

KLASIFIKASI WARNA KULIT MANUSIA DENGAN ALGORITMA K-NEAREST NEIGHBOUR (KNN) *Human Skin Color Classification With K-Nearest Neighbour (KNN) Algorithm*

Ana Try N.1¹, I Gede Pasek Suta W.2¹, L. A. Syamsu Irfan3¹

ABSTRAK

Pengolahan citra merupakan pemrosesan citra dengan menggunakan komputer untuk mendapatkan kualitas citra yang lebih baik. Secara sistematis citra merupakan fungsi kontinyu dengan intensitas cahaya pada bidang dua dimensi.

Citra harus direpresentasikan secara numerik dengan nilai diskrit agar dapat diolah dengan komputer. Informasi-informasi dari citra tersebut dapat digunakan untuk mengekstrak atau mengenali suatu objek yang nantinya akan dijadikan informasi untuk menyelesaikan masalah tertentu.

Penelitian ini mengimplementasikan algoritma K-Nearest Neighbor (KNN) untuk melakukan klasifikasi warna kulit manusia dan bukan kulit dari citra masukan. Hasil dari penelitian dengan menggunakan metode ini mendapatkan akurasi sebesar 78%.

Kata kunci : *Pengolahan citra, citra, K-Nearest Neighbor, warna kulit*

ABSTRACT

Image processing is process of image using computer to get better quality of image. Systematicly, image is continuous functions with light intensity in two dimensional plane.

Image should be represented numerically with discrete value in order to be processed by computer. Informations from image can be used to extract or recognizing an object which will be used as information to solve specific problem.

This research implements K-Nearest Neighbor (KNN) algorithm to classify human skin color and not skin from input image. Result of the research by using this method gets 78% of accuracy.

Keywords : *Image processing, image, K-Nearest Neighbor, skin color*

PENDAHULUAN

Perkembangan di bidang teknologi dan informasi (TI) pada saat ini sangat pesat. Perkembangan ini dimanfaatkan di berbagai bidang untuk membantu manusia dalam mengerjakan pekerjaannya dan juga dapat menjadi hiburan di berbagai kalangan, mulai dari anak-anak hingga dewasa. Salah satu bidang TI yang berkembang saat ini adalah pengolahan citra. Terdapat berbagai jenis aplikasi untuk pengolahan citra seperti *Photoshop, Corel draw, dan Paint*. Aplikasi-aplikasi pengolahan citra dapat digunakan untuk memperbaiki kualitas gambar, mengubah ukuran gambar serta dapat memodifikasi gambar.

Citra merupakan fungsi kontinyu dengan intensitas cahaya pada bidang dua dimensi. Citra harus direpresentasikan secara numerik dengan nilai diskrit agar dapat diolah dengan komputer. Salah satu domain pengolahan citra adalah mengklasifikasi citra dikelompokkan tertentu, misalnya bertujuan untuk menentukan warna kulit manusia.

Untuk mengklasifikasi warna kulit, diperlukan suatu metode yang dapat mendefinisikan suatu piksel dalam

membedakan warna kulit. Salah satu metode yang dapat digunakan adalah metode *K-Nearest Neighbour* (KNN). KNN memiliki beberapa kelebihan yaitu ketangguhan terhadap *training data* yang memiliki banyak *noise* dan efektif apabila *training data*-nya besar (Sikki, 2009). Oleh karena itu, pada penelitian ini akan diteliti tentang klasifikasi warna kulit dengan menggunakan algoritma *K-Nearest Neighbour*.

Pengolahan citra bertujuan memperbaiki kualitas citra agar mudah diinterpretasi oleh manusia atau mesin (dalam hal komputer). Teknik-teknik pengolahan citra mentransformasikan citra menjadi citra lain. Jadi, masukannya adalah citra dan keluarannya juga citra, citra keluaran (*output*) memiliki kualitas lebih baik daripada citra masukan (*input*) (Munir, 2004).

Model warna. Model warna RGB menurut penelitian, bahwa kombinasi warna yang memberikan rentang warna paling lebar adalah *red* (R), *green* (G), dan *blue* (B). Warna-warna ini disebut warna pokok (*primaries*), dan sering disingkat sebagai

¹Jurusan Teknik Elektro, Fakultas Teknik, Universitas Mataram Jln. Majapahit 62, Mataram NTB -Indonesia
E-mail :akako123@yahoo.com

warna dasar RGB. Warna-warna lain dapat diperoleh dengan mencampurkan ketiga warna pokok tersebut (Munir, 2004).

Model warna YCbCr. Model warna YCbCr memisahkan nilai RGB menjadi informasi *luminance* dan *chrominance* yang berguna untuk aplikasi kompresi. Transformasi RGB ke YCbCr dilakukan dengan formulasi operasi matriks pada persamaan berikut :

$$Y = 0.299900R + 0.58700G + 0.11400B$$

$$Cb = -0.16874R - 0.33126G - 0.50000B$$

$$Cr = 0.50000R - 0.41869G - 0.08131B$$

Warna kulit manusia. Jenis warna kulit manusia dikelompokkan menjadi lima jenis oleh Thomas B. Fitzpatrick pada penelitiannya. serta *brown* atau *black*. *Very fair* merupakan jenis warna kulit yang memiliki tingkat intensitas warna putih yang sangat terang dan cenderung pucat. Warna kulit ini meski terkena matahari tidak akan mencoklat atau menghitam. Masyarakat keturunan ras Eropa dan Australia memiliki kulit dengan warna jenis ini. Jenis warna kulit *Fair* cenderung lebih pink dan banyak muncul di jenis kulit bangsa Jepang dan Cina. Warna kulit fair terkadang bisa saja berubah menjadi gelap. Orang keturunan Indian Meksiko dan Indonesia pada umumnya memiliki warna kulit *medium*. Akan cukup mudah berubah menjadi gelap bila terpapar sinar matahari. Warna kulit *olive* walaupun jarang terpapar sinar matahari, warnanya akan tetap cokelat. Pada bangsa Afrika dan India warna kulit *brown* atau *black* banyak didapatkan. Warna kulit ini akan selalu menggelap atau hitam meskipun meskipun tidak pernah terkena sinar matahari langsung (Novintaclessence, 2015).

Confusion matrix. *Confusion matrix* berisi informasi tentang klasifikasi aktual dan prediksi yang dilakukan oleh sistem klasifikasi (Hamilton, 2012). Kinerja sistem menggunakan data dalam matriks untuk dievaluasi. Tabel berikut menunjukkan *confusion matrix*.

Tabel 1. confusion matrix

		Prediksi	
		Negative	Positive
Aktual	Negative	a	b
	Positive	c	d

Berikut merupakan penjelasan dari tabel *confusion matrix* diatas :
 Negatif a. adalah jumlah prediksi yang benar pada kelas negatif, positif b. adalah jumlah prediksi yang salah pada kelas positif, negatif

c. adalah jumlah yang salah prediksi pada kelas negatif, dan positif d. adalah jumlah prediksi yang benar pada kelas positif. Berikut merupakan Rumus untuk menghitung akurasi, *false positive rate* (FPR), dan *false negative rate*.

$$Accuracy = \frac{a+d}{a+b+c+d}$$

$$FPR = \frac{b}{a+b}$$

$$FNR = \frac{c}{c+d}$$

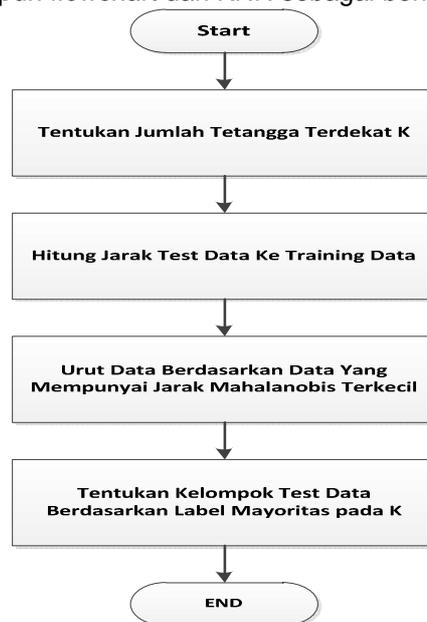
Accuracy : Persentase keberhasilan yang didapatkan. *FPR (False Positive Rate)*: proporsi dari nilai positif yang salah, di klasifikasikan sebagai nilai negatif. *FNR (False Negative Rate)*: proporsi dari nilai negatif yang salah diklasifikasikan sebagai nilai positif.

Algoritma KNN (K-Nearest Neighbor). KNN adalah suatu metode dimana hasil dari *query instance* yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada KNN dengan menggunakan algoritma *supervised* (Sikki, 2009).

Adapun tujuan dari algoritma ini yaitu mengklasifikasikan obyek baru berdasarkan atribut dan *training sample*. Untuk menentukan KNN-nya, algoritma KNN bekerja berdasarkan jarak terpendek dari *query instance* ke *training sample*. Dekat atau jauhnya tetangga bisa dihitung berdasarkan Mahalanobis Distance yang direpresentasikan sebagai berikut (Ker, 2010) :

$$D = (x_i - \mu)^T C^{-1} (x_i - \mu)$$

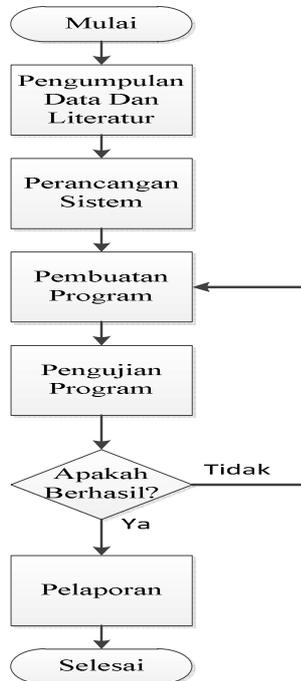
Adapun *flowchart* dari KNN sebagai berikut :



Gambar 1. Flowchart KNN

METODOLOGI PENELITIAN

Alur penelitian

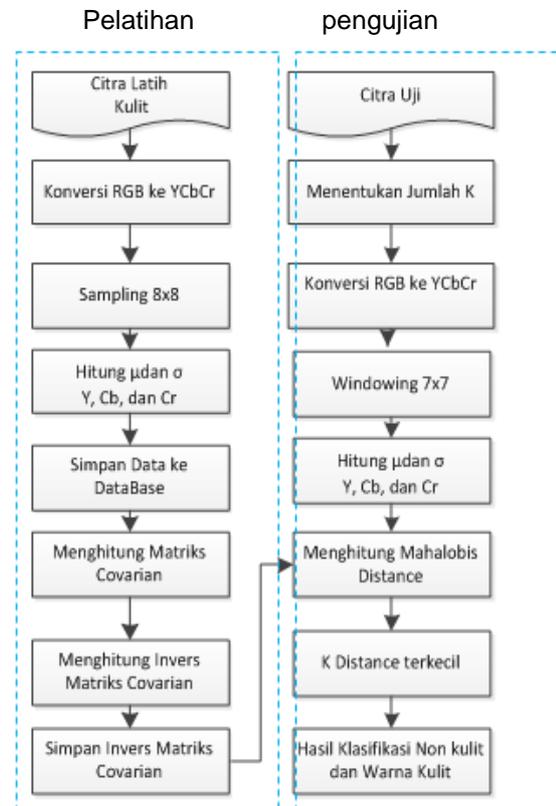


Gambar 2. Diagram alur penelitian

Dari Gambar 2 pada alur penelitian, dapat diketahui bahwa dari penelitian ini, dimulai dengan pengumpulan data dan literatur, dilanjutkan dengan perancangan sistem yang akan dibangun, pembuatan program serta pengujian program. Apabila program belum berhasil, maka proses akan kembali ke perancangan sistem yang di buat. Jika program yang di buat berhasil, maka akan dilanjutkan ke pembuatan laporan tugas akhir ini.

Pada blok pengumpulan data dan literatur, dilakukan pengumpulan data dan informasi yang digunakan sebagai acuan dalam menyelesaikan permasalahan. Blok selanjutnya yaitu perancangan sistem, pada blok ini dilakukan perancangan terhadap sistem yang akan di bangun. Setelah sistem telah dirancang, maka tahap selanjutnya yaitu merealisasikan rancangan yang telah dibuat, dengan membuat program. Apabila program sudah dibuat, maka program tersebut di uji, apakah sudah sesuai dari yang diharapkan atau belum. Jika program tersebut belum sesuai, maka proses akan kembali lagi ke proses perancangan sistem, tetapi jika program tersebut telah sesuai yang diharapkan maka akan dilanjutkan dengan pembuatan laporan mengenai penelitian yang telah dikerjakan.

Sistem klasifikasi kulit

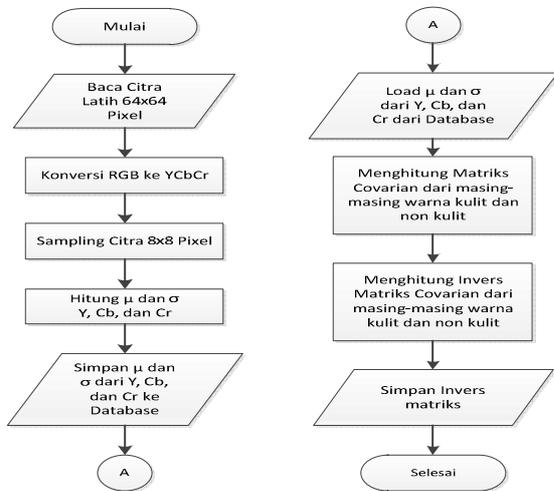


Gambar 3. Sistem klasifikasi kulit

Dari Gambar 3 terdapat 2 proses utama yaitu pelatihan dan pengujian. Proses pelatihan meliputi: Menyiapkan citra 5 warna kulit dan non kulit untuk dilatih. Masukkan citra yang berupa nilai RGB dikonversi ke nilai YCbCr. Melakukan sampling 8x8 pada citra masukan. Menghitung μ dan σ Y, Cb, dan Cr dari masing-masing sampling. Menyimpan bobot training μ dan σ Y, Cb, dan Cr kedalam database. Load bobot training untuk menghitung matriks kovarian training sampel. Menghitung invers dari matriks kovarian. Menyimpan nilai invers dari matriks kovarian.

Proses pengujian meliputi: Menentukan jumlah K (K=1,3,5,7,9,11,13,15, 17 dan 19) yang akan digunakan. Masukkan data citra yang akan diuji. Mengonversi nilai RGB pada citra ke nilai YcbCr. Melakukan windowing citra 7x7. Menghitung μ dan σ Y, Cb, dan Cr dari masing-masing window. Load nilai invers matriks kovarian dari data latih untuk menghitung mahalobis distance. Mengambil K distance terkecil untuk mengklasifikasi citra. Mendapatkan hasil klasifikasi warna kulit dari citra.

Proses pelatihan KNN

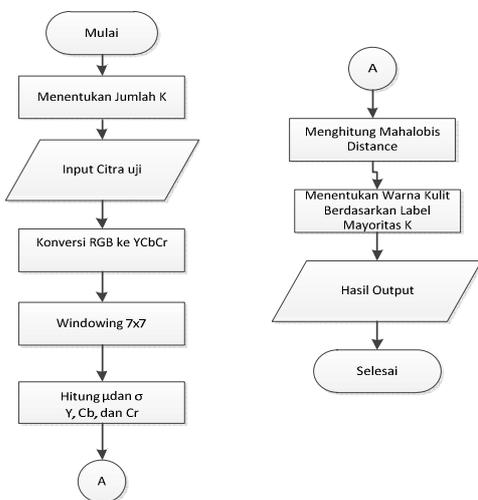


Gambar 4. Proses pelatihan KNN

Dari Gambar 4. dapat dilihat bahwa dalam proses pelatihan yaitu membaca citra latih 64x64 piksel. Masukan yang merupakan nilai RGB dikonversikan ke YCbCr. Melakukan sampling citra 8x8 piksel dan dihitung μ dan σ dari Y, Cb dan Cr dari masing-masing bagian.

Data μ dan σ dari Y, Cb dan Cr yang sudah didapat, disimpan kedalam *database*. Memanggil nilai μ dan σ dari Y, Cb dan Cr dari *database* untuk menghitung matriks kovarian dari 5 warna kulit dan bukan kulit. Menghitung invers matriks kovarian dari 5 warna kulit dan bukan kulit. Nilai invers yang didapatkan akan disimpan kedalam *file txt*.

Diagram Alir Pengujian Citra



Gambar 5. Diagram alir pengujian citra

Dari Gambar 5 dapat dilihat bahwa dalam proses pengujian citra, sebelum citra diuji terlebih dahulu kita tentukan jumlah K yang akan digunakan. citra input yang akan diuji terlebih dahulu dikonversi dari ruang warna RGB ke ruang warna YCbCr, lakukan proses *windowing* pada matriks citra, dalam hal ini penulis menggunakan metode *windowing* matriks 7x7 untuk melakukan pengecekan matriks gambar per piksel.

Setelah dilakukan *windowing*, hitung nilai μ dan σ Y,Cb, dan Cr dari masing-masing window. Nilai tersebut akan digunakan untuk menghitung *distance* pada *mahalanobis distance*. Setelah *distance*-nya didapatkan, maka diambil sejumlah K-*distance* terkecil. Sejumlah K-*distance* tersebut akan menentukan warna kulit atau bukan kulit dari citra uji.

Proses Windowing Citra. Pada proses pengenalan citra kulit dan bukan kulit ini dilakukan proses *filtering* citra dengan menggunakan filter *windowing*, filter *windowing* ini digunakan untuk mengecek masing-masing piksel dari citra yang dimasukkan. Adapun contoh cara kerja filter *windowing* adalah sebagai berikut:

```
int [][]A={{(5,4,7,6,5,8,6,7),(6,6,4,7,8,9,0,7),(8,6,7,1,2,5,6,10),(45,66,2,3,55,4,2,3),
{11,21,44,51,5,62,44,8),(5,6,3,7,8,2,4,11),(44,1,5,56,18,99,70,54),(8,5,3,9,6,4,3,3)}};
//windowing 7x7 dari array 2 dimensi
```

5	4	7	6	5	8	6	7
6	6	4	7	8	9	0	7
8	6	7	1	2	5	6	10
45	66	2	3	55	4	2	3
11	21	44	51	5	62	44	8
5	6	3	7	8	2	4	11
44	1	5	56	18	99	70	54

4	7	6	5	8	6	7	10
6	4	7	8	9	0	7	11
6	7	1	2	5	6	10	8
66	2	3	55	4	2	3	5
21	44	51	5	62	44	8	11
6	3	7	8	2	4	11	5
1	5	56	18	99	70	54	8

6	6	4	7	8	9	0	7
8	6	7	1	2	5	6	10
45	66	2	3	55	4	2	3
11	21	44	51	5	62	44	8
5	6	3	7	8	2	4	11
44	1	5	56	18	99	70	54
8	5	3	9	6	4	3	3

6	4	7	8	9	0	7	10
6	7	1	2	5	6	10	8
66	2	3	55	4	2	3	5
21	44	51	5	62	44	8	11
6	3	7	8	2	4	11	5
1	5	56	18	99	70	54	8
5	3	9	6	4	3	3	3

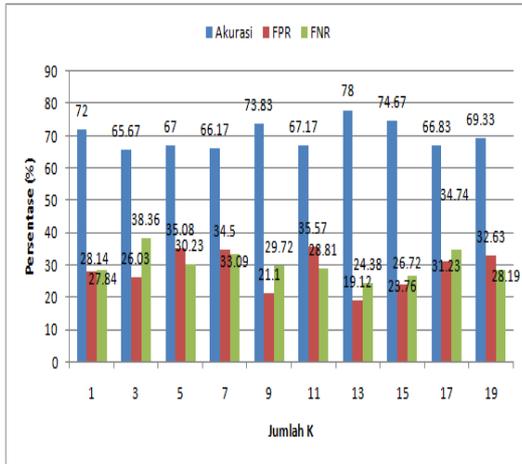
Gambar 6. Contoh windowing citra

Pada gambar 6. dapat dilihat proses *windowing* matriks, ukuran *windowing* yang digunakan yaitu 7x7. Proses *windowing* 7x7 yang tampak pada contoh diatas dicari nilai *mean* (rata-rata) dan standar deviasi dari masing-masing matriksnya terlebih dahulu. Nilai *mean* dan standar deviasi yang didapatkan dari matriks-matriks tersebut kemudian akan dijadikan sebagai nilai masukan untuk perhitungan KNN.

HASIL DAN PEMBAHASAN

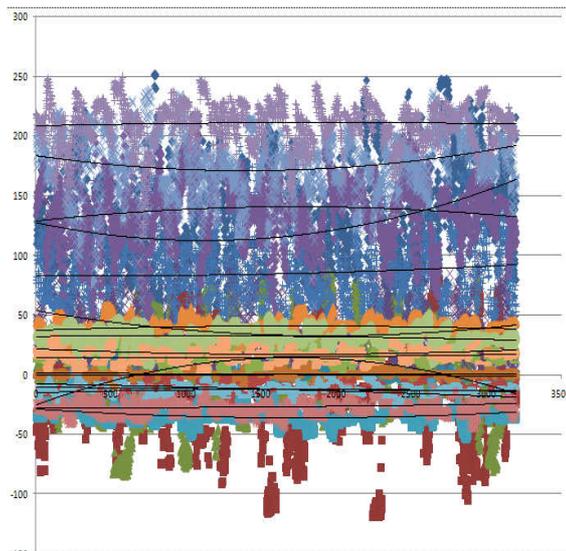
Hasil analisis *confusion matrix* dengan data latih 100 perkelas. Semua data latih di acak 50 sebagai data latih dan 50 sebagai data uji yang dilakukan disetiap kelas.

Berikut merupakan grafik akurasi, FPR dan FNR dari 10 variasi K :

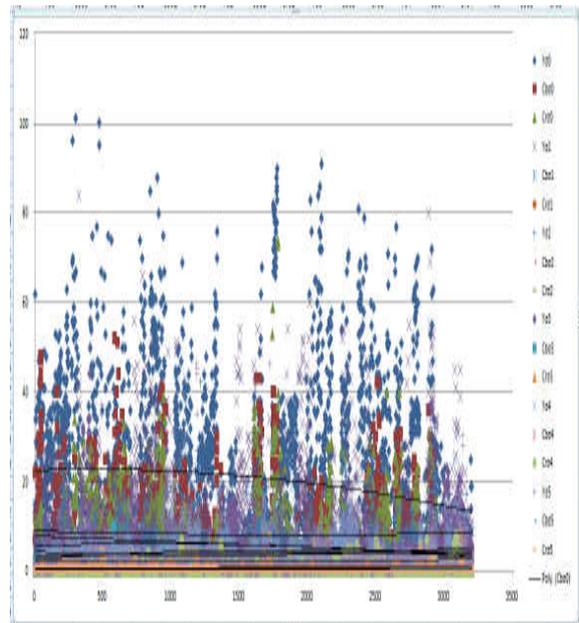


Gambar 7. Grafik ROC dari citra dengan 10 variasi K

Dari gambar 7 dapat dilihat hasil terbaik jatuh pada K=13 dengan nilai akurasi sebesar 78%, dengan nilai FPR 19.12% serta FNR 24.38. Nilai akurasi yang tidak terlalu tinggi disebabkan oleh penyebaran data dari satu kelas, memiliki data yang beririsan dengan data dari kelas lainnya.



Gambar 8. Penyebaran data mean Y,Cb, dan Cr masing-masing kelas



Gambar 9. Penyebaran data Standar Deviasi Y,Cb,Cr dari masing-masing kelas

Keterangan dari gambar 8 dan 9 :

- Y0, Cb0, Cr0 : data Cb dan Cr dari bukan kulit
- Y1, Cb1, Cr1 : data Cb dan Cr dari kelas hitam pekat
- Y2, Cb2, Cr2 : data Cb dan Cr dari kelas hitam
- Y3, Cb3, Cr3 : data Cb dan Cr dari kelas coklat
- Y4, Cb4, Cr4 : data Cb dan Cr dari kelas kuning langsung
- Y5, Cb5, Cr5 : data Cb dan Cr dari kelas putih

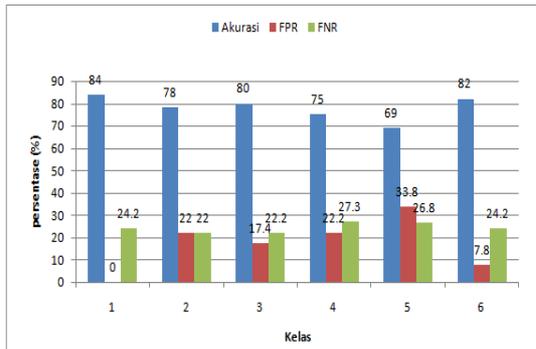
Dari gambar 8 dan 9 dapat dilihat penyebaran data dari masing-masing warna kulit dan bukan kulit. terjadi pengelompokan nilai dari setiap kelasnya dan ada nilai dari setiap kelas yang beririsan dengan kelas yang lain. Data diambil dari nilai μ dan σ data latih dari masing-masing kelas.

Hasil perhitungan *Confusion Matrix* dari tiap kelas dengan parameter K=13 adalah seperti pada tabel berikut:

Tabel (2) hasil kalkulasi dari K 13

Kelas	TP	FP	TN	FN	Akurasi (%)
Bukan kulit	50	0	34	16	84
Hitam Pekat	39	11	39	11	78
Hitam	42	8	38	12	80
Coklat	40	10	35	15	75
Kuning langsung	30	20	39	11	69
Putih	47	3	35	15	82
Rata-rata	41.33	8.67	36.67	13.33	78

Dari tabel (1) hasil kalkulasi dengan $K=13$ diatas, dapat dilihat bahwa akurasi terbaik terdapat pada kelas bukan kulit dengan akurasi sebesar 84%. Berdasarkan tabel, dapat dilihat grafik akurasi, FpR dan FnR dari K 13 pada gambar 4.17.

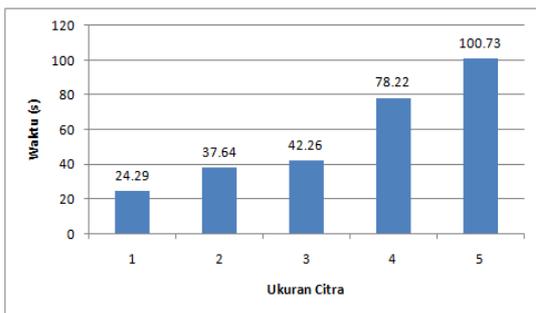


Gambar 10. Grafik akurasi, FPR dan FNR pada K 13

Keterangan kelas untuk Grafik di atas adalah : 1.Bukan kulit, 2.Hitam pekat, 3.Hitam, 4.Coklat, 5.Putih, 5. Kuning langsung

Dari gambar 4.17 dapat dilihat bahwa akurasi terbesar pada kelas bukan kulit yaitu sebesar 84%. FPR terkecil pada kelas bukan kulit dengan persentase sebanyak 0%. Sedangkan FNR terkecil pada kelas putih dengan persentase 7.89%.

Waktu komputasi. Berikut merupakan grafik dari waktu komputasi dari beberapa kombinasi ukuran citra yang digunakan.



Gambar 11. Grafik waktu komputasi citra

Keterangan ukuran citra pada gambar 11 yaitu :

1. Citra ukuran 64x64 piksel
2. Citra ukuran 100x100 piksel
3. Citra ukuran 200x200 piksel
4. Citra ukuran 300x300 piksel
5. Citra ukuran 400x400 piksel

Dari gambar 11 dapat dilihat bahwa waktu komputasi paling kecil yaitu citra ukuran 64x64 piksel dengan waktu komputasi

24.29 second. Sedangkan waktu komputasi paling lama yaitu pada citra ukuran 400x400 piksel dengan waktu yang dibutuhkan 100.73 second. Ukuran citra dapat mempengaruhi waktu komputasi dikarenakan semakin besar citra uji, semakin banyak perhitungan yang terjadi sehingga menyebabkan semakin lama waktu komputasi.

KESIMPULAN

Dari penelitian yang telah dilakukan, maka dapat disimpulkan bahwa:

1. Penentuan warna kulit cukup efektif dengan menggunakan metode KNN dengan dengan akurasi sebesar 78% , dengan nilai FPR 19.12% serta FNR 24.38 pada $K=13$.
2. Hasil dari kalkulasi pada $K=13$, kelas kulit hitam pekat memiliki akurasi sebesar 78%, nilai FPR 22%, nilai FNR 22%. Pada kulit hitam memiliki akurasi sebesar 80%, nilai FPR 17.4%, dan nilai FNR 22.2%. Kulit coklat memiliki akurasi sebesar 75%, nilai FPR 22.2%, dan nilai FNR 27.3%. Kulit kuning langsung memiliki akurasi sebesar 69% , nilai FPR 33.8% , dan nilai FNR 26.8%. Kulit putih memiliki akurasi 82%, nilai FPR 7.8%, dan nilai FNR 24.2% .

SARAN

1. Dalam melakukan proses pengkalsifikasian warna kulit menggunakan algoritma K-NN ini dibutuhkan spesifikasi komputer yang cukup baik, agar memperlancar proses komputasi.
2. Dalam melakukan pemilihan pada data pelatihan sebaiknya lebih bervariasi dan dilakukan pengecekan tiap kelasnya, agar di kelas yang 1 dengan kelas yang lain tidak memiliki kesamaan nilai.

DAFTAR PUSTAKA

Hamilton, H.J., 2012, *Confusion Matrix*, tersedia di http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html, diakses 2 Juli 2016.

Ker, E.D., 2010, *Stability Of The Mahalanobis Distance:A Technical Note*, Oxpord University Computing Labolatory, Oxpord.

Munir, R., 2004, *Pengolahan Citra Digital dengan Pendekatan Algoritmik*, Informatika, Bandung.

Novintaclenese, 2015, *Warna Kulit Manusia*, tersedia di <http://novintatutorial.blogspot.co.uk/2015/12/warna-kulit-manusia.html>, diakses 2 Juli 2016.

Sikki, M. I., 2009, *Pengenalan Wajah Menggunakan K-Nearest Neighbour dengan Praproses Transformasi Wavelet*, Jurnal paradigm vol x. no. 2, pp.159-172.