

## PENERAPAN KOHONEN NEURAL NETWORK PADA PENGENALAN HURUF TULISAN TANGAN

Akhmad Ridhani<sup>1</sup>, Fatma Indriani<sup>2</sup>, Dodon T. Nugrahadi<sup>3</sup>

Universitas Lambung Mangkurat<sup>1,2,3</sup>

akhridhani@gmail.com<sup>1</sup>

f.indriani@ulm.ac.id<sup>2</sup>

dodonturianto@ulm.ac.id<sup>3</sup>

### ABSTRACT

*In this paper, Kohonen Neural Network is used for OCR. Based on the results of research conducted, this method can perform character identification of handwritten text with an accuracy of 86.43% using training and testing from 10 respondents handwriting data. We also find the letters B, N, and R has low identification accuracy.*

**Keywords:** *artificial neural network, Kohonen, OCR, handwriting*

### ABSTRAK

Pada makalah ini kami memaparkan penerapan metode Kohonen Neural Network untuk melakukan pengenalan huruf tulisan tangan. Berdasarkan hasil penelitian yang dilakukan pada pengenalan huruf tulisan tangan dengan metode Kohonen Neural Network, didapatkan bahwa metode ini dapat melakukan analisis dengan tingkat keakurasian yang cukup tinggi, yaitu 86.43%. Hal ini dibuktikan dari pengujian tulisan tangan pada 10 responden. Pengujian juga dilakukan dengan jumlah data training yang beragam dan diketahui bahwa jumlah data training (responden) berpengaruh pada tingkat akurasi pengenalan. Selain itu, didapatkan bahwa huruf tertentu seperti B, N, dan R memiliki tingkat akurasi pengenalan yang rendah.

**Kata Kunci:** jaringan syaraf tiruan, Kohonen, OCR, tulisan tangan

### PENDAHULUAN

Perkembangan teknologi yang dulunya hanya pada bidang tertentu sekarang sudah hampir dimanfaatkan untuk segala bidang. Sehingga memicu perpindahan dokumen berupa cetak menjadi bentuk dokumen digital untuk mempermudah penyimpanan dan keamanan data yang terjaga. Pengenalan pola merupakan suatu ilmu yang digunakan untuk mengklasifikasi sesuatu berdasarkan pengukuran kuantitatif ciri atau sifat utama dari objek. *Optical Character Recognition (OCR)* merupakan salah satu area studi dalam bidang pengenalan pola.

*Optical Character Recognition (OCR)* dapat digunakan untuk solusi yang efektif dalam proses pengubahan dokumen cetak ke dokumen digital. Permasalahan yang kemudian muncul dalam proses pengenalan pola adalah teknik pengenalan pola dapat mengenali berbagai jenis huruf berupa cetak maupun tulisan tangan yang beraneka ragam.

Metode *Kohonen Neural Network* mempelajari distribusi seperangkat pola tanpa informasi kelas. Ide dasar teknik ini dipahami dari otak manusia menyimpan gambar/pola yang telah dikenali melalui mata, maka otak manusia mampu

mengungkapkan gambar/pola kembali. Oleh karena itu, model ini sering digunakan dalam banyak aplikasi karena *kohonen neural network* merupakan jaringan saraf yang tanpa pengawasan yang belajar secara kompetitif dalam proses adaptif (Chugh dan Jain, 2017). Penelitian ini fokus pada tingkat akurasi pengenalan huruf pada tulisan tangan dengan metode *kohonen neural network*. Huruf yang digunakan pada penelitian ini untuk dilakukan pengenalan adalah huruf abjad kapital.

Berdasarkan penelitian terdahulu, jaringan syaraf tiruan cukup baik digunakan untuk pengenalan pola. Misal Chugh & Jain (2017) mendesain sistem offline verifikasi tanda tangan menggunakan metode kohonen neural network. Penelitian ini memanfaatkan kekuatan pembelajaran kompetitif dari kohonen neural network untuk mengukur intra-variability tanda tangan seseorang. Hasil penelitian ini membuktikan bahwa sistem yang dirancang kuat untuk pemalsuan umum dan acak dengan FAR (False Acceptance Rate) and FRR (False Rejection Rate) untuk sampel asil masing-masing 2.8% dan 5%.

Pada penelitian lain, Kasaudhan et al. (2016) mengusulkan sebuah algoritma baru untuk mengekstrak teks/karakter dari gambar yang dipindai menggunakan jaringan syaraf tiruan dengan menggunakan metode untuk menggabungkan jaringan yang dilatih secara independen untuk mencapai kinerja yang lebih tinggi. Hasil penelitian ini diketahui sistem yang diusulkan lebih baik dari pada sistem yang ada dan mencoba untuk meningkatkan efisiensi dan ketepatan mengenali karakter dari citra yang dipindai. Chaudhary et al. (2017) melakukan pendekatan OCR pada Grid Infrastructure menggunakan Kohonen Neural Network. Hasil penelitian ini menyatakan sistem yang diusulkan pada penelitian ini dapat secara akurat mengenali karakter tulisan tangan dan dokumen cetak.

Penelitian Munggaran et al. (2014) bertujuan untuk mengetahui algoritma pengenalan tulisan tangan yang dapat menerima masukan tulisan tangan dan mengenali karakter tulisan tangan yang langsung diinput di komputer menggunakan Kohonen Neural Network dengan pixel karakter. Namun pixel yang digunakan hanya 5x5. Sedangkan Nirve & Shinde (2013) merancang aplikasi reorganisasi karakter dengan kinerja tinggi untuk bahasa India yaitu Hindi berdasarkan Devanagari script, hasil penelitian menyatakan Kohonen Neural Network dengan peta pengorganisasian sendiri pasti akan memberikan hasil yang lebih baik daripada sistem jaringan syaraf lainnya dan juga ada kenaikan tingkat persentase pengenalan.

Pada makalah ini, kami memberikan kontribusi berupa analisis pengaruh jumlah data training terhadap pengenalan huruf, serta analisis keberhasilan pengenalan berdasarkan huruf per huruf.

## METODE PENELITIAN

Metode penelitian yang dilakukan pada penelitian ini yaitu : pengambilan data, *input* data, *pre-processing*, ekstraksi fitur, klasifikasi, dan pengujian. Pada tahap pengujian dilakukan 10 kali percobaan dengan jumlah data *training* yang beragam. Data yang digunakan pada penelitian ini adalah data tulisan tangan dari 10 responden. Adapun prosedur kerja yang dilakukan dalam penelitian ini adalah sebagai berikut :

### 1. Pengambilan Data

Pengambilan data tulisan tangan pada 10 orang responden, data yang diambil berupa gambar tulisan tangan masing-masing huruf secara terpisah yaitu huruf abjad dari 'A' sampai dengan 'Z' untuk data *training* dan berupa kalimat untuk data pengujian.

## 2. Input Data

Pada tahap ini ditentukan jenis data yang digunakan yaitu gambar dengan format JPEG. Gambar yang sudah diambil dari tahap pengambilan data diinputkan untuk dilakukan tahap selanjutnya pada tahap *pre-processing*.

## 3. Pre-Processing

Pada tahap *Pre-processing* terdapat beberapa tahapan yang dilakukan pada hasil dari *input* data, yaitu pengubahan gambar RGB menjadi gambar *grayscale*, gambar *grayscale* diubah menjadi gambar biner, segmentasi, dan normalisasi.

## 4. Ekstraksi Fitur

Tahapan ini dilakukan ekstraksi fitur untuk mendapatkan nilai vector dengan mengubah gambar hasil *pre-processing* menjadi nilai 0 untuk warna putih dan 1 untuk warna hitam. Hasil dari ekstraksi fitur digunakan pada tahap klasifikasi .

## 5. Klasifikasi

Pada tahap klasifikasi ada dua proses yang dapat dilakukan yaitu *training* dan pengenalan data tulisan tangan dengan metode *Kohonen Neural Network*. Data huruf yang sudah diambil pada tahap pengambilan data digunakan pada proses *training* dan data kalimat digunakan pada tahap pengenalan. Algoritma *Kohonen Neural Network* sebagai berikut.

- a. Langkah 0 : Inisialisasi bobot  $W_{ij}$  diambil dari rata-rata tiap vector huruf yang digunakan untuk *training* dan parameter learning rate ( $\alpha$ ). Pada penelitian ini ditentukan learning rate ( $\alpha$ ) sebesar 0,5.
- b. Langkah 1 : Saat kondisi stop yaitu nilai epochs sama dengan nilai epochs maksimum, nilai alpha lebih kecil dari nilai alpha minimum, dan nilai *error* lebih kecil dari nilai *error* maksimum bernilai salah, maka lakukan langkah 2-8.
- c. Langkah 2 : Untuk tiap vektor *input*  $x$ , lakukan langkah 3-5
- d. Langkah 3 : Untuk tiap  $j$ , lakukan perhitungan dengan rumus sebagai berikut.

$$D(j) = \sum_i (w_{ij} - x_i)^2 \quad \dots(1)$$

- e. Langkah 4 : Temukan index  $J$  yang memiliki nilai  $D(J)$  terkecil
- f. Langkah 5 : Perbaharui semua bobot untuk tiap unit  $j$  dengan rumus sebagai berikut.

$$w_{ij}(\text{baru}) = w_{ij}(\text{lama}) + \alpha [x_i - w_{ij}(\text{lama})] \quad \dots(2)$$

- g. Langkah 6 : Perbaharui learning rate ( $\alpha$ ) dengan mengalikan learning rate yang lama dengan 0,5 untuk mengurangi nilai learning rate ( $\alpha$ ) dengan rumus sebagai berikut.

$$\text{learning rate } (\alpha) = \text{learning rate } (\alpha) * \text{pengurangan learning rate} \quad \dots(3)$$

- h. Langkah 7 : Kurangi radius tetangga pada waktu yang spesifik  
i. Langkah 8 : Cek kondisi stop

### Pengujian

Proses pengujian pada aplikasi OCR dengan metode *Kohonen Neural Network* dengan melakukan analisis data pengujian. Pengujian dilakukan pada data gambar tulisan tangan 10 orang sebanyak 10 macam pengujian untuk tiap data gambar tulisan tangan. Tingkat akurasi didapatkan dengan membandingkan jumlah kasus yang berhasil diidentifikasi dengan jumlah kasus secara keseluruhan.

## HASIL DAN PEMBAHASAN

### Pengambilan Data

Pada tahapan ini diambil data *training* dan pengujian contoh huruf tulisan tangan dari 10 responden. Untuk data training responden menulis huruf abjad kapital dari “A” sampai “Z” dengan latar putih dan huruf berwarna hitam. Sehingga didapatkan 260 data huruf tulisan tangan yang digunakan pada tahap *training*. Data pengujian diambil dari responden yang sama berupa kata yang diambil dari kalimat pangram yaitu kalimat yang mengandung seluruh huruf. Kalimat yang dipakai adalah “*Wibowo naik vespa dari roxy ke islamic center jakarta ikut program tahfizh Alquran*”, detail pembagian kata dapat dilihat pada tabel 1 dan detail jumlah huruf pada kalimat tersebut dapat dilihat di tabel 2.

**Tabel 1.** Pembagian kalimat pengujian menjadi kata

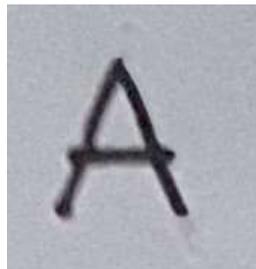
Kata Pengujian	
WIBOWO	CENTER
NAIK	JAKARTA
VESPA	IKUT
DARI	PROGRAM
ROXY	TAHFIZH
KE	ALQURAN
ISLAMIC	

**Tabel 2.** Detail jumlah huruf pada kalimat pengujian

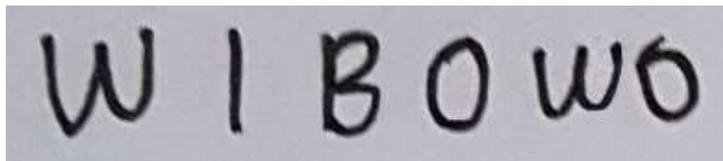
Huruf			
A	11	N	3
B	1	O	4
C	2	P	2
D	1	Q	1
E	4	R	7
F	1	S	2
G	1	T	4
H	2	U	2
I	7	V	1
J	1	W	2
K	4	X	1
L	2	Y	1
M	2	Z	1
TOTAL	70		

### Input Data

Data huruf tulisan tangan yang telah didapat, diambil gambar dengan menggunakan kamera yang memiliki ukuran 12MP, sehingga didapatkan 260 data berupa gambar RGB huruf tulisan tangan dan 10 kalimat tulisan tangan dengan format JPEG yang memiliki ukuran *pixel* beragam. Data yang didapat diinputkan secara bergiliran untuk dilakukan proses selanjutnya. Contoh gambar *input* dapat dilihat pada gambar 1 berupa huruf tulisan tangan untuk data *training* dan pada gambar 2 berupa kata tulisan tangan untuk pengenalan.



**Gambar 1.** Contoh *input* data gambar huruf tulisan tangan

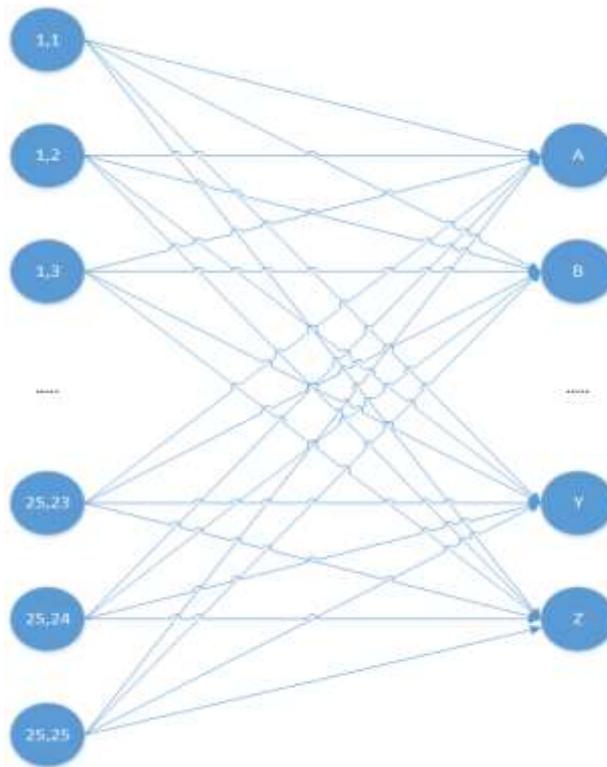


**Gambar 2.** Contoh *input* data gambar huruf tulisan tangan

### Pre-Processing

Pada tahapan *pre-processing* terdapat beberapa tahapan yang dilakukan yaitu perubahan gambar RGB menjadi gambar *grayscale*, Gambar *grayscale* diubah menjadi gambar *binary*, Segmentasi, dan Normalisasi. Pada Proses Pada Segmentasi digunakan metode *Otsu Thresholding* [2]. Normalisasi terdapat dua tahapan yaitu *scalling* dan *thinning*. Proses *scalling* mengubah gambar menjadi





**Gambar 5.** Kohonen Neural Network untuk pengenalan huruf tulisan tangan

**Pengujian**

Pengujian dilakukan untuk mengetahui tingkat akurasi pengenalan huruf tulisan tangan dengan menggunakan metode *kohonen neural network*. Pengujian dilakukan dengan mengujikan aplikasi terhadap 10 responden yang sama pada proses pengumpulan data. Pengujian dilakukan dengan jumlah *training* yang beragam dari data *training* 1 responden sampai dengan 10 responden secara berurutan. Berdasarkan hasil pengujian yang dilakukan sebanyak 10 kali untuk masing-masing jumlah data *training* akan dihitung tingkat akurasi untuk setiap huruf, akurasi rata-rata setiap huruf, dan setiap responden. Hasil pengujian dapat dilihat pada tabel berikut dengan huruf yang tidak sesuai diberi huruf tebal.

Hasil pengujian penerapan metode *kohonen neural network* pada pengenalan huruf tulisan tangan dari jumlah data *training* 1 responden sampai dengan data *training* 10 responden ada beberapa huruf yang tingkat akurasinya fluktuatif dan ada yang semakin meningkat akurasinya dengan bertambahnya data *training* yang digunakan.

**Tabel 3.** Tingkat Akurasi Setiap Huruf Dengan Data *Training* 1 Responden

HURUF	JUMLAH HURUF	JUMLAH BENAR	PERSENTASE BENAR
A	110	98	89.09%
B	10	0	0.00%
C	20	0	0.00%
D	10	5	50.00%

E	40	28	70.00%
F	10	7	70.00%
G	10	1	10.00%
H	20	11	55.00%
I	70	17	24.29%
J	10	6	60.00%
K	40	27	67.50%
L	20	18	90.00%
M	20	7	35.00%
N	30	18	60.00%
O	40	32	80.00%
P	20	13	65.00%
Q	10	5	50.00%
R	70	19	27.14%
S	20	10	50.00%
T	40	40	100.00%
U	20	15	75.00%
V	10	9	90.00%
W	20	5	25.00%
X	10	7	70.00%
Y	10	9	90.00%
Z	10	7	70.00%
<b>TOTAL BENAR</b>			<b>414</b>
<b>RATA-RATA PERSENTASE</b>			<b>56.65%</b>

**Tabel 4.** Tingkat Akurasi Setiap Huruf Dengan Data *Training* 5 Responden

HURUF	JUMLAH HURUF	JUMLAH BENAR	PERSENTASE BENAR
A	110	98	89.09%
B	10	3	30.00%
C	20	17	85.00%
D	10	8	80.00%
E	40	36	90.00%
F	10	10	100.00%
G	10	6	60.00%
H	20	19	95.00%
I	70	58	82.86%
J	10	10	100.00%
K	40	32	80.00%
L	20	19	95.00%
M	20	16	80.00%
N	30	17	56.67%
O	40	26	65.00%
P	20	12	60.00%
Q	10	5	50.00%
R	70	37	52.86%
S	20	20	100.00%
T	40	40	100.00%
U	20	18	90.00%

V	10	10	100.00%
W	20	20	100.00%
X	10	9	90.00%
Y	10	9	90.00%
Z	10	8	80.00%
<b>TOTAL BENAR</b>		<b>563</b>	
<b>RATA-RATA PERSENTASE</b>		<b>80.83%</b>	

**Tabel 5.** Tingkat Akurasi Setiap Huruf Dengan Data *Training* 10 Responden

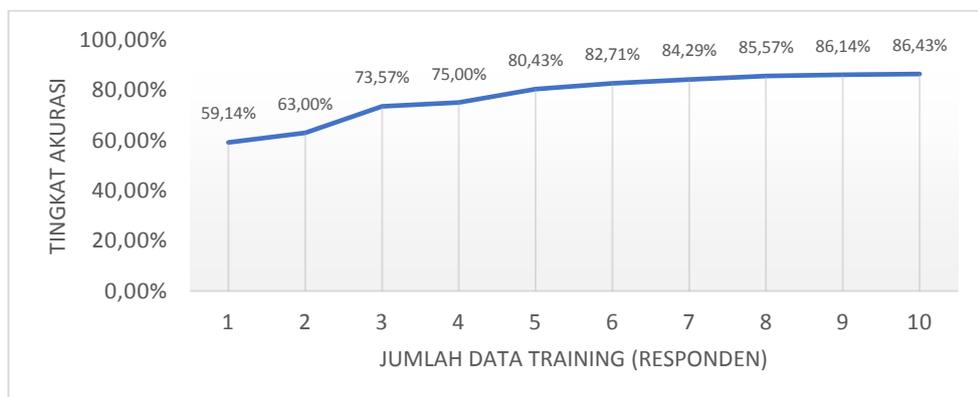
<b>HURUF</b>	<b>JUMLAH HURUF</b>	<b>JUMLAH BENAR</b>	<b>PERSENTASE BENAR</b>
A	110	102	92.73%
B	10	6	60.00%
C	20	20	100.00%
D	10	9	90.00%
E	40	35	87.50%
F	10	10	100.00%
G	10	9	90.00%
H	20	19	95.00%
I	70	67	95.71%
J	10	10	100.00%
K	40	35	87.50%
L	20	19	95.00%
M	20	16	80.00%
N	30	18	60.00%
O	40	30	75.00%
P	20	16	80.00%
Q	10	8	80.00%
R	70	41	58.57%
S	20	18	90.00%
T	40	40	100.00%
U	20	18	90.00%
V	10	10	100.00%
W	20	20	100.00%
X	10	10	100.00%
Y	10	10	100.00%
Z	10	9	90.00%
<b>TOTAL BENAR</b>		<b>605</b>	
<b>RATA-RATA PERSENTASE</b>		<b>88.35%</b>	

Tingkat akurasi pengenalan huruf yang didapatkan dari 10 responden dengan data *training* 10 responden adalah sebesar 88.35%, dengan tingkat akurasi pengenalan huruf terendah 58.57% dan tertinggi 100%. Dari hasil pengujian juga diketahui ada tiga huruf yang memiliki tingkat akurasi rendah yaitu huruf B dengan akurasi 60%, huruf N dengan akurasi 60%, dan huruf R dengan akurasi 58.57%. Rendahnya pengenalan tiga huruf ini ada kemungkinan disebabkan karena adanya kemiripan pola dengan huruf lain sehingga perlu dilakukan pengujian lebih lanjut.

Tingkat akurasi pengenalan huruf tulisan tangan dengan metode *kohonen neural network* pada tiap responden semakin meningkat dengan bertambahnya data *training*. Akurasi pengenalan yang didapatkan dari 10 responden dengan data *training* 10 responden adalah sebesar 86.43%, dengan tingkat akurasi terendah 59.14% dan tertinggi 91.43%. Pada Tabel 4 menunjukkan ada kenaikan total tingkat akurasi dengan semakin bertambahnya jumlah data *training*. Total tingkat akurasi terendah adalah 59.14% dan total tingkat akurasi tertinggi adalah 86.43% dengan kenaikan tingkat akurasi berkisar anantara 0.29% sampai 10.57%. Kenaikan tingkat akurasi tertinggi yaitu 10.57% terjadi pada jumlah data *training* 2 responden ke 3 responden. Tingkat akurasi terendah sudah cukup bagus karena dengan penggunaan data *training* yang sedikit sudah dapat melakukan pengenalan yang cukup bagus terhadap data tulisan yang belum pernah dilakukan *training*.

**Tabel 6.** Persentase Tingkat Akurasi Data *Training* Responden

JUMLAH DATA TRAINING (RESPONDEN)	JUMLAH HURUF PENGUJIAN	JUMLAH BENAR	PERSENTASE BENAR
1	700	414	59.14%
2	700	441	63.00%
3	700	515	73.57%
4	700	525	75.00%
5	700	563	80.43%
6	700	579	82.71%
7	700	590	84.29%
8	700	599	85.57%
9	700	603	86.14%
10	700	605	86.43%



**Gambar 6.** Grafik Persentase Tingkat Akurasi Data *Training* Responden

Pada gambar 6 menunjukkan grafik tingkat akurasi pengenalan huruf tulisan tangan dengan metode *kohonen neural network* dengan jumlah data *training* yang berbeda. Dari grafik tersebut dapat dilihat bahwa dari 1 sampai 5 data *training* tingkat akurasi meningkat cukup banyak, sedangkan pada data *training* selanjutnya peningkatan tingkat akurasi cukup sedikit. Dari data diatas dapat disimpulkan

bahwa jumlah data *training* sangat berpengaruh terhadap tingkat akurasi pengenalan huruf.

## KESIMPULAN

Pada penelitian ini, metode *Kohonen Neural Network* memiliki tingkat akurasi sebesar 86.43% dengan jumlah data *training* 10 responden dalam menangani kasus *Optical Character Recognition* pada tulisan. Jumlah data *training* yang digunakan mempengaruhi tingkat akurasi pengenalan huruf tulisan tangan dengan metode *Kohonen Neural Network*. Pengenalan huruf B, N, dan R memiliki tingkat akurasi yang rendah dibanding huruf yang lain.

## DAFTAR PUSTAKA

- Chaudhary, Shantanu., Swarnima Garg & Sathyaraj R. 2017. An Approach for Optical Character Recognition on Grid Infrastructure Using Kohonen Neural Network. *International Journal of Advanced Research in Computer Science*. Vol. 8 No. 3:485-490.
- Chugh, Aarti & Charu Jain. 2017. "Kohonen Networks for Offline Signature Verification". *International Journal of Recent Research Aspects* Vol. 4, Issue 2, pp. 18-23.
- Dash, Jyotiprava & Nilamani Bhoi. 2018. "Retinal Blood Vessel Segmentation Using Otsu Thresholding With Principal Component Analysis". *IEEE 2nd International Conference on Inventive Systems and Control (ICISC)*, 2018.
- Guo, Z. & Hall, R.W. 1989. "Parallel thinning with two sub-iteration algorithms". *Communications of the ACM* 32(3): 359–373.
- Kasaudhan, S., S. Saxena, V. Upadhayay, & L. Kumar. 2016. Optical Character Recognition using Grid Infrastructure. *IJSRD - International Journal for Scientific Research & Development*. Vol 4:1043-1045.
- Munggaran, L. C., S. Widodo, Cipta A.M, & Nuryuliani. 2014. "Handwritten Pattern Recognition Using Kohonen Neural Network Based on Pixel Character." (*IJACSA*) *International Journal of Advanced Computer Science and Applications*. Vol. 5 No. 11.
- Nirve, Sheetal A. & U. B. Shinde, 2013. "Hindi Charecter Recognition Using Kohonen Network." *International Journal of Scientific & Engineering Research* Vol. 4:1786-1790.
- Priya, A., S. Mishra, S. Raj, S. Mandal, & S. Datta. 2017. "Online and Offline Character Recognition: A Survey". *IEEE International Conference on Communication and Signal Processing*, 2016.
- Shatil, Adnan Md. Shoeb & Mumit Khan. 2006. "Bangla Optical Character Recognition Using Kohonen Network". *Computer Science and Engineering*, BRAC University.

Utami et al., A. E., O. D. Nurhayati & K. T. Martono. 2016. “Aplikasi Penerjemah Bahasa Inggris – Indonesia dengan Optical Character Recognition Berbasis Android”. *Jurnal Teknologi dan Sistem Komputer*. Vol.4:167-177.