

# PERBANDINGAN ALGORITMA K-MEANS DAN FUZZY C-MEANS UNTUK PENGELOMPOKAN KEYPOINT DESCRIPTOR DALAM CITRA SIDIK JARI

Indra Apriyanto, Catur supriyanto

Program studi teknik informatika S1, fakultas Ilmu computer

Universitas dian nuswantoro

Jl. Nakula I No, 5-11 Semarang 50131

Telp : (024) 3517261, Fax : (024) 3520165

Email : indraapri82@gmail.com

## ABSTRAK

*Memasuki era globalisasi, diperlukan informasi yang cepat, tepat dan akurat untuk kemajuan disegala bidang. Demikian juga dalam pengolahan citra, pemrosesan citra khususnya menggunakan komputer menjadikan citra yang kualitasnya lebih baik. Analisis kelompok (cluster analysis) merupakan salah satu teknik dalam analisis multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek pengamatan menjadi beberapa kelompok berdasarkan karakteristik yang dimilikinya. Dalam ilmu pengolahan citra ada sebuah metode untuk melakukan recognition sebuah citra yang dimana dapat dilakukan dalam jumlah data yang banyak. Sehingga untuk melakukan identifikasi citra dapat dilakukan dengan mudah. Metode yang digunakan untuk recognition citra adalah Bag of visual words. Proses metode Bag of visual words ini memerlukan sebuah clustering terhadap beberapa keypoint yang dihasilkan dari sebuah algoritma matching point. Kemudian setelah keypoint dikelompokkan, maka keypoint tersebut akan di digunakan untuk proses classification. Tugas akhir ini hanya bertujuan untuk mengetahui kinerja algoritma clustering dalam pengelompokan keypoint untuk pembentukan bag of visual words. Fuzzy c-means (FCM) dan K-means dipilih untuk proses clustering tersebut.*

*Kata kunci : Bag of visual word, Clustering, Keypoint, Fuzzy c-means, K-means*

## ABSTRACT

*Entering of the globalisation era, required information quickly and accurately to progress in all fields. Likewise, in image processing, image processing, particularly using a computer made image quality was better. Analysis group (cluster analysis) was one technique in multivariate analysis that had the main purpose to classify the objects of observation into groups based on characteristics they had. In the science of image processing there was a method to perform an image recognition which could be done in the amount of data that a lot. With the result that identify the image could be done easily. Methods used for image recognition is a bag of visual words. Bag method of visual words process required a -clustering toward several keypoint generated from a point matching algorithm. Therefore, after keypoint grouped, then the keypoint would be used for the classification process. The final task was only intended to determine the performance of clustering algorithm for the formation of the grouping keypoint bag of visual words. Fuzzy c-means (FCM) and K-means clustering process chosen for them*

*Keywords: Bag of visual word, Clustering, Keypoint, Fuzzy c-means, K-means*

## 1. PENDAHULUAN

Sidik jari merupakan salah satu variable penting yang digunakan untuk menegakkan keamanan dan memelihara identifikasi yang dapat diandalkan setiap individu dan digunakan sebagai variable keamanan, pemeriksaan, operasi rekening bank, untuk mengontrol akses ketempat-tempat yang sangat aman seperti kantor, ruang peralatan, pusat control dan sebagainya [1]. Sidik jari dipercaya memiliki sifat yang unik dari seluruh individu dan seluruh individu yang sama. Bahkan kembar identik pun yang memiliki DNA sama, namun memiliki sidik jari yang berbeda. Segmentasi dalam pengambilan gambar sidik jari terdiri dari dua bagian yaitu bagian depan dan bagian belakang. Dalam modus identifikasi sidik jari, penggunaan input sidik jari dan sistem mengidentifikasi seseorang yang sesuai dengan potensi dalam database [2]. Banyak informasi yang bisa kita dapatkan mengenai sifat sidik jari antara lain : layak (*feasible*), berbeda satu sama lain (*distinct*), tetap (*penent*), akurat (*accurate*), handal (*reliable*) dan dapat diterima (*acceptable*).

Dalam beberapa tahun terakhir, ketertarikan pada sidik jari berdasarkan sistem biometrik telah tumbuh secara signifikan. Tentu saja harapan adanya sistem identifikasi yang cepat dan tepat bukan hanya untuk pencarian dan pembuktian pelaku kejahatan, tetapi juga diharapkan dapat menjadi pengganti cara-cara pengamanan konvensional misalnya dengan password dan PIN, Akan tetapi ada masalah dengan menggunakan keamanan yang bersifat konvensional [3].

Dalam penelitian ini kita menggunakan gambar yang berbeda dan dikombinasikan Fuzzy C-mean (FCM) pengelompokan algoritma dengan segi statistik untuk segmentasi gambar sidik jari. Ini adalah teknik pengelompokan tanpa pengawasan yang telah banyak digunakan dalam segmentasi gambar. Menunjukkan sebuah sekat kabur dari kumpulan data yang diberikan. Keuntungan dari FCM pengelompokan algoritma yang pelaksanaannya langsung, ini adalah tindak tanduk yang tegap/kuat, penerapannya data multichannel dan ketidakpastian kemampuan pemodelan data, juga memiliki kemampuan

untuk menjaga informasi lebih lanjut dari gambar asli [4].

Keypoint descriptor merupakan salah satu algoritma yang bekerja cukup baik dalam mendeteksi ciri pada suatu citra, output dari suatu algoritma ini berupa titik-titik kunci yang berada disekitar pola dari citra [5].

Dalam kinerja pengambilan keypoint penulis menggunakan SURF (speeded-up robust features) dibandingkan SIFT (scale invariant feature transform), terbukti menjadi salah satu yang cocok [6].

Clustering adalah pengelompokan data yang heterogen menjadi subkelompok homogen yang disebut “cluster” [7]. Sebuah cluster adalah sebuah koleksi yang terdiri dari data yang homolog tetapi berbeda dari komponen klaster lainnya. Algoritma *clustering* dibagi menjadi dua kategori utama: pengelompokan hirarki dan partitional (*flat*).

*Bag of visual word (BOVW)* adalah kumpulan *keypoint* yang di ekstraksi sebagai *patch* gambar menonjol dan sering digunakan dalam klasifikasi data citra [8]. Dalam proses BOVW ada beberapa tahap yang harus dilakukan yaitu : 1) *Extraction keypoint*, 2) *Clustering keypoint*, 3) *Visual-word vocabulary*.

Dan berawal dari semua latar belakang yang telah dibahas sebelumnya, maka penulis tertarik untuk mencoba melakukan eksperimen mengevaluasi peforma dari clustering keypoint dengan menggunakan algoritma fuzzy c-means. Kemudian melakukan analisa dari hasil peforma untuk dibandingkan dengan algoritma k-means.

## 2. METODE

### 2.1 Metode Pengumpulan Data

Data yang digunakan dalam penelitian diperoleh dari internet dengan format \*.tif dengan ukuran piksel yang bervariasi. Data uji yang digunakan berjumlah 4 dataset FVC 2004<sup>1</sup> DB1, DB2, DB3 dan DB4.

### 2.2 Pengukuran Kinerja Model

*Residual Sum of Square (RSS)* Teknik statistik yang digunakan untuk mengukur jumlah variasi dalam suatu kumpulan data yang tidak dijelaskan oleh model regresi. Sisa jumlah kuadrat adalah ukuran dari jumlah yang tersisa antara fungsi regresi dan kumpulan data error. Sebuah sisa jumlah yang lebih kecil dari angka kotak merupakan fungsi regresi yang menjelaskan sejumlah besar data. [7]

$$RSS = \sum_{k=1}^K RSS_k \text{ where } RSS_k = \sum_{x \in \omega_k} |x - \mu(\omega_k)|^2 \quad (2.1)$$

contoh evaluasi atau perhitungan RSS

keterangan :

$K$  : jumlah *cluster*

$x$  : *keypoint*

$\mu$  : centroid

$\omega$  : *Cluster*

dimana  $K=1$  adalah jumlah *cluster*, dan  $x$  adalah *keypoint*,  $c$  adalah centroid

### 2.3 Instrumen Penelitian

Perangkat yang digunakan untuk melakukan eksperimen ini sebagai berikut :

- Sistem operasi yang digunakan adalah Microsoft Windows 7 Ultimate 64 bit.
- Aplikasi code editor Matlab untuk menulis code program
- Menggunakan matlab versi R2014a 64 bit

## 3. LANDASAN TEORI

### 3.1 Speeded Up Robust Features (SURF)

Algoritma SURF mirip dengan algoritma SIFT, deteksi titik fitur SURF juga didasarkan pada teori ruang skala. Perbedaan nya bahwa algoritma SURF diterapkan pada DoH (Determinan of Hessian), ekstraksi poin fitur dalam gambar yang diwujudkan melalui penyederhanaan dan pendekatan, dibandingkan dengan algoritma SIFT, mengurangi kompleksitas ekstraksi titik fitur dan juga memiliki kinerja real-time yang lebih baik [13]. SURF

menggunakan penggabungan algoritma citra integral dan *blob detection* berdasarkan determinan dari matriks Hessian [14]. Ada beberapa tahap implementasi SURF, sebagai berikut :

#### a. Interest poin detector

Interest poin digunakan untuk memilih titik yang mengandung banyak informasi dan sekaligus stabil terhadap gangguan lokal atau global dalam citra digital. Algoritma SURF mempunyai sifat invariant dalam skala. Jika diberikan titik  $x = (x,y)$  pada citra  $I$ , matrik Hessian  $H(x,\sigma)$  pada  $x$  dengan skala  $\sigma$  didefinisikan sebagai berikut :

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix}$$

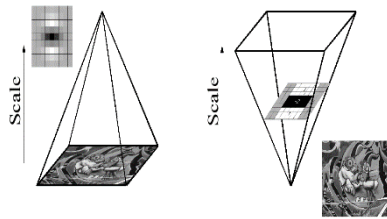
Dimana  $L_{xx}(x, \sigma)$  konvolusi dari turunan kedua fungsi Gaussian  $\frac{\partial^2}{\partial x^2} g(\sigma)$  dengan citra  $I$  pada titik  $x$ . Definisi ini berlaku juga untuk  $L_{xy}(x, \sigma)$  dan  $L_{yy}(x, \sigma)$ . Fungsi Gaussian didefinisikan sebagai :

$$g(\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

#### b. Scale Space Reprntasion

Ukuran citra yang berbeda-beda, akan sangat sulit untuk membandingkan fitur-fitur pada citra. Maka dari itu citra secara berulang akan diperhalus (*smoothing*) dengan fungsi Gaussian dan secara berurutan dengan cara *sub-sampling* untuk mencapai tingkat tertinggi pada piramida. Dalam

metode ini menggunakan metode *scale space* dimana citra di implementasikan dalam bentuk sebuah gambar piramida [14].



Gambar 2.1 Scale Space Reprantasion

- c. Feature description  
 Didefinisikan sebagai bagian yang mengandung banyak informasi suatu citra, dan fitur ini digunakan sebagai titik awal untuk algoritma deteksi objek. Langkah pertama [15] adalah melihat orientasi yang dominan pada titik perhatian yang terdapat pada citra, kemudian membangun suatu area yang akan diambil nilainya dan mencari fitur korespondensi pada citra pembanding. Dalam menentukan orientasi suatu citra menggunakan filter wavelet haar. Selanjutnya untuk deskripsi fitur dalam algoritma SURF, digunakan perhitungan *gradient histogram* dalam empat kelompok (*bins*) untuk mempercepat perhitungan.

### 3.2 Descriptor keypoint

*Keypoint* dari sebuah gambar. *Keypoint* [17] adalah titik-titik dari sebuah gambar yang nilainya tetap ketika mengalami perubahan skala, rotasi, blurring, pencahayaan, dan juga perubahan bentuk. Perubahan

bentuk ini bisa terjadi dikarena bentuk gambar *query* yang tidak utuh atau tidak sesempurna gambar yang ada didalam database gambar tersebut. Gambar *query* yang tidak utuh mungkin karena objek lain yang menutupi, atau pengambilan gambar yang tidak sempurna, atau keadaan objek itu sendiri yang mengalami perubahan. Agar invariant terhadap perubahan skala maka proses yang dilakukan pertama kali adalah membuat ruang skala (*scale space*). *Ruang Skala (Scale Space) Scale space* terbagi kedalam bilangan yang disebut *octave*. Setiap *octave* merepresentasikan respon filter yang diperoleh dengan melakukan proses konvolusi gambar yang diinputkan dengan ukuran filter yang menaik.

### 3.3 Algoritma K-means

Algoritma k-means merupakan salah satu partisi berbasis teknik pengelompokan berbasis non hierarki. Untuk setiap himpunan angka benda X dan jumlah integer K, pencarian algoritma K-Means untuk partisi dari X ke k cluster yang meminimalkan dalam kelompok jumlah kesalahan kuadrat. dimulai dengan menginisialisasi pusat k klaster. Titik data masukan tersebut kemudian dialokasikan ke salah satu cluster yang ada sesuai dengan kuadrat jarak Euclidean dari cluster, memilih yang paling dekat. Mean (centroid) dari setiap cluster kemudian dihitung sehingga untuk memperbarui pusat klaster. Pembaruan ini terjadi sebagai akibat dari perubahan dalam keanggotaan setiap

cluster. K-Means metode pengelompokan dapat dianggap, sebagai metode licik karena di sini, untuk mendapatkan yang lebih baik menghasilkan centroid, disimpan sejauh mungkin dari satu sama lain. Langkah-langkah untuk K-berarti algoritma yang diberikan di bawah [20]:

1. Memilih secara acak vector masukan K (titik data) untuk menginisialisasi cluster.
2. Pencarian terdekat : untuk setiap vector input, menemukan pusat cluster yang paling dekat, dan menetapkan bahwa vector memasukan cluster yang sesuai.
3. Update pusat cluster disetiap cluster dengan menggunakan rata-rata (centroid) dari vector untuk cluster itu.
4. Ulangi langkah 2 dan 3 sampai tidak ada lagi

cluster ini masih belum akurat. Tiap-tiap data memiliki derajat

Dengan cara memperbaiki pusat *cluster* dan nilai keanggotaan tiap-tiap data secara berulang, maka dapat dilihat bahwa pusat *cluster* akan menuju lokasi yang tepat. Perulangan ini didasarkan pada minimasi fungsi obyektif.

Fungsi obyektif yang digunakan FCM adalah :

$$J(U,V;X) = \sum_{k=1}^n \sum_{i=1}^C (\mu_{ik}) (d_{ik})$$

Dengan  $W \in [1, \infty]$ ,

$$d_{ik} = d(x_k - v_i) = [\sum_{j=1}^m (x_{kj} - v_{ij})] \quad (3.1)$$

X adalah data yang dicluster :

perubahan nilai yang lebih dalam.

### 3.4 Fuzzy C-means

Fuzzy c-means clustering (FCM) [21] atau dikenal juga sebagai Fuzzy ISODATA, Merupakan salah satu metode clustering yang merupakan bagian dari metode Hard K-Means. FCM menggunakan model pengelompokan Fuzzy sehingga data dapat menjadi anggota dari semua kelas atau *cluster* terbentuk dengan derajat atau tingkat keanggotaan yang berbeda antara 0 hingga 1. Tingkat keberadaan data dalam suatu kelas atau *cluster* ditentukan oleh derajat keanggotaannya. Teknik ini pertama kali diperkenalkan oleh Jim Bezdek pada tahun 1981. Konsep dasar FCM, pertama kali adalah menentukan pusat cluster yang akan menandai lokasi rata-rata untuk tiap-tiap cluster. Pada kondisi awal, pusat

keanggotaan untuk tiap-tiap *cluster*

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \quad (3.2)$$

dan V adalah matrik pusat cluster :

$$V = \begin{bmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & & \vdots \\ v_{c1} & \cdots & v_{cm} \end{bmatrix} \quad (3.3)$$

Nilai  $J_w$  terkecil adalah yang terbaik, sehingga :

$$J_w^* (U^*, V^* ; X) = \min_{m_{fc}} J (U, V, X) \quad (3.4)$$

Jika  $d_{ik} > 0, \forall i, k; w > 1$  dan X setidaknya memiliki m elemen, maka  $(U, V) \in$

$M_{fm} \times R^{mp}$  dapat meminimasi  $J_w$  hanya jika :

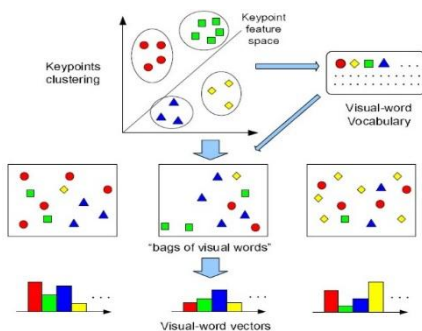
$$\mu_{ik} = \frac{[\sum_{j=1}^m (x_{ij} - v_{kj})^2]_{w-1}^{-1}}{\sum_{k=1}^m [\sum_{j=1}^m (x_{ij} - v_{kj})^2]_{w-1}^{-1}}; 1 \leq i \leq m; 1 \leq k \leq n \quad (3.5)$$

Keterangan :

1. Menentukan data yang akan di cluster  $X$ , berupa matriks berukuran  $n \times m$  ( $n$  = jumlah sampel data,  $m$  = atribut setiap data).  $X_{ij}$  = data sampel ke- $i$  ( $i=1,2,\dots,n$ ), atribut ke- $j$  ( $j=1,2,\dots,m$ ).
2. Menentukan jumlah cluster ( $c$ ), pangkat ( $w$ ), maksimum interaksi (MaxIter), error terkecil yang diharapkan ( $\epsilon$ ), fungsi obyektif awal ( $P_0 = 0$ ), interaksi awal ( $t=1$ ).

### 3.5 Bag of Visual Words

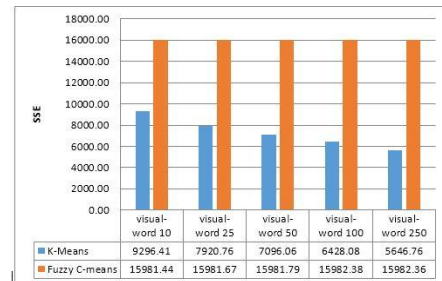
*Bag of Visual Words (BOVW)* digunakan di bidang computer vision [18]. BOVW sering digunakan dalam pengenalan dan penyaringan pada citra. Metode BOVW mempunyai tiga tahapan [19] : 1) Ekstrak descriptor fitur local (seperti SIFT), 2) Mengkuantisasi descriptor lokal ke codebook, 3) Konversi gambar menjadi satu set kata visual (*vocabulary*). Berikut contoh proses *Bag of Visual Words* :



Gambar 3.1 proses dari bag of visual words

## 4. EKSPERIMEN DAN HASIL

Dalam penelitian ini penulis melakukan eksperimen data untuk mencari hasil (SSE) yang terbaik dan waktu komputasi tercepat dalam pengelompokan. Sehingga penulis melakukan perbandingan menggunakan SURF-128 dan SURF-64 dengan dua algoritma yaitu K-Means dan Fuzzy C-Means. Berikut hasil eksperimen memperoleh *Sum of Square Errors* yang dipaparkan ke dalam grafik dibawah ini :



Gambar 4.1 SSE SURF-128 db 1

Pada gambar 4.1 algoritma K-Means lebih unggul dalam menentukan hasil SSE SURF-128 Db 1, karena nilainya lebih kecil dari algoritma Fuzzy C-Means, dan hasil dari grafik Fuzzy C-means stabil sedangkan K-means dari visual-word 10 sampai 250 mengalami penurunan grafik yang signifikan.



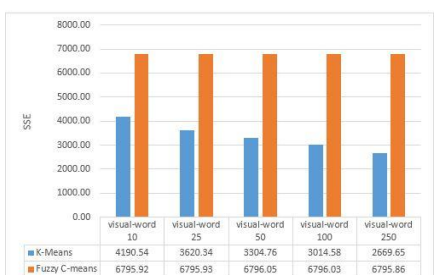
Gambar 4.2 SSE SURF-128 db 2

Pada gambar 4.2 algoritma K-Means masih unggul dalam menentukan hasil SSE SURF-128 Db 2 dibandingkan algoritma Fuzzy C-means



Gambar 4.3 SSE SURF-128 db 3

SSE SURF-128 Db 3 pun algoritma K-means masih terlihat unggul dalam menentukan hasil SSE dibandingkan algoritma Fuzzy C-Means pada gambar 4.3



Gambar 4.4 SSE SURF-128 db 4  
Sampai gambar 4.4 SSE SURF-128 Db 4 algoritma K-Means terlihat masih lebih unggul dalam menentukan hasil SSE dibandingkan algoritma Fuzzy C-means. Yang dimana dari mulai 10 visual-word sampai 250 visual-word nilai SSE

terkecil ditunjukkan oleh algoritma K-Means.



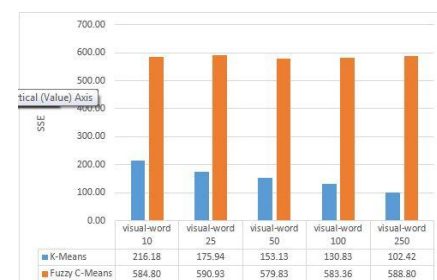
Gambar 4.5 SSE SURF-64 db 1

Pada gambar 4.5 algoritma K-means menentukan hasil SSE pada SURF-64 Db1 memperoleh nilai yang rendah dibandingkan dengan algoritma Fuzzy C-means yang memperoleh nilai tinggi.



Gambar 4.6 SSE SURF-64 db 2

Pada gambar 4.6 di atas menjelaskan bahwa nilai SSE Surf-64 pada Db2 yang terkecil di tunjukan pada algoritma k-means. Sedangkan algoritma Fuzzy C-means memperoleh nilai yang besar dalam menentukan SSE.



Gambar 4.7 SSE SURF-64 db 3



Algoritma K-means menjelaskan bahwa nilai SSE Surf-64 pada Db3 yang terkecil. Sedangkan algoritma Fuzzy C-means memperoleh nilai yang besar dalam menentukan SSE. Dapat dilihat pada gambar 4.7



Gambar 4.8 SSE SURF-64 db 4

Di gambar 4.8 diatas bahwa algoritma k-means dalam melakukan *clustering keypoint* untuk menentukan SSE pada Surf-64 di Db4 baik, akan tetapi pada visual word ke 100 algoritma Fuzzy C-means memiliki nilai terkecil.

## 5. KESIMPULAN DAN SARAN

### 1. Kesimpulan

Berdasarkan penelitian yang telah dilakukan, maka dapat disimpulkan beberapa hal sebagai berikut :

1. Berdasarkan eksperimen yang telah dilakukan dalam mengevaluasi performa *clustering keypoint* menggunakan algoritma K-means lebih baik dalam menentukan hasil SSE dalam pengelompokan descriptor pada SURF-64 dan SURF-128
2. Algoritma K-Means dalam melakukan evaluasinya membutuhkan waktu

komputasi yang cukup cepat dibandingkan algoritma Fuzzy C-Means yang membutuhkan waktu komputasi yang cukup lama.

### 2. Saran

Saran peneliti untuk mengembangkan penelitian ini antara lain :

1. Penelitian bisa dikembangkan dengan mencoba algoritma lain sebagai pembanding
2. Menerapkan metode reduksi dimensi untuk mempercepat waktu komputasi k-means. Metode reduksi dimensi yang diusulkan seperti *Singular Value Decomposition (SVD)* dan *Non-negative Matrix Factorization (NMF)*.

## 6. DAFTAR PUSTAKA

- [1] I. G. Babatunde, A. O. Charles, A. B. Kayode dan a. O. Olatubosun, "Fingerprint Image Enhancement," *International Journal of Advanced Computer Science and Applications*, vol. 3, no. 1, 2012.
- [2] C. Z. a. P. H. Shihua He, "Clustering-Based Descriptors for Fingerprint Indexing Fast Retrieval," *Asian Conference on Computer Visison*, pp. 354-363, 2010.
- [3] M. a. R. Mishra, "Fingerprint Recognition Using Robust Local Features," *Internarional Journal of Advanced Research in Computer Science and Software Engineering*,

- vol. 2, no. 6, 2012.
- [4] A. Balti dan M. S. a. F. fnaiech, "Fingerprint Images Segmentation based on fuzzy C-means theory and statistical features," *International Journal of the Physical Sciences*, vol. 7, pp. 462-470, 2012.
- [5] E. Priyatno, M. Hariadi dan A. Zaini, "Perkiraan Bentuk dan Pose Model 3D," *Prosiding Penelitian Bidang Studi TKT*, 2009-2010.
- [6] S. He dan C. Z. a. P. Hao, "Clustering-Based Descriptors for Fingerprint Indexing and Fast Retrieval," *Asian Conference on Computer Vision*, pp. 354-363, 2010.
- [7] K. S. a. K. Yildiz, "A Web Based Clustering Analysis Toolbox (WBCA) design using," *Procedia Sosial and Behavioral Sciences*, vol. 2, pp. 5276-5280, 2010.
- [8] L. N. a. A. Lumini, "Descriptor for Image-Based Fingerprint Matchers," *Expert System With Applications*, vol. 36, pp. 12414-12422, 2009.
- [9] X. W. L. Q. Feng Qi, "Research of Image Matching Based on Improved SURF Algorithm," *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 12, pp. 1395 - 1402, 2014.
- [10] P. G. A. a. W. G. A. A. S. Gunawan, "Pendeteksian Rambu Lalu Lintas Dengan Algoritma Speeded Up Robust Features (SURF)," *Mathematics & Statistics Department School of Computer Science*, vol. 13, pp. 91 - 96, 2013.
- [11] W. N. S. H. Hu Shuo, "Object Tracking Method Based On SURF," *AASRI Conferense on Modelling, Identification and Control*, Vol. %1 dari %2351 - 356, 2012.
- [12] K. D. Tania, "Pengenalan Gambar Menggunakan Sebagian Data Gambar," *Jurnal Generik*, vol. 5, 2010.
- [13] T. G. K. Somasundaram, "Binarization of MRI with Intensity Inhomogeneity using K-Means Clustering for Segmenting Hippocampus," *IJMA The International Journal of Multimedia & Its Applications*, vol. 5, 2013.
- [14] Y.-G. J. A. H. a. a. C.-W. N. J. Yang, "Evaluating Bag-of-Visual-Words Representations in Scene Classification," 2007.
- [15] L. S. L. Z. Z. L. a. Y. Y. J. Zhang, "An approach of bag-of-words based on visual attention model for pornographic images recognition in compressed domain," *Neurocomputing*, vol. 110, pp. 145 - 152, 2013.