusi untuk hal tersebut adalah membuat sistem yang bernama *Identifikasi Huruf Hijaiyah* yang dapat mengenali dan mengidentifikasi Huruf Hijaiyah.

Sebelum melakukan identifikasi, diperlukan langkah untuk memproses citra dengan Pengolahan Citra Digital. Pengolahan citra digital bertujuan memperbaiki kualitas citra agar mudah diinterpretasi oleh manusia atau komputer. [2]

Hasil pengolahan citra digital dilakukan ekstraksi fitur berbasis *Gray Level Co-Occurrence Matrix* (GLCM). Parameter ciri dari citra Huruf Hijaiyah yang digunakan adalah tekstur. Tekstur pada citra merupakan frekuensi perubahan rona yang terjadi karena adanya keteraturan pola piksels tertentu yang muncul secara berulang-ulang sehingga

membentuk pola tertentu. Ekstraksi GLCM Huruf Hijaiyah menghasilkan fitur diantaranya yaitu Energi, Kontras, Homogenitas, Entropi dan Korelasi. [1] Variabel ini yang digunakan sebagai *input* untuk algoritma pengenalan pola menggunakan metode Jaringan Syaraf Tiruan (JST) *Backpropagation*.

JST adalah bentuk representasi buatan dari otak manusia. Istilah buatan digunakan karena dibuat berdasarkan fungsi syaraf biologis yang merupakan dasar unit sinyal dari sistem syaraf [3] Salah satu algoritma pembelajaran dalam JST adalah propagasi balik (*backpropagation*) yang merupakan algoritma pembelajaran terawasi di mana masukan dan keluarannya telah diketahui sebelumnya. Pasangan data ini dipakai untuk melatih bobot-bobot masukan untuk mencari keluaran yang akan dibandingkan dengan target awal sehingga meminimalisir tingkat kesalahan (*error*). Metode *backpropagation* merupakan metode yang sangat baik dalam menangani masalah pengenalan pola-pola kompleks. Beberapa contoh aplikasi yang melibatkan JST *Backpropagation* adalah pengoperasian data, peramalan (prediksi) dan pengenalan pola. [4] Proses identifikasi ini mengenali gambar tulisan Huruf Hijaiyah dengan gambar yang sudah ada di database sebagai data latih, sehingga hasilnya valid dengan data uji objek tulisan pada citra.

Berdasarkan permasalahan tersebut, maka dibangun suatu sistem Identifikasi Huruf Hijaiyah berbasis komputer untuk mengidentifikasi kumpulan citra Huruf Hijaiyah. Penelitian ini bertujuan untuk mengidentifikasi untuk mempelajari Huruf Hijaiyah dengan membangun sistem Identifikasi Huruf Hijaiyah berbasis GLCM menggunakan Jaringan Syaraf Tiruan *Backpropagation* untuk menentukan ketepatan hasil akurasi sistem.

## 2. Metodologi Penelitian

Peralatan, perangkat dan bahan yang digunakan adalah *Printer/Scanner*, Kertas HVS A4, Spidol Hitam, *Adobe Photoshop CC 2019 Portable* dan *Matlab R2018a*.

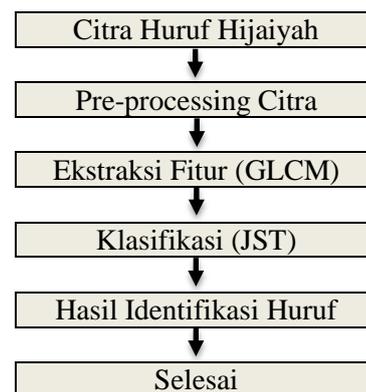
### 2.1 Pengumpulan Data Citra Huruf Hijaiyah

Data citra Huruf Hijaiyah diisi dengan tulisan tangan oleh lima orang responden yang telah diberi blangko penulisan Huruf Hijaiyah.

Pada blangko yang dibagikan, terdapat empat kolom kosong disamping gambar contoh Huruf Hijaiyah. Responden diberikan kesempatan menulis tulisan tangan Huruf Hijaiyah sebanyak 4x30 huruf. Blangko tersebut kemudian dipindai dengan scanner untuk mengubah data gambar Huruf Hijaiyah dari blangko menjadi citra digital berekstensi .JPG. Huruf Hijaiyah yang telah dipindai, kemudian dipotong (*crop*) menggunakan *Adobe Photoshop Portable* dengan ukuran 300x300 piksel untuk setiap hurufnya. Total keseluruhan Huruf Hijaiyah setelah dipotong diperoleh sebanyak 600 data citra Huruf Hijaiyah.

### 2.2 Perancangan dan Pembuatan Sistem

Ada beberapa langkah yang dilakukan untuk melakukan implementasi sistem yaitu dengan langkah awal mengumpulkan data citra Huruf Hijaiyah dan mengubahnya dalam bentuk citra digital. Skema perancangan Sistem Identifikasi Huruf Hijaiyah secara umum dilihat pada Gambar 1.



**Gambar 1.** Diagram Perancangan Sistem

Langkah pertama, data citra Huruf Hijaiyah yang telah dikumpulkan dilakukan *pre-processing* citra dengan mengubahnya ke dalam bentuk *grayscale*. Selanjutnya data citra Huruf Hijaiyah *grayscale* dilakukan ekstraksi fitur menggunakan GLCM. Nilai hasil ekstraksi fitur data citra Huruf Hijaiyah yaitu Energi, Kontras, Homogenitas, Entropi dan Korelasi. digunakan sebagai variabel *input* pada proses pelatihan JST *Backpropagation* untuk mendapatkan output hasil Identifikasi Huruf Hijaiyah. Variabel input JST dilihat pada Tabel 1.

**Tabel 1.** Keterangan Variabel Input

VARIABEL	KETERANGAN
X1	Nilai Fitur Energi

X2	Nilai Fitur Kontras
X3	Nilai Fitur Korelasi
X4	Nilai Fitur Homogenitas
X5	Nilai Fitur Entropi

Pada penelitian ini target *output* diklasifikasikan menjadi 3 kelas, seperti terlihat pada Tabel 2.

**Tabel 2.** Target Output

KLASIFIKASI	KETERANGAN
BENAR	Berhasil melakukan Identifikasi Huruf Hijaiyah
SALAH	Gagal melakukan Identifikasi Huruf Hijaiyah
HURUF TIDAK TERIDENTIFIKASI	Input Huruf bukan Huruf Hijaiyah

Pembagian penggunaan data citra Huruf Hijaiyah untuk proses pelatihan (*training*) dan pengujian (*testing*) pada JST Backpropagation memiliki perbandingan 70 data *training* dan 30 data *testing*. Data citra Huruf Hijaiyah yang digunakan untuk proses *training* berjumlah 420 huruf dengan satu jenis Huruf Hijaiyah memiliki sebanyak 14 citra latih. Sedangkan untuk proses *testing*, sebanyak 180 citra huruf dengan satu jenis Huruf Hijaiyah sebanyak 6 citra uji.

### 2.3 Pengujian Sistem

Sistem yang telah dibuat harus melalui pengujian untuk mengetahui hasil keakuratan

sistem apakah sesuai dengan yang diharapkan yaitu sistem berhasil mengidentifikasi Huruf Hijaiyah. Hasilnya dihitung dengan persamaan berikut:

$$\text{Akurasi} = \frac{\text{Huruf yang diuji benar}}{\text{Total huruf yang diuji}} \times 100\% \quad (1)$$

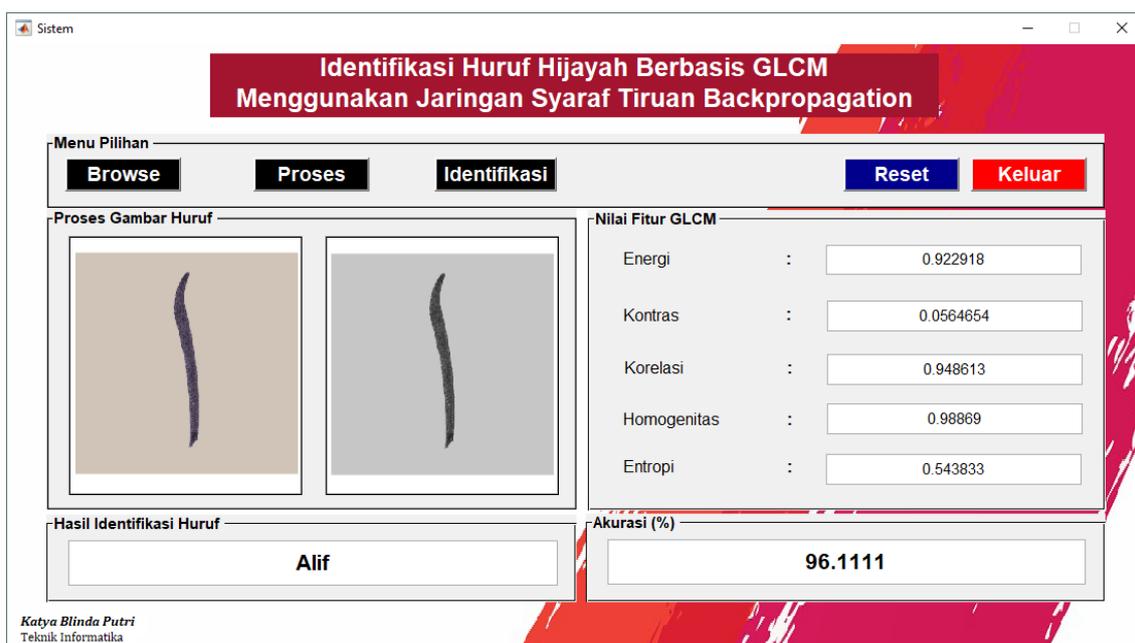
$$\text{Error} = \frac{\text{Huruf yang diuji salah}}{\text{Total huruf yang diuji}} \times 100\% \quad (2)$$

## 3. Hasil dan Pembahasan

### 3.1 Tampilan Sistem Identifikasi Huruf Hijaiyah

Pada Gambar 3. menampilkan hasil identifikasi Huruf Hijaiyah dengan antarmuka sistem memiliki tampilan nama sistem, dan beberapa panel diantaranya panel menu pilihan yang terdiri dari *button Browse*, *Proses*, *Identifikasi*, *Reset* dan *Keluar*, panel *Proses Gambar Huruf*, panel *Nilai Fitur GLCM*, panel *Hasil Identifikasi Huruf* dan panel *Akurasi (%)*. Sebelum dilakukan identifikasi Huruf Hijaiyah, pada Panel *Akurasi (%)* hasil ketepatan pengujian data citra Huruf Hijaiyah ditampilkan secara default sebesar 96.1111% berdasarkan hasil pelatihan JST *Backpropagation*.

Saat klik *button* identifikasi, sistem akan mengeluarkan output suara untuk cara membaca Huruf Hijaiyah dan tulisan hasil identifikasinya. Terlihat pada Gambar 3. huruf yang diidentifikasi adalah gambar Huruf Alif menghasilkan output tulisan Alif dan suara pelafalannya.



**Gambar 2.** Tampilan Identifikasi Huruf Hijaiyah Alif

**3.2 Pengujian dan Perhitungan Akurasi dan Error**

Pembagian penggunaan data citra Huruf Hijaiyah untuk proses pelatihan (training) dan pengujian (testing) pada JST *Backpropagation* memiliki perbandingan 70 data *training* dan 30 data *testing*. Data citra Huruf Hijaiyah yang digunakan untuk proses training berjumlah 420 citra huruf dengan satu jenis Huruf Hijaiyah memiliki sebanyak 14 citra latih. Sedangkan untuk proses *testing*, sebanyak 180 citra huruf dengan satu jenis Huruf Hijaiyah sebanyak 6 citra uji.

Pengujian data uji yang digunakan dalam penelitian ini menggunakan konfigurasi parameter JST yaitu *epoch* 1000, *goal* 0.001, *neuron hidden layer* pertama 100 *neuron*, *neuron hidden layer* kedua 95 *neuron*.

Berdasarkan jumlah benar data uji yang sesuai dengan *output* citra Huruf Hijaiyah yang telah diuji, diperoleh persentase akurasi sistem sebagai berikut:

$$Akurasi = \left[ \frac{173}{180} \right] \times 100\% = 96.11\% \tag{2}$$

Berdasarkan jumlah data uji yang tidak sesuai dengan *output* citra Huruf Hijaiyah, diperoleh persentase *error* sistem sebagai berikut:

$$Error = \left[ \frac{7}{180} \right] \times 100\% = 3.89\% \tag{3}$$

**Tabel 3.** Hasil *Output* Huruf Hijaiyah dan Persentase Pengujian

Input	Output Huruf Hijaiyah						Dikenali	Tidak Dikenali	Rata-rata
	1	2	3	4	5	6			
Alif	Alif	Alif	Alif	Alif	Alif	Alif	6	0	100
Ba'	Ba'	Ba'	Ba'	Ba'	Ba'	Ba'	6	0	100
Ta'	Ta'	Ta'	Ta'	Ta'	Ta'	Ta'	6	0	100
Tsa'	Tsa'	Tsa'	Tsa'	Tsa'	Tsa'	Tsa'	6	0	100
Jim	Jim	Jim	Jim	Jim	Jim	Jim	6	0	100
Ha'	Ha'	Ha'	Ha'	Ha'	Ha'	Ha'	6	0	100
Kha	Kha	Kha	Kha	Lam Alif	Kha	Kha	5	1	83.33
Dal	Dal	Dal	Dal	Dal	Dal	Dal	6	0	100
Dzal	Dzal	Dzal	Dzal	Dzal	Dzal	Dzal	6	0	100
Ra'	Ra'	Dho	Ra'	Syin	Ra'	Ra'	4	2	66.67
Zai	Zai	Zai	Zai	Zai	Zai	Zai	6	0	100
Sin	Syin	Sin	Sin	Sin	Sin	Sin	5	1	83.33
Syin	Syin	Syin	Syin	Syin	Syin	Shod	5	1	83.33
Shod	Shod	Shod	Shod	Shod	Shod	Shod	6	0	100
Dhod	Dhod	Dhod	Dhod	Dhod	Dhod	Dhod	6	0	100
Tha'	Tha'	Tha'	Tha'	Tha'	Tha'	Tha'	6	0	100
Za'	Za'	Za'	Za'	Za'	Za'	Za'	6	0	100
'Ain	Ain	Ain	Ain	Ain	Ain	Ain	6	0	100
Ghain	Ghain	Ghain	Ghain	Ghain	Ghain	Ghain	6	0	100
Fa'	Fa'	Fa'	Fa'	Fa'	Fa'	Fa'	6	0	100
Qaf	Qaf	Qaf	Qaf	Qaf	Qaf	Qaf	6	0	100
Kaf	Kaf	Kaf	Qaf	Kaf	Kaf	Kaf	5	1	83.33

<b>Lam</b>	Lam	Lam	Lam	Lam	Lam	Lam	6	0	100
<b>Mim</b>	Mim	Mim	Mim	Mim	Mim	Mim	6	0	100
<b>Nun</b>	Nun	Nun	Nun	Nun	Nun	Nun	6	0	100
<b>Waw</b>	Waw	Waw	Waw	Sin	Waw	Waw	5	1	83.33
<b>Ha</b>	Ha	Ha	Ha	Ha	Ha	Ha	6	0	100
<b>Hamzah</b>	Hamzah	Hamzah	Hamzah	Hamzah	Hamzah	Hamzah	6	0	100
<b>Lam Alif</b>	Lam Alif	6	0	100					
<b>Ya</b>	Ya	Ya	Ya	Ya	Ya	Ya	6	0	100
<b>Jumlah</b>							<b>173</b>	<b>7</b>	
<b>Rata-rata Akurasi (%)</b>							<b>96.11</b>		

Berdasarkan hasil percobaan yang dilakukan, tingkat akurasi sistem dapat mengenali Huruf Hijaiyah sebesar 96.11% dengan *error* 3.89%.

#### 4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan untuk Sistem Identifikasi Huruf Hijaiyah dapat ditarik kesimpulan sebagai berikut:

1. Penelitian ini menggunakan Data Citra Huruf Hijaiyah.
2. Variabel masukan yang digunakan berupa hasil ekstraksi fitur GLCM diantaranya Energi, Kontras, Korelasi, Homogenitas dan Entropi
3. GLCM sebagai ekstraksi fitur tekstur dan Metode Klasifikasi Jaringan Syaraf Tiruan *Backpropagation* dapat digunakan untuk Identifikasi Huruf Hijaiyah.
4. Berdasarkan hasil penelitian yang telah dilakukan, akurasi yang didapatkan menggunakan Metode JST *Backpropagation* dengan *epoch* 1000,

*goal* 0.001, *Neuron Hidden Layer* pertama 100 *neuron*, *Neuron Hidden Layer* kedua 95 *neuron* mencapai akurasi sebesar 96,11% dan *error* 3.89% sehingga sistem ini layak untuk digunakan

5. *Output* dari tulisan Huruf Hijaiyah dapat ditampilkan jenis huruf yang dimaksud dan *output* suara sesuai dengan cara membacanya.

#### DAFTAR PUSTAKA

- [1] M. Haralick, R. 1973. *Textural Features for Image Classification*. IEEE. USA.
- [2] Munir, R. 2004. *Pengolahan Citra Digital*. (<http://informatika.stei.itb.ac.id/~rinaldi.munir/Buku/Pengolahan%20Citra%20Digital/>) Diunduh tanggal 16 Mei 2018.
- [3] Putra, D. 2010. *Pengolahan Citra Digital*. Penerbit Andi: Yogyakarta.
- [4] Puspitaningrum, D. 2006. *Pengantar Jaringan Saraf Tiruan*. Penerbit Andi: Yogyakarta.