PENGENALAN CITRA SASIRANGAN BERBASIS FITUR *GLCM* DAN *MEDIAN FILTER* MENGGUNAKAN *LEARNING VECTOR QUANTITATION*

Muharir

Fakultas Teknologi Informasi
Universitas Islam Kalimantan Muhammad Arsyad Al Banjari Banjarmasin
muharir17@gmail.com

Abstrak

Sasirangan adalah kain khas suku Banjar provinsi Kalimantan Selatan, kain sasirangan merupakan salah satu budaya yang dimiliki bangsa Indonesia yang harus dijaga dan dilestarikan. Sasirangan saat ini memiliki beragam motif dan sebagian motif-motif yang ada belum dikenal masyarakat. Pengenalan citra saat ini sudah banyak dilakukan namun akurasi yang dihasilkan masih rendah dan belum diketahui algoritma apa yang menghasilkan akurasi terbaik untuk mengenali citra sasirangan. Pada penelitian ini, teknik yang digunakan untuk ekstraksi fitur adalah metode Grey Level Co-occurrence Matriec. Untuk pengenalan pola digunakan metode Linear Vector Quantization dengan melakukan beberapa tahapan pengujian, menggunakan pengujian dengan training dan testing, dan menggunakan cross validation. Pengujian menggunakan training dan testing dilakukan dengan 4 ketentuan 60%:40%, 70%:30%, 80%:20%, 90%:10%, untuk sampling type yang terbagi menjadi 3 bagian : yaitu Linear Sampling, Shuffled Sampling, dan Stratified Sampling menggunakan teknik Fold Cross Validation dengan masing-masing di uji dari 2, 3, 4, 5, 6, 7, 8, 9, dan 10 . Dari hasil penelitian menunjukkan bahwa pada pengujian tersebut akurasi tertinggi untuk mengenali citra sasirangan menggunakan sampling type dengan teknik Shuffled Sampling di dapatkan (93,33%) dengan nilai validasi 5, dengan perbandingan data training 70% dan testing 30% .

Keyword : *Median Filter, Grey Level Co-occurrence Matriec, Learning Vector Quantization*

1. PENDAHULUAN

Indonesia adalah negara yang kaya akan sumber daya alam dan memiliki keberagaman suku, agama, dan ras. Indonesia memiliki lebih dari 300 suku bangsa. Dimana setiap suku bangsa memiliki kebudayaan yang berbeda-beda antara satu dengan yang lainnya. Diantara warisan budaya yang dikenal baik di dalam negeri dan luar negeri adalah warisan berupa kain batik. Batik juga telah diakui *United Nations Educational, Scientific, and Culture Organization* (UNESCO) pada tanggal 2 Oktober 2009, sebagai hak intelektual kebudayaan bangsa Indonesia [1]. Bila di Batak terkenal dengan kain bernama Ulos dan Jawa berupa batik, maka di Kalimantan yang masyarakatnya dikenal dengan suku Banjar juga memiliki kain bernama sasirangan sebagai warisan budayanya. Berbeda dengan kain batik dan ulos, kain sasirangan dibuat menggunakan teknik tusuk jelujur. Kain Sasirangan merupakan salah satu kebudayaan yang dimiliki bangsa Indonesia yang harus dijaga dan dilestarikan. Sasirangan memiliki karakteristik yang beragam pada motifnya. Motif-motif dan ragam hiasnya lahir dari proses kognitif manusia yang diperoleh dari lingkungan sekitarnya. Sasirangan saat ini memiliki beragam motif dan sebagian dari motif-motif yang ada belum dikenal masyarakat. Hal inilah yang dianggap menarik untuk diteliti menggunakan sains dan teknologi. Pengenalan citra saat ini sudah banyak dilakukan namun akurasi yang dihasilkan masih rendah.

Berdasarkan permasalahan diatas, peneliti akan mengusulkan metode *Learning Vector*

Quantization untuk meningkatkan akurasi dalam mengenali citra sasirangan.

2. METODE PENELITIAN

Metode penelitian yang dilakukan adalah metode penelitian eksperimen, dengan tahapan penelitian sebagai berikut :

a. Pengumpulan Data

Data primer yang diporeh dari peneliti adalah data hasil foto sasirangan dari penelitian sebelumnya sebanyak 3 motif yaitu Abstrak dengan label (a), Gagatas dengan label (g), Kulat Kurukit dengan label (k), dan setiap motif diambil sebanyak 10 gambar.



Gambar 1. Motif Gagatas (g)



Gambar 2. Motif Kulat Kurikit (k)



Gambar 3. Motif Abstrak (a)

b. Pengolahan Awal Data (Data Preprocess-ing)

Merupakan tahap pengolahan awal data dalam mempersiapkan data yang telah diperoleh dari tahap pengumpulan data sebelumnya untuk dilakukan proses cropping, resizesing, dan selanjutnya dilakukan filter menggunakan Median Filter untuk menghilangkan derau.

c. Median Filter

Median Filter merupakan salah satu metode yang digunakan dalam restorasi citra. Istilah Restorasi mempunyai perbedaan dengan peningkatan citra. Peningkatan citra (*Image Enhancement*) merupakan istilah yang menyatakan usaha untuk membuat gambar agar lebih baik dari sudut pandang pengolahnya. Hal ini dilakukan misalnya melalui pengubahan kontras dan kecerahan. Berbeda dengan peningkatan citra, restorasi citra merupakan proses untuk membuat citra yang kualitasnya turun akibat adanya tambahan derau agar menjadi mirip dengan kedadaan aslinya.

Median filter dapat dipakai untuk menghilangkan derau dalam citra. Filter ini menggunakan nilai median piksel-piksel di dalam sebagai keluaran $\hat{f}$. Median Filter dapat ditulis sebagai berikut :

$$\hat{f}(y,x) = median_{(p,q)\in S_{yx}}(g(p,q)) \quad (1)$$

d. GLCM (*Greey Level Coocurence Matrix*)

Melakukan proses ekstraksi *Greey Level Coocurence Matrix* (GLCM) dengan atribut yang digunakan hanya beberapa besaran yang diusulkan Haralick, Newsam dan Kammath (2005) yaitu hanya menggunakan lima besaran, berupa *angular second moment* (ASM), contrast, inverse different moment (IDM), entropi, dan korelasi.

ASM yang merupakan ukuran homogenitas citra dihitung dengan cara seperti berikut:

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

Dalam hal ini, L menyatakan jumlah level yang digunakan untuk komputasi.

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

$$Kontras =$$

$$\sum_{n=1}^{L} n^2 \left\{\sum_{|i-j|=n} GLCM(i,j)\right\} \quad (3)$$

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

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

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:

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

Korelasi yang merupakan ukuran ketergantungan linear antarnilaiaras keabuan dalam citra dihitung dengan menggunakan rumus:

$$Korelasi = \frac{\sum_{i=1}^{L}\sum_{j=1}^{L}(ij)(GLCM(i,j)-\mu_i'\mu_j'}{\sigma_i'\sigma_j'} \quad (6)$$

dengan

$$\mu_i' = \sum_{i=1}^{L}\sum_{j=1}^{L} i * GLCM(i,j) \quad (7)$$

$$\mu_j' = \sum_{i=1}^{L}\sum_{j=1}^{L} j * GLCM(i,j) \quad (8)$$

$$\sigma_j^2 = \sum_{i=1}^{L}\sum_{j=1}^{L} GLCM(i,j)(i-\mu_i')^2 \quad (9)$$

$$\sigma_i^2 = \sum_{i=1}^{L}\sum_{j=1}^{L} GLCM(i,j)(i-\mu_i')^2 \quad (10)$$

e. Learning Vector Quantization

Metode LVQ menggunakan vector acuan (*vector reference*) dari satu unit keluaran yang menjadi acuan bagi kelas/ kategori yang mewakili oleh keluaran tersebut. Pendekatan yang dilakukan adalah mengelompokkan vektor input berdasarkan kedekatan jarak vektor input terhadap bobot (metode kuadrat jarak euclidean minimum). Proses pem-belajaran LVQ merupakan pembelajaran supervised atau dengan kata lain menggunakan pengarahan, dengan tujuan untuk mendapatkan vektor-vektor pewakil yang akan melakukan kuantisasi terhadap vektor masukan.

Pada metode ini, pada saat menentukan vektor-vektor pewakil untuk digunakan pada saat pelatihan, maka dengan pengarahan vektor pewakil tersebut akan mengenali target yang telah diberikan bersamaan dengan input. Dalam proses training, unit-unit output diarahkan kepada suatu *decision surface* teori dengan mengupdate bobot pada training. LVQ diarahkan untuk menentukan unit keluaran yang paling sesuai dengan target dari vektor masukannya melalui cara pergeseran posisi vektor pewakil. Apabia vektor data training dikelompokkan sama dengan vektor pewakil pemenang, maka vektor pewakil digeser mendekati vektor training dengan notasi :

$$W(t+1) = w(t) + \alpha(t)(x(t)-w(t)) \quad (11)$$

Bila vektor data training dikelompok-kan tidak sama dengan vektor pewakil pemenang, maka vektor pewakil digeser menjauhi vektor training yang dinyata-kan dengan notasi :

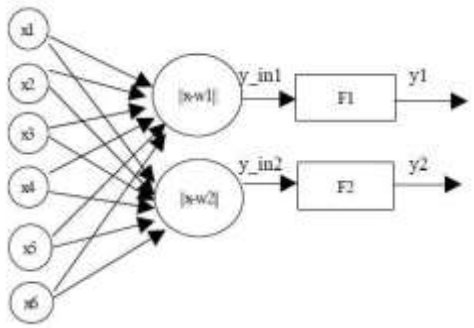
$$W(t+1) = w(t) + \alpha(t)(x(t)-w(t)) \quad (12)$$

Dimana :

α = laju pembelajaran

w = posisi vektor pewakil saat **t**

x = posisi vektor masukan saat t



Gambar 4. Arsitektur Jaringan LVQ [15]

Gambar diatas merupakan contoh struktur jaringan LVQ yang memiliki empat input layer dengan dua unit neuron pada output layers. W1 dan W2 merupakan bobot yang menghubungkan input layer ke output layers. Setiap fungsi aktivitas F melakukan pemetaan setiap y diklasifikasikan y1 atau y2. Pada F1, jika | x – w1 | < | x – w2|, maka y_in1 dipetakan ke y1=1 dan dipetakan ke y1=0 jika sebaliknya. Kondisi ini juga berlaku untuk F2 dengan kondisi yang sesuai.

Secara garis besar, LVQ akan mencari unit keluaran yang paling miri dengan vektor masukan. Jika vektor pelatihan adalah bagian dari kelas yang sama, maka vektor bobot digeser mendekati vektor masukan tersebut. Sebaliknya jika vektor pelatihan bukan bagian dari kelas yang sama, maka vektor bobot digeser menjauhi vektor masukan tersebut.
Berikut algoritma metode pembelajaran LVQ

Langkah 0 : Inisialisasi Inisialisasi vektor-vektor acuan Inisialisasi laju pembelajaran alpha=0
Langkah 1 : Jika kondisi stop salah
Langkah 2 : Untuk setiap training vektor, lalukan langkah 2-6
Langkah 3 : Dapatkan nilai j sehingga ||x-wj|| bernilai minimum

$$Di = \sqrt{\sum_{j-1}^{m} \quad (Wji - Xi^2)}$$

Langkah 4 : Update nilai bobot wj

Jika T = Cj

Wj(baru)=wj(lama) + α(x-wj(lama))

Jika T ≠ Cj

Wj(baru)=wj(lama) + α(x-wj(lama))

Langkah 5 : Update nilai *learning rate*

Langkah 6 : Uji kondisi stop

Uji kondisi stop ini dapat dilakukan berdasarkan jumlah iterasi tertentu (mulai dari langkah 1) atau setelah laju pembeljaran yang telah mencapai harga yang sangat kecil.
Dimana :

X = training vektor (x1,…, xi,…, xn)

T = kategori training vektor yang benar untuk pelatihan

Wj = vektor bobot untuk unit keluaran ke-j (w1,…,wij,…,win)

Cj = kategori atau kelas yang mewakili oleh nilai unit keluaran ke-j (hasil training)

||x-wj|| = *Euclidian distances* antara 18 vektor masukkan dan vektor bobot dari unit keluaran ke-j.

f. Tahapan Penelitian

Penelitian ini dilakukan berdasarkan tahapan-tahapan dengan orientasi pada hasil peningkatan akurasi dari metode LVQ, maka tahapan yang penelitian ini adalah sebegai berikut :

1. Pengumpulan Data (Data *Gathering*), Data primer yang diperoleh peneliti adalah data hasil foto sasirangan dari penelitian

sebelumnya sebanyak 3 motif yaitu Abstrak, Gagatas, Kulat Kurikit tiap motif diambil sebanyak 10 gambar.

2. Pengolahan Awal Data (Data *Pre-processing*), Merupakan tahap pengolahan awal data dalam mempersiapkan data yang telah diperoleh dari tahap pengumpulan data sebelumnya untuk dilakukan proses *cropping, resizesing*, dan selanjutnya dilakukan filter menggunakan Median Filter untuk menghilangkan derau pada citra, setelah itu dilanjutkan dengan ektraksi fitur menggunakan GLCM.
3. Model/Metode Yang Diusulkan (*Proposed Model/Method*), Metode yang diusulkan adalah metode ektrasi GLCM yang diuji dengan *Linear Vector Quantization* untuk klasifikasinya.
4. Eksperimen dan Pengujian Metode (*Method Test and Experiment*), Eksperimen dan pengujian metode merupakan perhitungan dan simulasi pengujian metode yang digunakan pada penelitian ini.
5. Evaluasi dan Validasi Hasil (*Result Evaluation and Validation*), Pada tahapan ini dilakukan evaluasi terhadap kesimpulan atau informasi yang diperoleh dari tiga kelas data yang diproses hasil yang diperoleh yaitu tingkat akurasi.

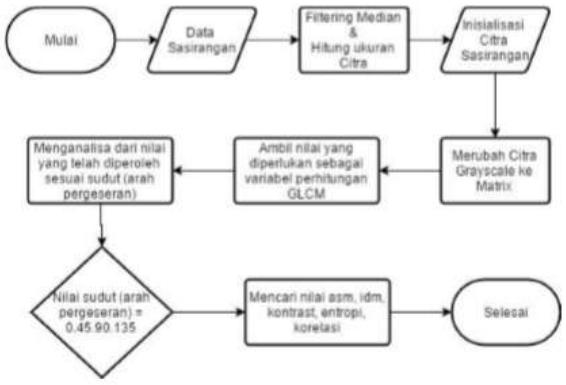
3. HASIL DAN PEMBAHASAN

A. Eksperimen dan Pengujian Model-/Metode

Tujuan dari penelitian ini adalah mengenali citra sasirangan dengan cara mengklasifikasi-kan citra berdasarkan motif, data diuji menggunakan Rapidminer dengan metode Algoritma *Linear Vector Quantization*.

B. Ekstraksi Fitur GLCM

Ekstraksi fitur GLCM ini menganalisa berdasarkan 5 atribut berupa, *angular second moment* (asm), *contrast, inverse different moment* (idm), *entropi, dan korelasi*.



Gambar 4. Digram alir ekstraksi GLCM

Ekstraksi GLCM sudut 0^0, sudut 45^0, sudut 90^0, sudut 135^0 dengan menggunakan Matlab akan digunakan sebagai file repository pada aplikasi RapidMiner 7.4 untuk melanjutkan ke tahap selanjutnya yaitu klasifikasi menggunakan *Linear Vector Quantization*.

Setelah didapatkan 5 besaran dari masing-masing sudut, selanjutnya menjadikan dataset baru dengan cara menggabungkan nilai besaran dari 4 sudut tadi menjadi satu buah dataset baru yang terjadi dari 20 atribut dan 1 buah label yang akan menjadi target class seperti pada tabel n berikut :

Tabel 1. atribut dataset

No	Nama Kolom	Tipe Data	Jenis
1	Label	Polynominal	Label
2	$ASM0^0$	Real	Atribut
3	$IDM0^0$	Real	Atribut
4	$Entropi0^0$	Real	Atribut
5	$Korelasi0^0$	Real	Atribut
6	$Kontras0^0$	Real	Atribut
7	$ASM45^0$	Real	Atribut
8	$IDM45^0$	Real	Atribut
9	$Entropi45^0$	Real	Atribut
10	$Korelasi45^0$	Real	Atribut
11	$Kontras45^0$	Real	Atribut
12	$ASM90^0$	Real	Atribut
13	$IDM90^0$	Real	Atribut
14	$Entropi90^0$	Real	Atribut
15	$Korelasi90^0$	Real	Atribut
16	$Kontras90^0$	Real	Atribut
17	$ASM135^0$	Real	Atribut
18	$IDM135^0$	Real	Atribut
19	$Entropi135^0$	Real	Atribut
20	$Korelasi135^0$	Real	Atribut
21	$Kontras135^0$	Real	Atribut

C. Training dan Testing

Pengenalan citra sasirangan menggunakan Algoritma LVQ (*Linear Vector Quantization*) berbasis median filter yang dilakukan kepada 2 data hasil ekstraksi GLCM yang dengan menggunakan filter dan tidak menggunakan filter dengan membagi data terlebih dahulu menjadi dua bagian data training dan data testing dengan 4 pembagian sebagai berikut :

1. Membagi data 60% untuk data training dan 40% untuk data testing
2. Membagi data 70% untuk data training dan 30% untuk data testing.
3. Membagi data 80% untuk data training dan 20% untuk data testing.
4. Membagi data 90% untuk data training dan 10% untuk data testing.

Pada tahap pembagian data training dan data testing ini dimaksudkan agar merupakan menjadi proses pembelajaran terhadap pola data yang akan dikenali, proses ini digunakan untuk menguji validasi data yang telah dilakukan pada proses training dan testing dengan memasukkan data baru untuk mengetahui nilai akurasi yang tertinggi yang dihasilkan.

D. Pengujian dengan Proses Training dan Testing

Hasil pengujian dari data sasirangan dengan proses training dan testing akan membentuk matrik yang terdiri dari true gagatas atau tupel gagatas, true kulatkurikit atau tupel kulatkurikit, dan true abstrak atau tupel abstrak, kemudian training dan testing yang sudah disiapkan dalam confusion matrix sehingga menghasilkan sebagai berikut :

Tabel 2. Confusion Matrix dengan training dan testing berfilter median

Ket.	Pembagian Data	Jumlah Data	Accuracy
Training	60%	18	82.50%
Testing	40%	12	
Training	70%	21	83.33%
Testing	30%	9	
Training	80%	24	60.00%
Testing	20%	6	
Training	90%	27	80.00%
Testing	10%	3	

Dari tabel hasil pengujian diatas dapat disimpulkan bahwa pembagian data 70% untuk data training dan 30% untuk data testing menghasilkan nilai akurasi yang tinggi yaitu 83.33%.

Berikut ini adalah data sasirangan tidak berfilter dalam pengujian testing dan training yang sudah disiapkan dalam confusion matrix sehingga menghasilkan Tabel dan grafik sebegai berikut :

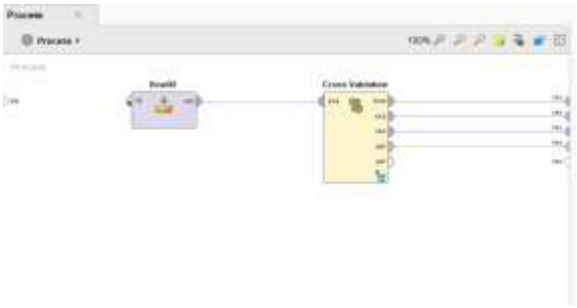
Tabel 3. Confusion Matrix dengan training dan testing tanpa filter

Ket.	Pembagian Data	Jumlah Data	Accuracy
Training	60%	18	74.17%
Testing	40%	12	
Training	70%	21	80.00%
Testing	30%	9	
Training	80%	24	65.00%
Testing	20%	6	
Training	90%	27	73.33%
Testing	10%	3	

Dari tabel hasil pengujian diatas dapat disimpulkan bahwa pembagian data 70% untuk data training dan 30% untuk data testing menghasilkan nilai akurasi yang tinggi yaitu 80.00%.

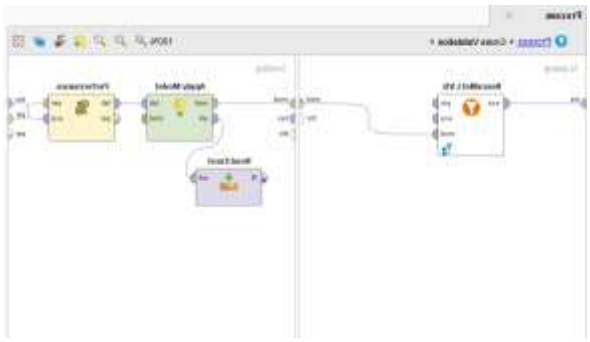
E. Validasi Hasil

Data ekstraksi fitur yang sebelumnya dibuat *file repository* pada aplikasi RapidMiner 7.4



Gambar 5. *Repository rapidminner*

Membuat model menentukan jenis validasi dan algoritma *Linear Vector Quantization* klasifikasi tahap selanjutnya adalah pengujian data menggunakan *cross validation* dengan ketentuan Number validation uji 2 sampai dengan 10, type validasi yang diuji mulai dari *Statified Sampling*, *Shufflled Sampling* dan *Linear Sampling*.



Gambar 5. Model dan menentukan validasi

Skenario penggujian yang dilakukan adalah dengan menggunakan data training sebagai pelatihan terhadap sistem. Hasil klasifikasi data uji oleh sistem dibandingkan dengan klasifikasi data uji sesungguhnya. Analisis dilakukan dengan membandingkan antara data uji yang berhasil diklasifikasi dengan benar dengan semua data uji, maka didapatkan akurasi sistem klasifikasi motif tersebut berdasarkan tekstur kain.

F. HASIL VALIDASI

Validasi citra hasil ektraksi dengan metode GLCM dibedakan menjadi 2 untuk melihat pengaruh median filter terhadap akurasi yang dihasilkan menggunakan LVQ.

Tabel 4. Hasil Validasi Citra Tanpa Filter

Validasi	Akurasi Linear Sampling	Akurasi Statified Sampling	Akurasi Shufflled Sampling
10	83.33%	77.78%	81.11%
9	81.48%	80.25%	82.72%
8	83.33%	81.94%	83.33%
7	80.95%	84.13%	79.37%
6	68.52%	81.48%	83.33%
5	73.33%	82.22%	84.44%
4	75.00%	75.00%	77.78%
3	59.26%	74.07%	77.78%
2	55.56%	77.78%	66.67%

Dari table diatas dapat disimpulkan bahwa akurasi Shuffled Sampling dengan nilai validasi 5 memiliki nilai tertinggi dengan akurasi 84.44% sedangkan akurasi statified sampling dengan nilai validasi 5 yaitu 82.22% ,dan akurasi Linear Sampling dengan validasi 8 yaitu 83.33%.

Table 5. Hasil Validasi Citra Berfilter Median

Validasi	Akurasi Linear Sampling	Akurasi Statified Sampling	Akurasi Shufflled Sampling
10	86.67%	83.33%	88.89%
9	86.42%	83.95%	90.12%
8	88.89%	87.50%	91.67%
7	87.30%	87.30%	90.48%
6	74.07%	81.48%	85.19%
5	73.33%	86.67%	93.33%
4	77.78%	83.33%	83.33%
3	59.26%	74.07%	85.19%
2	55.56%	83.33%	72.22%

Dari table diatas dapat disimpulkan bahwa akurasi Shuffled Sampling dengan nilai validasi 5 memiliki nilai tertinggi dengan akurasi 93.33% sedangkan akurasi statified sampling dengan nilai validasi 8 yaitu 87.50% ,dan akurasi Linear Sampling dengan validasi 8 yaitu 88.89%.

Dari 2 tabel diatas dapat disimpulkan pula bahwa metode median filter dapat meningkatkan akurasi pada pengenalan citra sasirangan.

4. KESIMPULAN

Berdasarkan hasil penelitian dan pengukuran didapatkan kesimpulan bahwa pengujian menggunakan metode Linear Vector Quantization dengan GLCM dan Median Filter dapat disimpulkan algoritma ini dapat digunakan untuk mengklasifikasikan sasirangan, dengan tingkat akurasi lebih baik. yaitu 93,33% dari penelitian sebelumnya yang menggunakan Naive Bayes dengan akurasi tertinggi yang didapat adalah 80,00%.

5. REFERENSI

Mustafa Zayid, "*Ekstraksi Fitur Motif Sasirangan Menggunakan Algoritma Naive Bayes Berbasis Color Histogram dan Gray Level Cooccurrence Matrices (GLCM)".*

Prabowo Anindito, Sarwoko Adi Eko, Er Riyanto Djalal. "*Perbandingan Antara Metode Kohonen Neural Network dengan Metode Learning Vector Quantization Pada Pengenalan Pola Tandatangan*".

Andrie, "*Implementasi Segmentasi Citra dan Algoritma Learning Vector Quantization (LVQ) Dalam Pengenalan Bentuk Botol*".

Kadir, Abdul & Susanto, Adhi. 2013. Teori dan Aplikasi pengolahan citra

Prasetyo Eko.2014. Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab, Penerbit Andi, Yogyakarta.

Padeiroot, Alfa Kiven.2014. *Pengembangan Aplikasi Mobile Untuk Pengenalan Pola Kain Bentenan Dengan Learning Vector Quantization*.

Putra, Darma, 2010, "Pengolahan Citra Digital", Penerbit Andi, Yogyakarta.