

IDENTIFIKASI BIOMETRIK POLA ENAMEL GIGI MENGGUNAKAN METODE GREY LEVEL RUN LENGTH MATRIX (GLRLM) DAN KLASIFIKASI K-NEAREST NEIGHBOR (K-NN) SEBAGAI APLIKASI FORENSIK KEDOKTERAN GIGI BERBASIS MATLAB

BIOMETRIC IDENTIFICATION OF TOOTH ENAMEL PATTERN USING GREY LEVEL RUN LENGTH MATRIX (GLRLM) METHOD AND K-NEAREST NEIGHBOR (K-NN) CLASSIFICATION AS A MATLAB-BASED APPLICATION FOR FORENSIC DENTISTRY

Diny Hafizha Amelia¹, Dr. Ir. Bambang Hidayat, DEA², drg. Yuti Malinda, MM, M.Kes³

^{1,2} Prodi S1 Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom

³ Fakultas Kedokteran Gigi, Universitas Padjajaran

¹ dinyhafizha.amelia@gmail.com, ² bhidayat@telkomuniversity.ac.id, ³ yuti.malinda@fkg.unpad.ac.id

Abstrak

Enamel gigi merupakan bagian penyusun terluar dari gigi. Tiap individu memiliki pola enamel gigi yang berbeda. Unikunya pola enamel gigi tiap individu menjadikannya sebagai aspek penting dalam proses identifikasi. Identifikasi pola enamel gigi bertujuan untuk mempermudah dalam mengidentifikasi korban dalam suatu kecelakaan atau bencana.

Tugas akhir ini bertujuan untuk menghasilkan sistem yang dapat mengidentifikasi pola enamel pada gigi secara personal. Metode yang digunakan untuk mendapatkan nilai ekstraksi ciri pada sistem ini yaitu metode *Grey Level Run Length* (GLRL) yang memiliki 5 ciri analisis tekstur yaitu SRE (*Short Run Emphasis*), LRE (*Long Run Emphasis*), RLU (*Run Length Uniformity*), GLU (*Grey Level Uniformity*), dan RPC (*Run Percentage*). Digunakan 10 gigi sebagai sampel citra yang akan di akuisisi. Kemudian citra tersebut diubah ke *grayscale* lalu dilakukan ekstraksi ciri dan dilakukan proses pengenalan citra dengan klasifikasi *K-Nearest Neighbor*.

Hasil yang diperoleh dari tugas akhir ini adalah sebuah aplikasi berbasis Matlab dengan nilai akurasi mencapai 84 % dalam waktu komputasi rata-rata 0,7707 detik. Sistem menggunakan data latih yang terdiri dari 100 citra dan data uji sebanyak 200 citra dari 10 *sample* gigi.

Kata kunci: pola enamel gigi, *Grey Level Run Length*, *K-Nearest Neighbor*

Abstract

Dental enamel is the outermost constituent part of the tooth. Each individual has a different tooth enamel pattern. The unique pattern of individual tooth enamel makes it an important aspect for the identification process. Identification of tooth enamel patterns aims to make it easier to identify victims in an accident or disaster.

This final task aims to produce a system that can identify the enamel pattern on the teeth personally. The method used to obtain the extraction value of this system is the Gray Level Run Length (GLRL) method which has 5 characteristics of texture analysis i.e. SRE, LRE, RLU, GLU, and RPC. The sample images were taken from 10 teeth for the data acquisition. Then the images were converted to grayscale and then performed feature extraction and carried out image recognition process with the K-Nearest Neighbor classification.

The result obtained from this final task is a Matlab-based application with an accuracy rate 84% and average computation time 0.7707 seconds. The system uses training data consisting of 100 images and 200 images for the testing data from 10 teeth samples.

Keywords : dental enamel pattern, *Grey Level Run Length*, *K-Nearest Neighbor*

1. Pendahuluan

Forensik merupakan salah satu bidang pengetahuan yang dapat digunakan untuk membantu proses identifikasi seseorang baik itu berdasarkan pada usia, ras ataupun jenis kelamin. Salah satu inovasi ilmu forensik yang dapat digunakan adalah sidik gigi atau pola enamel gigi karena polanya yang unik dan berbeda untuk tiap individu bahkan pada saudara kembar sekalipun.

Pola enamel gigi tidak dapat terlihat jelas hanya dengan menggunakan mata telanjang. Dalam hal ini, *image processing* dapat membantu mempercepat proses identifikasi pola enamel gigi dengan cara mendeteksi pola pada foto dari salah satu gigi, lalu dengan menggunakan metode *Grey Level Run Length Matrix* dan klasifikasi *K-Nearest Neighbor* akan didapatkan pola enamel gigi yang sesuai. Gigi *incisivus* atas dipilih karena dinilai memiliki permukaan paling luas dibandingkan gigi jenis lain sehingga pola enamel akan lebih mudah terlihat pada gigi tersebut.

Metode *Grey Level Run Length Matrix* dipilih karena dinilai dapat mengekstraksi ciri dengan tingkat akurasi yang baik karena metode tersebut dapat digunakan untuk membedakan citra halus dan citra kasar, dimana citra

yang memiliki tekstur halus akan memiliki sedikit piksel tetangga yang intensitasnya sama. Sedangkan citra yang memiliki tekstur kasar akan memiliki banyak piksel tetangga yang intensitasnya sama. Selain itu, klasifikasi dilakukan dengan *K-Nearest Neighbor* karena dapat mengklasifikasi suatu objek berdasarkan data yang jaraknya paling dekat dengan objek yang diteliti. Cara tersebut menghasilkan akurasi sistem yang baik.

2. Dasar Teori dan Perancangan

1. Gigi dan Enamel

Gigi merupakan bagian terkeras dari seluruh tubuh manusia. Gigi pada manusia terbagi menjadi tiga bagian, yaitu bagian mahkota, bagian leher dan bagian akar. Bagian mahkota merupakan bagian gigi yang terlihat dalam mulut yang memiliki lapisan email atau enamel, sedangkan bagian akarnya merupakan bagian gigi yang biasanya tidak terlihat karena tertanam dalam tulang rahang. [1]

Pola enamel disebut juga sebagai sidik gigi. Bila diamati dengan mikroskop akan terlihat struktur enamel yang mirip dengan sarang lebah (heksagonal) yang dinamakan prisma/batang enamel. Prisma enamel ini berfungsi seperti serat optik apabila diarahkan oleh cahaya secara langsung. Sehingga, jika diamati dari luar, permukaan enamel tersebut seperti memiliki garis-garis yang mirip dengan sidik jari dan menjadikan pola enamel disebut juga sebagai sidik pada gigi.

2. Pengolahan Citra Digital

Pengolahan citra digital merupakan metode untuk mengolah citra digital sehingga menghasilkan gambar lain yang sesuai dengan kebutuhan masing-masing [2]. Citra yang diolah merupakan citra digital yang merupakan sekumpulan bilangan yang direpresentasikan oleh bit berhingga.

Dalam tugas akhir ini pengolahan citra digital akan lebih dititikberatkan pada *image analysis*. Dengan menggunakan *image analysis* maka ekstraksi ciri-ciri tertentu yang dimiliki citra akan diambil untuk membantu dalam pengidentifikasian objek.

Terdapat beberapa tipe dasar citra digital, di antaranya:

2.2.1 Citra RGB

Red (Merah), *Green* (Hijau) dan *Blue* (Biru) merupakan warna dasar yang dapat diterima oleh mata manusia. Setiap piksel pada citra warna mewakili warna yang merupakan kombinasi dari ketiga warna dasar RGB. Setiap titik pada citra warna membutuhkan data sebesar 3 byte [8]. Setiap warna dasar memiliki intensitas tersendiri dengan nilai minimum nol (0) dan nilai maksimum 255 (8 bit). RGB didasarkan pada teori bahwa mata manusia peka terhadap panjang gelombang 630 nm (merah), 530 nm (hijau), dan 450 nm (biru) [3].

2.2.2 Citra Biner

Citra biner adalah citra digital yang hanya memiliki dua kemungkinan nilai piksel, yaitu hitam dan putih. Citra biner juga disebut sebagai citra B&W (*black and white*) atau citra monokrom. Hanya dibutuhkan 1 bit untuk mewakili nilai setiap piksel dari citra biner [5]. Citra biner sering kali muncul sebagai hasil dari proses pengolahan seperti segmentasi, pengembangan, morfologi ataupun *dithering*.

2.2.3 Citra Grayscale

Citra *grayscale* merupakan citra digital yang hanya memiliki satu nilai kanal pada setiap pikselnya, artinya nilai dari *Red = Green = Blue* [2]. Nilai-nilai tersebut digunakan untuk menunjukkan intensitas warna. Citra *grayscale* disimpan dalam format 8 bit untuk setiap sample piksel yang memungkinkan sebanyak 256 intensitas [2].

3. Grey Level Run Length Matrix (GLRLM)

GLRLM merupakan metode analisis pengambilan ciri dengan mengestimasi piksel-piksel yang memiliki derajat keabuan yang sama. GLRLM membuat rangkaian pasangan nilai (i,j) pada tiap baris dari piksel/matriks. *Run length* sendiri merupakan jumlah piksel berurutan dalam arah tertentu (0°, 45°, 90°, dan 135°) yang punya derajat keabuan / nilai intensitas yang sama. Analisis tekstur dengan metode GLRLM digunakan untuk membedakan citra halus dan citra kasar.

Setelah mendapat matriks nilai pada piksel, maka nilai tersebut dapat dilihat berdasarkan ciri matriks dengan menggunakan persamaan, sebagai berikut:

Tabel 2.1 Persamaan Ekstraksi Ciri GLRLM

No	Nama	Persamaan
1	SRE (<i>Short Run Emphasis</i>)	$\frac{1}{K} \sum_{a=1}^L \sum_{r=1}^{Nr} \frac{B(a, r)}{r^2}$
2	LRE (<i>Long Run Emphasis</i>)	$\frac{1}{K} \sum_{a=1}^L \sum_{r=1}^{Nr} r^2 B(a, r)$

3	GLU (<i>Grey Level Uniformity</i>)	$\frac{1}{K} \sum_{a=1}^L \left(\sum_{r=1}^{Nr} B(a,r) \right)^2$
4	RLU (<i>Run Length Uniformity</i>)	$\frac{1}{K} \sum_{r=1}^{Nr} \left(\sum_{a=1}^L B(a,r) \right)^2$
5	RPC (<i>Run Percentage</i>)	$\frac{K}{\sum_{a=1}^L \sum_{r=1}^{Nr} rB(a,r)} = \frac{K}{MN}$

4. K-Nearest Neighbor

K-Nearest Neighbor atau K-NN adalah metode yang digunakan untuk mengklasifikasikan suatu data. Pada data latih biasanya diambil lebih dari satu tetangga terdekat dengan data uji kemudian akan digunakan algoritma ini untuk ditentukan kelasnya [4]. Tujuan dari algoritma ini adalah untuk mengklasifikasikan obyek baru berdasarkan atribut dan sampel-sampel dari data latih. K-NN bekerja berdasarkan jarak minimum dari data baru ke sample data latih untuk menentukan K tetangga terdekat. Setelahnya akan didapatkan nilai mayoritas sebagai hasil prediksi dari data yang baru tersebut.

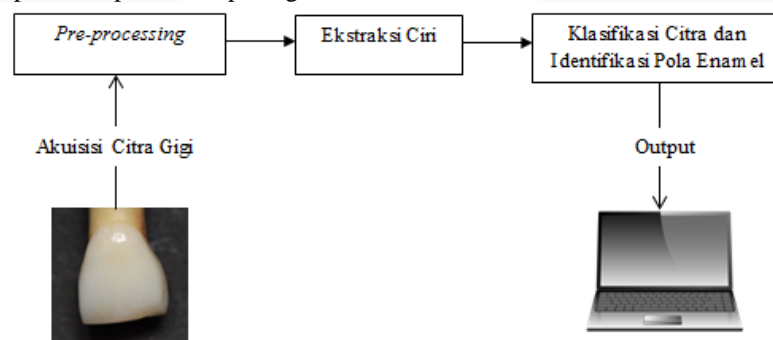
Pada K-NN terdapat beberapa rumus aturan jarak yang dapat digunakan jika terdapat masalah dimana penentuan kelas dengan voting mayoritas sederhana tidak bisa dilakukan, yaitu:

Tabel 2.2 Rumus Aturan Jarak pada K-NN

No	Nama	Persamaan
1	<i>Euclidean Distance</i>	$L_2(X, Y) = \sqrt{\sum_{i=1}^d (X_i - Y_i)^2}$
2	<i>City Block</i>	$L_1(X, Y) = \sum_{i=1}^d X_i - Y_i $
3	<i>Cosine Similarity</i>	$\cos(d_i, d_j) = \frac{\sum_k a_{i,k} \cdot a_{j,k}}{\sqrt{\sum_k a_{i,k}^2} \sqrt{\sum_k a_{j,k}^2}}$
4	<i>Correlation Distance</i>	$S_{ij} = \frac{\sum_{k=1}^n (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)}{\left(\sum_{k=1}^n (x_{ik} - \bar{x}_i)^2 \sum_{k=1}^n (x_{jk} - \bar{x}_j)^2 \right)^{1/2}}$

5. Perancangan Sistem

Dalam model sistem ini dijelaskan tentang alur dalam proses identifikasi pola enamel gigi. Secara umum tahapan proses dapat dilihat pada gambar berikut:

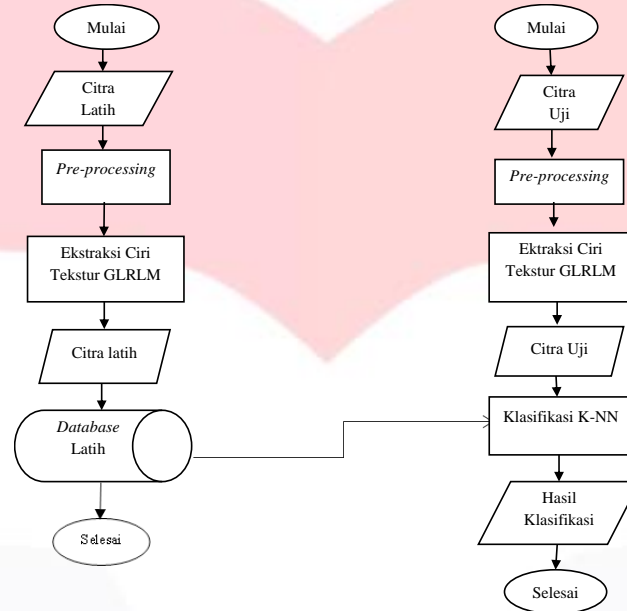


Gambar 2.1 Perancangan Sistem Secara Umum

Pada tugas akhir ini, dirancang sistem perangkat lunak yang terdiri atas dua tahapan, yaitu tahap pelatihan dan pengujian. Pada tahap pelatihan serta pengujian, data citra dimasukkan ke dalam perangkat dan kemudian digunakan metode *Grey Level Run Length Matrix* untuk segmentasinya yang kemudian hasil segmentasi tersebut digunakan untuk keperluan ekstraksi ciri dan proses klasifikasi dengan *K-Nearest Neighbor*.

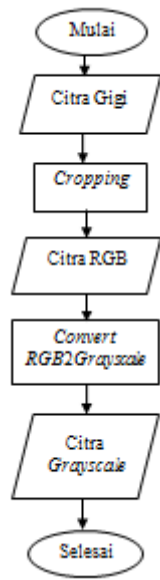
Pada tahap pelatihan, proses pertama yang dilakukan adalah proses *pre-processing* terhadap citra latih yang dimasukkan, kemudian GLRLM digunakan untuk memperoleh beberapa fitur dari data latih

yang dijadikan *database* untuk keperluan klasifikasi pada proses pengujian. Sedangkan pada tahapan pengujian digunakan untuk menguji data citra uji sehingga dapat diklasifikasi oleh perangkat lunak. Tahapan pada tahap pengujian terdiri atas *pre-processing* terhadap data citra uji yang dimasukkan, kemudian ekstraksi ciri lalu masuk ke dalam tahapan klasifikasi. Ciri latih yang diperoleh dari tahap pelatihan digunakan sebagai masukan pada tahap klasifikasi, kemudian dari hasil klasifikasi ini dibuat sebuah keputusan yang mengenai pola enamel serta pemilik dari pola enamel tersebut. Untuk lebih jelasnya, dapat dilihat pada gambar berikut ini:



Gambar 2.2 Blok Diagram Citra Latih dan Citra Uji

Untuk proses ekstraksi ciri dengan GLRLM dapat dilihat pada blok diagram di bawah ini:



Gambar 2.3 Diagram Alir Ekstraksi Ciri dengan GLRLM

Setelah didapatkan ciri tekstur dari citra gigi, dilakukan klasifikasi dengan K-NN untuk dapat mengidentifikasi identitas dari pemilik ciri berdasarkan hasil ekstraksi ciri sebelumnya. Pada penelitian ini, dilakukan pemilihan kelas klasifikasi dengan cara mencari kelas terdekat dari suatu data latih. Metode ini digunakan karena dinilai handal dalam meminimalisasi kesalahan pada proses pengklasifikasian. Selain itu, K-NN merupakan metode klasifikasi yang sederhana, sehingga waktu komputasi yang dihasilkan relatif singkat, namun tetap menghasilkan hasil yang maksimal.

3. Hasil Analisis Sistem

3.1 Pengujian Skenario 1

Pada skenario pertama, dilakukan pengujian terhadap metode GLRLM untuk mendapatkan nilai dari kelima ciri yang dimiliki oleh metode tersebut. Ciri tersebut yaitu SRE, LRE, GLU, RLU, dan RPC. Adapun tujuan dari mendapatkan nilai-nilai tersebut adalah untuk mengetahui karakteristik ciri dari tiap sampel/kelas yang diujikan. Berikut merupakan hasil perolehan dari masing-masing ciri bila dibandingkan pada citra yang diakuisisi dengan bantuan lampu dan tanpa lampu:

Tabel 3.1 Nilai *Mean* Ciri GLRLM

Ciri	Lampu	Tanpa Lampu
SRE	0,4856	0,4926
LRE	4,3909	4,3333
GLU	$1,2886 \times 10^{11}$	$1,3959 \times 10^{11}$
RLU	$4,3893 \times 10^4$	$4,3104 \times 10^4$
RPC	0,518	0,522

Dari hasil pengujian tersebut didapatkan kesimpulan bahwa nilai GLU dan RLU sangat besar dibandingkan nilai ciri yang lain. Selain itu, penggunaan lampu pada proses akuisisi citra berpengaruh terhadap nilai ekstraksi ciri, karena dengan digunakannya lampu dapat memperlihatkan tekstur kasar dari pola enamel gigi secara lebih jelas.

3.2 Pengujian Skenario 2

Pada skenario kedua, dilakukan pengujian terhadap akurasi sistem dengan menggunakan parameter K-NN yaitu jarak dan nilai k. Selain itu, pengujian ini membandingkan derajat keabuan/*grey level* yang digunakan pada sistem. Hal ini bertujuan untuk mengetahui level keabuan mana yang paling baik untuk sistem sehingga dapat menghasilkan sistem yang cukup handal. Berikut merupakan hasil dari pengujian akurasi sistem dalam mengklasifikasikan pola enamel gigi sesuai dengan kelasnya dilihat dari parameter aturan jarak K-NN, penggunaan nilai k, serta derajat keabuan:

Tabel 3.2 Nilai Akurasi dengan Parameter K-NN dan Derajat Keabuan GLRLM

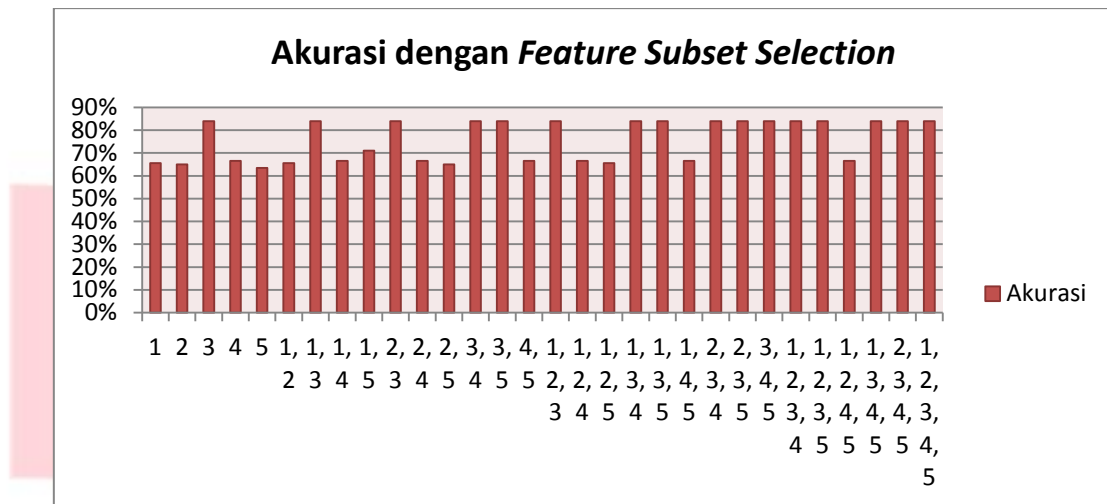
K-NN	Akurasi (%)											
	16				8				4			
	K=1	K=3	K=5	K=7	K=1	K=3	K=5	K=7	K=1	K=3	K=5	K=7
<i>Euclidean</i>	84	72	67	58	59,5	57,5	52	48	58,5	58	48,5	41
<i>City Block</i>	84	72	67	58	59,5	57,5	52	48	58,5	58	48,5	41
<i>Cosine</i>	43,5	42	47,5	50,5	27	31	30	31,5	18	19	19	20,5
<i>Correlation</i>	37	52,5	56	54	27	24,5	28,5	32	21	22	20	20,5

Dari hasil pengujian tersebut, terlihat bahwa akurasi tertinggi diperoleh dengan derajat keabuan = 16 dan pada saat memakai aturan jarak *euclidean* atau *city block* dengan nilai k = 1. Sedangkan, akurasi terendah diperoleh dengan derajat keabuan = 4 dan pada saat memakai aturan jarak *cosine* dengan nilai k = 1.

3.3 Pengujian Skenario 3

Pengujian pada skenario ketiga dilakukan dengan *Feature Subset Selection*. Fungsi dari *Feature Subset Selection* ini adalah untuk mengetahui ciri mana yang paling dominan pada sistem. Hal tersebut dilakukan dengan cara mencari nilai akurasi untuk masing-masing ciri yang digunakan, kemudian tiap-tiap ciri dikombinasikan satu sama lain sehingga dapat diketahui ciri mana yang paling berpengaruh dan dapat menghasilkan nilai akurasi tertinggi.

Pada tugas akhir ini, ciri dari GLRL yang digunakan yakni SRE, LRE, GLU, RLU, dan RPC yang selanjutnya akan dibaca pada tabel sebagai ciri 1, 2, 3, 4, dan 5 secara berurutan. Hasil dari pengujian *Feature Subset Selection* dapat dilihat pada Gambar 3.1 berikut:



Gambar 3.1 Akurasi dengan *Feature Subset Selection*

Berdasarkan Gambar 3.1 di atas, terlihat bahwa nilai akurasi tertinggi yang dihasilkan sistem adalah pengaruh dari ciri nomor 3 yaitu GLU (*Grey Level Uniformity*) yang menghasilkan akurasi sebesar 84 %. Sedangkan, nilai akurasi terendah yang dihasilkan sistem adalah dengan penggunaan ciri nomor 5 yaitu RPC (*Run Percentage*).

4. Kesimpulan

1. Sistem ini dapat direalisasikan dengan mengikuti prosedur yang telah dirancang sebelumnya dan ekstraksi ciri dengan menggunakan metode GLRLM serta klasifikasi dengan K-NN dapat digunakan dalam pembuatan sistem identifikasi biometrik pola enamel gigi berbasis Matlab.
2. Analisis dari penelitian ini yaitu sistem dapat mengidentifikasi pola enamel gigi secara personal dengan menggunakan metode ekstraksi ciri GLRLM dan klasifikasi K-NN dengan nilai akurasi klasifikasi citra latihan 100 %, yang berarti sistem dapat bekerja dengan sangat baik bila citra yang diujikan sama dengan citra dalam database pelatihan. Selain itu, untuk pengujian terhadap citra uji, diperoleh akurasi sebesar 84 %.
3. Parameter yang dapat menghasilkan nilai akurasi terbaik yaitu sebesar 84 % adalah dengan nilai derajat keabuan = 16, $k = 1$, dan pada aturan jarak euclidean atau cityblock dengan waktu komputasi 0,7707 detik dan 0,7290 detik.

DAFTAR REFERENSI

- [1] Vickie P. Overman, *Introduction to Dental Anatomy*, Crest + Oral-B at dentalcare.com Continuing Education Course, 2001.
- [2] Darma, Putra. 2010. *Pengolahan Citra Digital*. Yogyakarta: Penerbit Andi
- [3] Hestningsih, Idhawati. (2012) *Pengolahan Citra*. [Online] tersedia di <http://www.scribd.com/doc/33661393/Pengolahan-Citra>. [diakses 1 Maret 2017]
- [4] Pdraig. Cunningham, and Sarah Jane Delany, "*K-Nearest Neighbor Classifier*". (-): Technical Report UCD-CSI, vol. 4, pp. 1-2, 2007.
- [5] Agus Prijono dan Marvin Ch. Wijaya, 2007. *Pengolahan Citra Digital Menggunakan Matlab Image Processing Toolbox*. Bandung: Informatika.