

KLASIFIKASI NAMA OBAT TULISAN TANGAN DOKTER DENGAN METODE GLCM DAN BACKPROPAGATION NEURAL NETWORK

Arrahman¹, Purwanto², Pulung Nurtantio³

¹²³Pasca Sarjana Teknik Informatika Universitas Dian Nuswantoro

ABSTRACT

Pharmaceutical personnel in the work in general is always associated with the reading prescription, where the required accuracy, speed, and accuracy in reading prescription to avoid medication errors. This research show how to classify the doctor's handwriting drug name. Research conducted by the image processing prescription taken by scanner. Then the image manually cropped to take 200 drug names. Refining the drug name image has done twice with median filter and wiener filter, then dilation and erosion, feature extraction with GLCM (Grey-Level Co-occurrence matrix) methods to obtain data sets that will be classified by the software RapidMiner. From the test we find that Backpropagation Neural Network had more accurate than Naive Bayes and C 4.5.

Keywords: Drug Name Classification, GLCM, Backpropagation Neural Network, Naïve Bayes, C 4.5

1. PENDAHULUAN

Pada dasarnya resep dokter kurang dapat dibaca oleh tenaga kefarmasian lama atau tenaga kefarmasin baru dan mahasiswa farmasi disebabkan oleh jarang nya membaca resep dokter, kurangnya dalam perbendaharaan nama-nama obat dan obat yang diresepkan jarang ditulis oleh dokter. Tulisan tangan dokter yang kurang dapat dibaca sangat menyulitkan sehingga berpotensi menimbulkan kesalahan interpretasi terutama pada nama obat, dosis, aturan pakai dan cara pemberian yang selanjutnya dapat menyebabkan kesalahan pengobatan [1]. Tulisan di resep yang tidak dapat dibaca, tulisan tangan yang buruk disebut sebagai sumber primer dalam 15% insiden *medication error*. *Medication error* merupakan insiden keselamatan yang melibatkan obat, yang dapat terjadi di semua sistem pelayanan kesehatan. *Medication error* yang serius dapat membahayakan pasien dan memaparkan tenaga kesehatan pada tuntutan kriminal [2].

Pengenalan tulisan tangan merupakan kemampuan komputer untuk menerima dan menterjemahkan input yang berupa tulisan tangan dari sumber seperti dokumen digital, layar sentuh, kamera digital dan scanner. Pengenalan tulisan tangan dapat dibagi menjadi dua tipe yaitu pengenalan tulisan tangan secara off line dan pengenalan tulisan tangan secara on line. Pengenalan tulisan tangan secara off line merupakan pengambilan image berasal dari scanner atau kamera digital sedangkan pengenalan tulisan tangan on line adalah pengambilan image yang berasal dari gerakan pena pada layar sentuh.

Neural Network merupakan metode yang banyak digunakan oleh peneliti pengenalan tulisan tangan yang memberikan hasil pengenalan dengan akurasi yang tinggi dan merupakan metode yang banyak digunakan pada bidang-bidang yang melibatkan pemecahan masalah pengelompokan dan pengenalan pola dengan menghasilkan keakuratan yang tinggi [3]. Dalam Pengenalan tulisan tangan, *Neural Network* banyak diterapkan dan terbukti dapat memberikan akurasi pengenalan yang tinggi [4][5][6][7], dan algoritma yang banyak digunakan adalah algoritma *Backpropagation Neural Network*.

Penelitian ini bertujuan untuk meningkatkan akurasi pengenalan tulisan tangan dengan menerapkan metode GCLM dan algoritma *Backpropagation Neural Network* khususnya untuk klasifikasi nama obat tulisan tangan dokter.

2. TINJAUAN PUSTAKA

2.1. Filtering

2.1.1 Median Filter

Median filter adalah teknik pemrosesan sinyal nonlinier yang dikembangkan oleh Turkey yang berguna untuk menghilangkan *noise* dalam citra [8]. Median dari sekumpulan nilai adalah nilai yang berada di tengah dari sekumpulan nilai yang telah diurutkan. Pemrosesan median filter pada sebuah titik dalam citra yang pertama dilakukan adalah mengurutkan nilai piksel dalam tetangga (termasuk titik pusatnya), menentukan median, dan memasukkan nilai hasilnya pada piksel yang berkorespondensi (titik pusat tetangga) dalam citra yang difilter [9][10].

2.1.2 Wiener Filter

Wiener Filter merupakan jenis filter adaptif yang merubah karakteristiknya sesuai dengan nilai-nilai keabuan yang berada di bawah *mask* [11]. Filter ini biasanya digunakan dalam domain frekuensi.

2.2. Morphologi Citra

Morphologi merupakan tahapan dalam pengolahan citra digital yang secara khusus dimaksudkan untuk menganalisis bentuk dalam citra, pada umumnya digunakan untuk analisis citra biner tetapi juga bisa digunakan untuk citra abu-abu. Nilai setiap piksel pada sebuah citra diperoleh melalui proses perbandingan antara piksel yang bersesuaian pada citra digital masukan dengan piksel tetangganya. Operasi morphologi bergantung pada urutan kemunculan dari piksel, tidak memperhatikan nilai numerik dari piksel. Pemrosesan citra secara morphologi, citra dipandang sebagai himpunan.

Secara umum proses citra morphologi dilakukan dengan mem-*passing* sebuah *structuring element* terhadap sebuah citra. *Structuring element* dapat diibaratkan dengan *mask* pada pemrosesan citra biasa.

Hasil analisis morphologi dapat digunakan untuk pengambilan keputusan dengan menganalisis lebih lanjut. Terdapat dua operasi dasar dari yaitu dilasi dan erosi.

2.2.1 Dilasi

Operasi dilasi dilakukan dengan melebarkan segmen objek dengan menambahkan lapisan di sekeliling objek. Terdapat 2 cara untuk melakukan operasi ini yaitu dengan cara mengubah semua titik latar yang bertetangga dengan titik batas menjadi titik objek, atau lebih mudahnya set setiap titik yang tetangganya adalah titik objek menjadi titik objek. Cara kedua yaitu dengan mengubah semua titik di sekeliling titik batas menjadi titik objek atau lebih mudahnya set semua titik tetangga sebuah titik objek menjadi titik objek.

2.2.2 Erosi

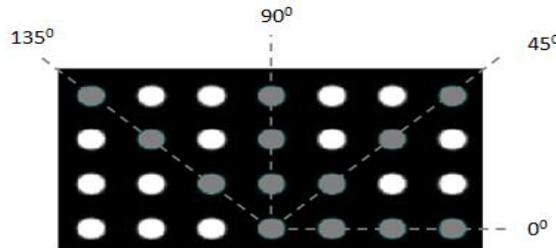
Operasi erosi adalah kebalikan dari operasi dilasi. Pada operasi ini ukuran objek diperkecil dengan mengikis sekeliling objek. Cara yang dapat dilakukan ada 2. Cara pertama yaitu dengan mengubah semua titik batas menjadi titik latar dan cara kedua dengan menset semua titik di sekeliling titik latar menjadi titik latar.

2.3. Grey-Level Co-Occurance Matrix (GLCM)

Salah satu teknik untuk memperoleh ciri statistik orde dua adalah dengan menghitung probabilitas hubungan ketetanggaan antara dua piksel pada jarak dan orientasi sudut. Pendekatan ini bekerja dengan membentuk sebuah matriks kookurensi dari data citra, dilanjutkan dengan menentukan ciri sebagai fungsi dari matriks antara tersebut.

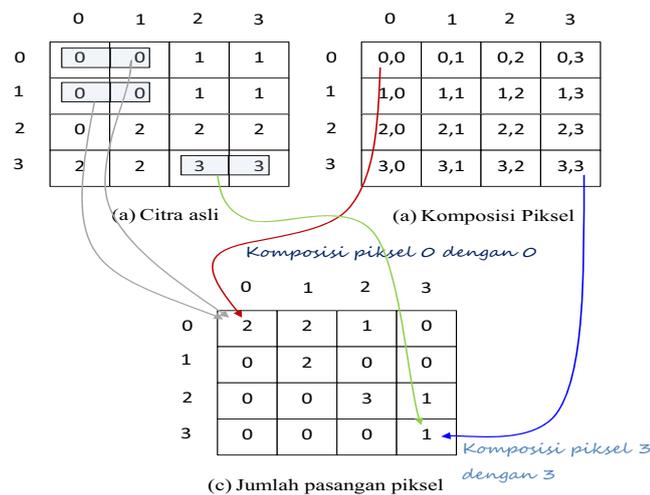
Kookurensi berarti kejadian bersama yaitu jumlah kejadian satu level nilai piksel bertetangga dengan satu level nilai piksel lain dalam jarak (d) dan orientasi (θ) tertentu. Jarak dinyatakan dalam piksel dan orientasi dinyatakan dengan derajat. Matriks kookurensi merupakan matriks bujur sangkar dengan jumlah elemen sebanyak kuadrat jumlah intensitas piksel citra. Setiap (p,q) pada matriks kookurensi berorientasi

θ berisi peluang kejadian piksel bernilai p bertetangga dengan piksel bernilai q pada jarak d serta orientasi θ dan $(180-\theta)$.



Gambar 1. Contoh arah dengan sudut 0° , 45° , 90° , 135° (Kadir & Susanto, 2013)

Ketetanggaan piksel dapat dipilih ke arah timur (kanan). Salah satu cara untuk merepresentasikan hubungan ini yaitu berupa $(1,0)$, yang menyatakan hubungan dua piksel yang berjajar horizontal dengan piksel bernilai 1 diikuti dengan piksel bernilai 0. Berdasarkan komposisi tersebut, jumlah kelompok piksel yang memenuhi hubungan tersebut dihitung [12].



Gambar 2. Penentuan awal matrik bebasis pasangan dua piksel

Untuk mendapatkan fitur hanya menggunakan lima besaran untuk GLCM, berupa *angular second moment* (ASM), *contrast*, *inverse different moment* (IDM), entropi, dan korelasi [12].

a. *Angular Second Moment*

Angular Second Moment yang merupakan ukuran homogenitas citra dihitung dengan cara seperti berikut:

$$ASM = \sum_{i=1}^L \sum_{j=1}^L (GLCM(i,j))^2 \tag{1}$$

b. *Contrast*

Kontras yang merupakan ukuran keberadaan variasi aras keabuan piksel citra dihitung dengan cara seperti berikut:

$$contrast = \sum_{n=1}^L n^2 \{ \sum_{|i-j|=n} GLCM(i,j) \} \tag{2}$$

c. *Correlation*

Correlation yang merupakan ukuran ketergantungan linear antarnilai aras keabuan dalam citra dihitung dengan menggunakan rumus:

$$Korelasi = \frac{\sum_{i=1}^L \sum_{j=1}^L (i,j) (GLCM(i,j) - \mu_i' \mu_j')}{\sigma_i' \sigma_j'} \tag{3}$$

d. *Inverse Different Moment*

IDM digunakan untuk mengukur homogenitas. IDM dihitung dengan cara seperti berikut:

$$IDM = \sum_{i=1}^L \sum_{j=1}^L \frac{(GLCM(i,j))^2}{1+(i-j)^2} \tag{4}$$

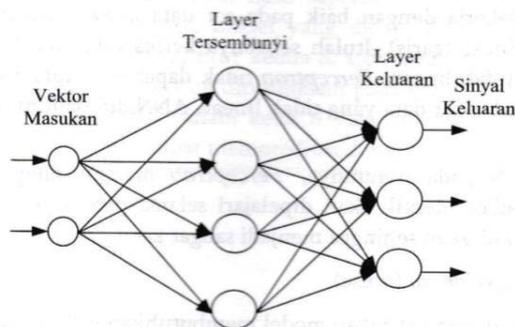
e. *Entropy*

Entropi menyatakan ukuran ketidakteraturan aras keabuan di dalam citra. Nilainya tinggi jika elemen-elemen GLCM mempunyai nilai yang relatif sama. Nilai rendah jika elemen-elemen GLCM dekat dengan nilai 0 atau 1. Rumus untuk menghitung entropi:

$$Entropi = - \sum_{i=1}^L \sum_{j=1}^L (GLCM(i,j) \log(GLCM(i,j))) \tag{5}$$

2.4. Multilayer Perceptron dengan Algoritma Backpropation

Multilayer Percetroon (MLP) merupakan *Artificial Neural Network* turunan dari *perceptron*, berupa *feedforward neureal network* dengan satu atau lebih layer tersembunyi (hidden layer) dan merupakan jaringan yang pembelajarannya terawasi, mentrasformasikan data input sedemikian rupa sehingga menghasilkan keluaran/ respon seperti yang diinginkan. Jaringan ini terdiri dari atas satu layer masukan, setidaknya satu layer neuron komputasi di tengah (tersembunyi), dan sebuah layer neuron komputasi keluaran. Sinyal masukan dirambatkan dengan arah maju pada pada layer-per-layer[13].



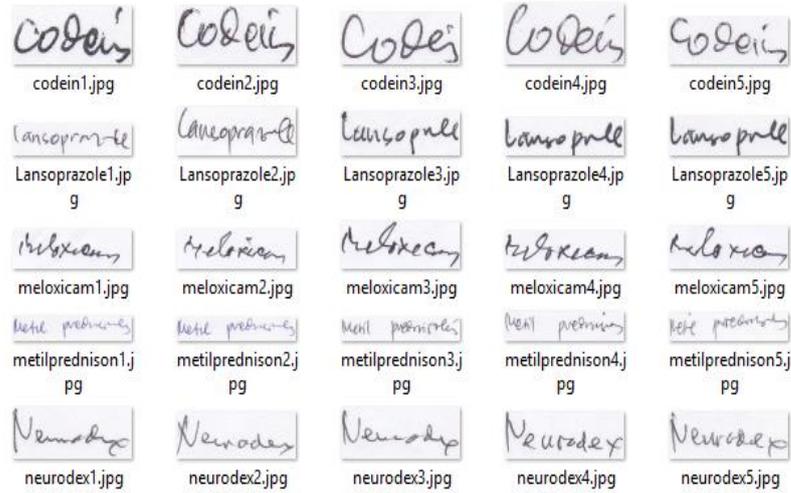
Gambar 3. Arsitektur Backpropagation [13]

3. METODE PENELITIAN

3.1. Akuisisi Citra

Citra nama obat diperoleh dengan men-*scanner* resep dokter, kemudian dilakukan *cropping* secara manual untuk mendapatkan citra nama obat sebanyak 200 citra yang terdiri 5 macam nama obat. Tiap

macam nama obat terdapat 40 citra nama obat. Contoh citra nama obat yang telah dikroping terdapat pada gambar di bawah ini :



Gambar 4. Citra Nama Obat yang telah Dikroping

3.2. Penyesuaian Ukuran Citra

Untuk mendapat hasil yang lebih baik maka dilakukan penyesuaian ukuran citra dengan ukuran 400 x 100 piksel.



Gambar 5. Citra Nama Obat setelah Penyesuaian Ukuran 400 x 100

3.3. Merubah Citra RGB menjadi Grayscale

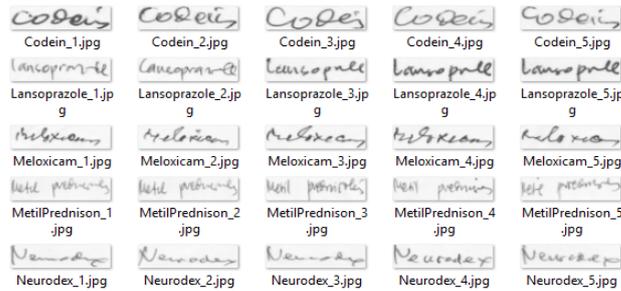
Melakukan perubahan citra RGB menjadi citra *Grayscale*



Gambar 6. Citra Nama Obat setelah Dirubah Menjadi *Grayscale*

3.4. Penghalusan Citra

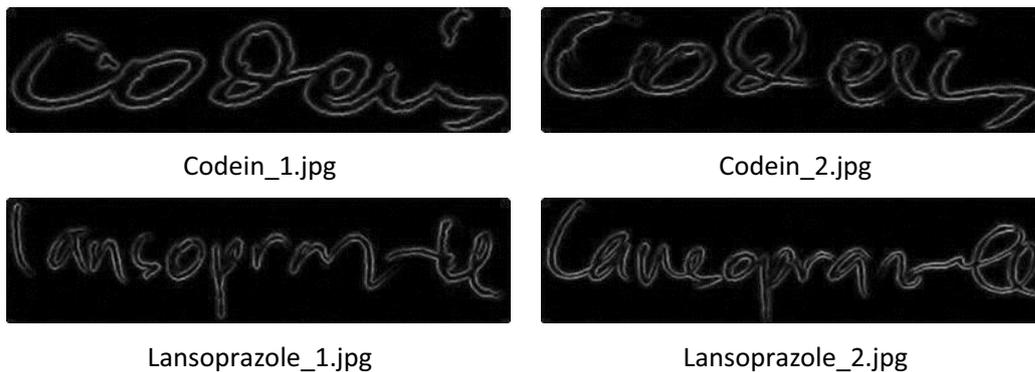
Penghalusan citra dimaksudkan untuk menghilangkan *noise* yang terjadi pada waktu *scanner* citra resep atau keadaan awal citra resep yang ber-*noise*. Penghalusan dilakukan dengan filter median dan filter *wiener*.

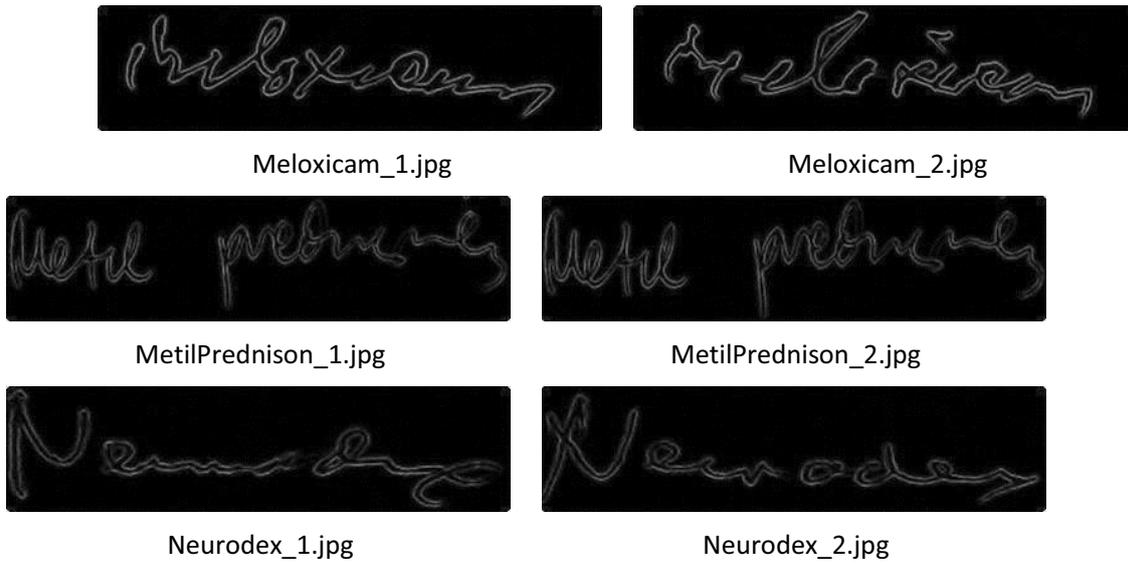


Gambar 7. Citra Nama Obat Setelah Dilakukan Penghalusan

3.5. Proses Morphologi

Pada tahapan ini dilakukan proses dilasi dan erosi pada citra nama obat.



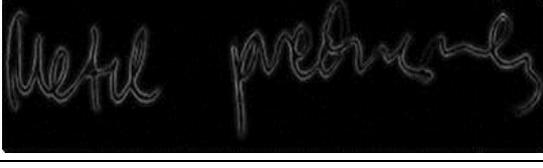


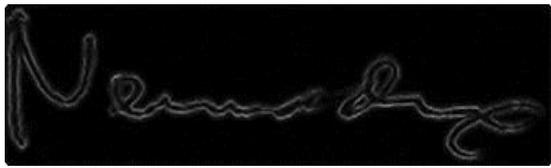
Gambar 7. Citra Nama Obat setelah Dilakukan Proses Morfologi

3.6. Ekstraksi Fitur GLCM 0°

Contoh hasil ekstraksi fitur nama obat dengan GLCM 0° dapat dilihat pada tabel di bawah ini :

Tabel 1. Hasil Ekstraksi Fitur GLCM 0°

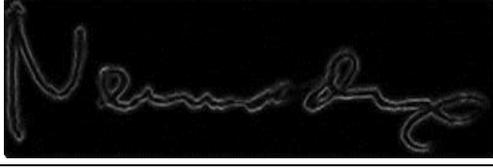
Contoh Citra Nama Obat Setelah Proses Morphologi	Nilai GLCM 0°	
	ASM	0,090969
	Contras	66,38793
	IDM	0,564243
	Entropi	5,077645
	Korelasi	0,001816
	ASM	0,166112
	Contras	155,4118
	IDM	0,644701
	Entropi	4,290141
	Korelasi	0,002049
	ASM	0,210295
	Contras	219,4173
	IDM	0,689614
	Entropi	3,871816
	Korelasi	0,001372
	ASM	0,185662
	Contras	140,7609
	IDM	0,654877
	Entropi	4,136852
	Korelasi	0,002486

	ASM	0,24803
	Contras	48,3785
	IDM	0,727499
	Entropi	3,605751
	Korelasi	0,003339

3.7. Ekstraksi Fitur GLCM 45 °

Contoh hasil ekstraksi fitur nama obat dengan GLCM 45 ° dapat dilihat pada tabel di bawah ini.

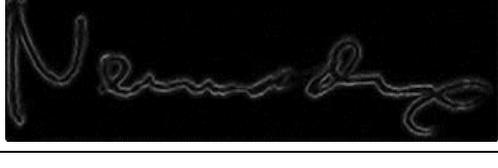
Tabel 2. Hasil Ekstraksi Fitur GLCM 45 °

Contoh Citra Nama Obat Setelah Proses Morphologi	Nilai GLCM 45 °
	ASM 0,071964 Contras 192,283 IDM 0,461309 Entropi 5,416944 Korelasi 0,001577
	ASM 0,141005 Contras 213,7205 IDM 0,564756 Entropi 4,513961 Korelasi 0,001856
	ASM 0,181782 Contras 275,8246 IDM 0,607096 Entropi 4,088062 Korelasi 0,001291
	ASM 0,159831 Contras 169,2797 IDM 0,58466 Entropi 4,331154 Korelasi 0,002334
	ASM 0,212337 Contras 132,5317 IDM 0,6312 Entropi 3,930426 Korelasi 0,002769

3.8. Ekstraksi Fitur GLCM 90 °

Contoh hasil ekstraksi fitur nama obat dengan GLCM 90 ° dapat dilihat pada tabel Hasil Ekstraksi Fitur GLCM 90 ° berikut ini.

Tabel 3. Hasil Ekstraksi Fitur GLCM 90 °

Contoh Citra Nama Obat Setelah Proses Morphologi	Nilai GLCM 90 °	
	ASM	0,080562
	Contras	197,5422
	IDM	0,495497
	Entropi	5,324496
	Korelasi	0,001569
	ASM	0,156106
	Contras	104,9747
	IDM	0,610523
	Entropi	4,338601
	Korelasi	0,002213
	ASM	0,199557
	Contras	252,4928
	IDM	0,643205
	Entropi	3,985406
	Korelasi	0,001325
	ASM	0,176948
	Contras	78,62017
	IDM	0,637016
	Entropi	4,142527
	Korelasi	0,002804
	ASM	0,22421
	Contras	105,5086
	IDM	0,663279
	Entropi	3,833629
	Korelasi	0,002954

3.9. Ekstraksi Fitur 135 °

Contoh hasil ekstraksi fitur nama obat dengan GLCM 135 ° dapat dilihat pada tabel sebagian hasil ekstraksi fitur nama obat dengan GLCM 135 ° berikut ini.

Tabel 4. Sebagian Hasil Ekstraksi Fitur Nama Obat dengan GLCM 135 °

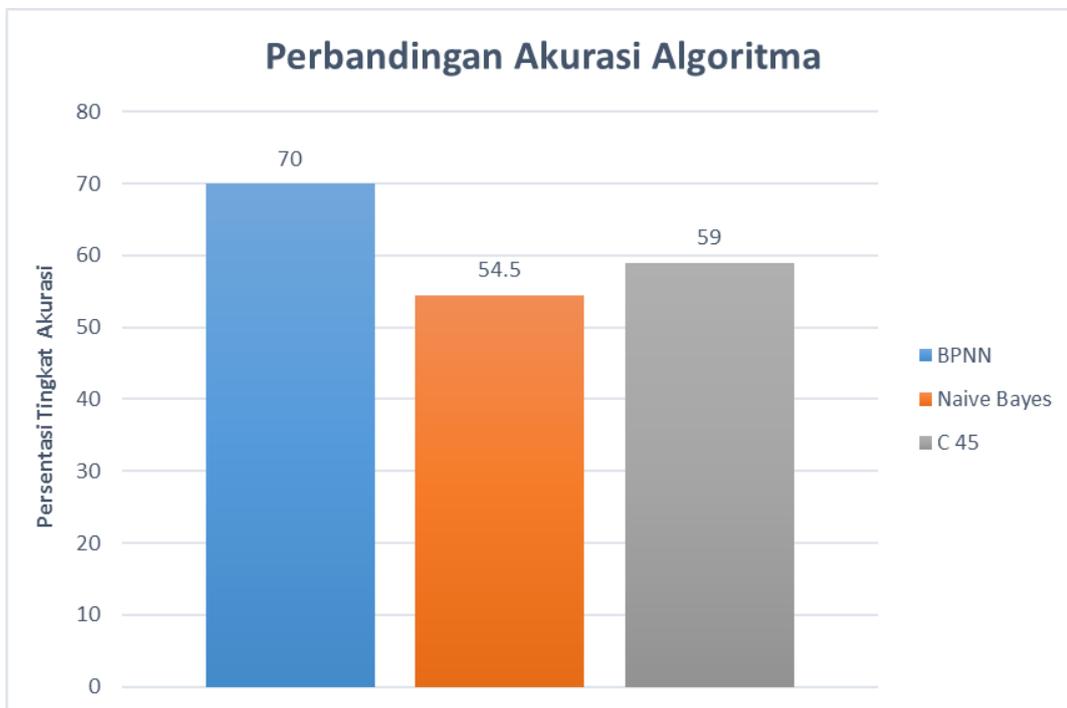
Contoh Citra Nama Obat Setelah Proses Morphologi	Nilai GLCM 135 °
	ASM 0,071485 Contras 277,1828 IDM 0,44954 Entropi 5,479339 Korelasi 0,001418
	ASM 0,140601 Contras 222,7341 IDM 0,559688 Entropi 4,53417 Korelasi 0,001827
	ASM 0,18105 Contras 479,7683 IDM 0,602416 Entropi 4,137835 Korelasi 0,001003
	ASM 0,159701 Contras 184,386 IDM 0,580466 Entropi 4,353525 Korelasi 0,002257
	ASM 0,212815 Contras 134,044 IDM 0,62991 Entropi 3,936157 Korelasi 0,002759

4. HASIL EKSPERIMEN

Pengujian dilakukan dengan membandingkan hasil akurasi algoritma *Back Propagation* dengan *Naive Bayes* dan C 4.5 dengan menggunakan tool rapid miner dan diuji dengan 2 sampai 10 *fold validation*.

Tabel 5 Perbandingan Algoritma *BPNN*, *Naive Bayer*, dan C 4.5

Algoritma	Akurasi	Fold Validation	GLCM
BPNN	70.00	5	0°
Naive Bayes	54.50	4	0°
C 45	59.00	5 & 10	0°



Gambar 5. Grafik Perbandingan Algoritma

Pada tabel 5 dan gambar 5 terlihat bahwa akurasi yang dihasilkan oleh algoritma *Backpropagation Neural Network* jauh lebih tinggi dari *Naive bayes* dan C 45 dengan menggunakan ekstraksi Fitur GLCM 0°.

Dari keseluruhan hasil perbandingan ekstraksi fitur GLCM 0°, GLCM 45°, GLCM 90° dan GLCM 135° dengan pengujian algoritma *Backpropagation Neural Network*, *Naive Bayes* dan C 45 dapat disimpulkan bahwa algoritma *Backpropagation Neural Network* dengan pengujian 5 *fold validation* dapat memberikan nilai akurasi tertinggi 70.00 % daripada algoritma *Naive Bayes* dan C 45.

Dengan demikian algoritma *backpropagation* dapat memberikan solusi terbaik dalam klasifikasi nama obat tulisan dokter dan data set fitur ekstraksi GLCM dapat mendukung klasifikasi nama obat tulisan tangan dokter.

5. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan dengan mengekstrak fitur citra nama obat dengan GLCM dapat diambil kesimpulan sebagai berikut.

- a. Data set ekstraksi Fitur GLCM dengan 0° dapat mendukung hasil klasifikasi algoritma *Backpropagation* dengan akurasi tertinggi.
- b. *Backpropagation Neural Network* secara umum dapat memberikan nilai akurasi tertinggi dibanding dengan *Naive Bayes* dan *C 4.5*.

DAFTAR PUSTAKA

- [1] Rahmawati, F. & Oetari, R. A., 2002. Kajian Penulisan Resep : Tinjauan Aspek Legalitas dan kelengkapan resep di apotek-apotek kotamadya yogyakarta. *Majalah Farmasi Indonesia*, Volume 13 (2), pp. 86-94.
- [2] Kementerian Kesehatan Republik Indonesia, 2012. Penggerakan Penggunaan Obat Rasional. Dalam: Jakarta: s.n.
- [3] Puspitaningrum, D., 2006. Pengantar Jaringan Syaraf Tiruan. Yogyakarta: Andi Offset.
- [4] Gharoie, R. A. & Farajpoor, M. A., 2009. Handwritten Farsi Character Recognition Using Artificial Neural Network. *International Journal of Computer Science and Information Security*, Volume 4.
- [5] Al-Omari, S. a. K., Sumari, P., Al-Taweel, S. A. & J.A.Husain, A., 2009. Digital Recognition Using Neural Network. *Journal Of Computer Science*, Volume 5 (6), pp. 427-434.
- [6] Noeri, O., Mumtazah Syed Ahmad, S. & Shakil, A., 2011. Offline Malay Handwritten Cheque Words Recognition Using Artificial Neural Network. *Journal Of Applied Sciences*, Volume 11, pp. 86 - 95.
- [7] Pal, A. & Singh, D., 2010. Handwritten English Character Recognition Using Neural Network. *International Journal of Computer Science & Communication*, Volume 1, pp. 141-144.
- [8] K. Pratt, W., 2007. *Digital Image Processing*. Los Altos, California: Willey.
- [9] Gonzales, R. C. & E Woods, R., 2002. *Digital Image Processing second edition*. s.l.:Pearson Poentice Hall.
- [10] Prasetyo, E., 2011. *Pengolahan Citra Digital dan Aplikasinya menggunakan Matlab*. Yogyakarta: ANDI.
- [11] Sianipar, R. H., 2013. *Pemrograman Matlab dalam contoh dan penerapan*. Bandung: Informatika.
- [12] Kadir, A. & Susanto, A., 2013. *Teori dan Aplikasi pengolahan Citra*. yogya: Andi.
- [13] Prasetyo, E., 2012. *Data Mining Konsep dan Aplikasi menggunakan matlab*. Yogyakarta: Andi.