Analisis dan Penerapan Perhitungan Orang Menggunakan Metode Histogram Of Oriented Gradients-Local Binary Pattern Dengan Deteksi Kepala-Bahu Studi Kasus: Perhitungan Orang Dalam Kelas

Analysis and Implementation Of People Counting Using Histogram Of Oriented Gradients-Local Binary Pattern Based on Head-Shoulder Detection Case Study: People Counting in Class Room

Rivan Ardyanto Sutoyo[1], Bedy Purnama, S.Si., MT.[2]Febryanti Sthevanie, ST., MT.[3]

[1,2,3] Teknik Informatika, Universitas Telkom
Jl. Telekomunikasi, Dayeuh Kolot Bandung 40257 Indonesia

rivanardyanto19@gmail.com, bedy.purnama@gmail.com, febryantisthevanie@gmai.com

ABSTRAKSI

Menghitung jumlah orang (*people counting*) yang berada dalam video merupakan salah satu fungsionalitas yang saat ini sedang dikembangkan para peneliti di dunia *computer vision*. *People Counting* memiliki berbagai macam manfaat seperti sistem pengawasan keamanan, estimasi jumlah kepadatan manusia, analisis bisnis di sebuah tempat, dan masih banyak lagi manfaatnya. Adapun studi kasus yang dipakai untuk pembuatan Tugas Akhir ini yaitu menghitung jumlah orang di dalam kelas dengan menggunakan kombinasi metode *Histogram Of Oriented Gradients* (HOG) dan *Local Binary Pattern* (LBP) untuk deteksi kepala dan bahu dari seorang individu. HOG-LBP digunakan untuk fitur ekstraksi untuk identifikasi objek. Kemudian perhitungan orang dilakukan dengan cakupan individu yang berada di dalam video. Keluaran dari sistem ini adalah sistem dapat menunjukkan jumlah orang dalam ruang kelas serta menunjukkan lokasi masing-masing dari individu tersebut berdasarkan kepala dan bahu. Dari hasil penelitian dengan menggunakan metode Multilevel HOG-LBP didapatkan rata-rata hasil akurasi orang duduk berjarak satu sebesar 93.53% dan rata-rata akurasi data orang duduk berdampingan sebesar 80.33%.

Kata Kunci: *people counting, histogram of oriented, local binary pattern, kepala,bahu*

ABSTRACT

Counting the number of people which are in the video is one of the functionality that is currently being developed by researchers in the world of computer vision. People counting have a wide range of benefits such as security surveillance system, the estimated number of human density, in a place of business analysis, and many more benefits. The case studies used for the preparation of this final project is to calculate the number of people in the classroom by using a combination of methods Histogram Of Oriented histogram (HOG) and Local Binary Pattern (LBP) for the detection of head and shoulders of an individual. HOG -LBP is used for feature extraction for object identification. Then the calculation is done with individual coverage inside the video. The outcomes of this system is the system can indicate the number of people in the classroom and show the location of each of the individual based on the head and shoulders. From the research of this system the average of accuration for case people sit in class room separate one seat is 93.53% and people sit in class room next to each other is 80.33%.

Keywords: *people counting, histogram of oriented, local binary pattern, head, shoulder*

1. Pendahuluan

1.1. Latar Belakang

Proses perhitungan jumlah orang atau yang biasa dikenal dengan *people counting* saat ini sedang mengalami perkembangan yang sangat signifikan di dalam dunia *computer vision*. *People counting* memiliki banyak sekali manfaatnya antara lain: mengawasi sistem keamanan jumlah antrian tiket di stadium bola, estimasi jumlah kepadatan manusia atau pengunjung kawasan bisnis untuk membantu analisis bisnis di sebuah tempat dan estimasi jumlah mahasiswa di dalam suatu ruangan. Dalam kurun waktu ini ada begitu banyak metode yang diusulkan oleh para peneliti di seluruh dunia. Metode-metode untuk perhitungan orang dua dimensi dapat dibagi menjadi 4 kategori yaitu *Mapping-based Methods* yang biasa dipakai untuk estimasi perhitungan orang di keramaian, *Segmentation-based Methods* dimana tergantung pada akurasi masing-masing metode, *Clustering-based Methods*, dan *Detection-based Methods* [9]. Sebuah

penelitian *people counting* menyebutkan bahwa deteksi manusia dengan hanya kepala saja belum terbukti akurat [10]. Sehingga penelitian baru untuk deteksi manusia dikembangkan kembali dengan cara mencoba menggabungkan deteksi manusia dari kepala dan bahu [18]. Fitur esktraksi *Histogram Of Oriented Gradients* (HOG) dan *Local Binary Pattern* (LBP) sangat efektif dalam pengembangan deteksi bagian atas tubuh manusia termasuk kepala dan bahu kemudian [3] ditambah lagi dengan penelitian yang mengatakan bahwa deteksi dengan menggunakan kepala dan bahu terbukti sangat ampuh untuk mengenali ciri-ciri dari seorang individu(manusia) [12]. HOG adalah deskriptor yang sangat baik untuk mengenali garis-garis tepi bentuk manusia [19,20] sedangkan LBP dikenal baik untuk deskriptor tekstur [17].

Metode yang dipakai di dalam tugas akhir ini adalah metode *Histogram Of Oriented Gradients* (HOG) dan *Local Binary Pattern* (LBP) untuk deteksi kepala dan bahu dari seorang individu dengan bantuan klasifikasi *Support Vector Machine* (SVM). Studi kasus dalam pembuatan tugas akhir ini yaitu menghitung jumlah orang di dalam ruangan kelas dengan menggunakan metode di atas.

1.2. Perumusan Masalah

Masalah-masalah yang akan dirumuskan pada tugas akhir ini adalah:

1. Bagaimana menerapkan metode HOG-LBP untuk mengidentifikasi bentuk seorang individu melalui deteksi kepala dan bahu dalam proses perhitungan jumlah orang di dalam ruangan kelas?
2. Bagaimana menganalisis akurasi dan performansi sistem yang dibangun?

1.3. Tujuan

Adapun tujuan dari tugas akhir ini yaitu:

1. Menerapkan metode HOG-LBP untuk mengidentifikasi bentuk seorang individu melalui deteksi kepala dan bahu dalam proses perhitungan jumlah orang di dalam ruangan kelas.
2. Menghitung jumlah orang yang berada di dalam ruangan kelas dengan metode HOG-LBP serta menganalisis performansi dan akurasi sistem.

1.4. Batasan Masalah

Adapun batasan masalah dalam tugas akhir ini yaitu:

1. Diasumsikan bahwa ruangan kelas seperti ruangan di B107, B109, Gedung K, Multimeda LC yang memiliki baris tempat duduk yang berbeda-beda ketinggiannya.
2. Kamera yang digunakan merupakan kamera yang statik, tinggi kamera ± 2.2 meter, dan sudut pengambilan gambar ± $30^{o.}$
3. Penerapan *background substraction* menggunakan *frame* pertama sebagai acuan karena diasumsikan pada sistem bahwa tidak terdapat objek bergerak pada setiap awalan video.

2. Landasan Teori

2.1. Citra Digital

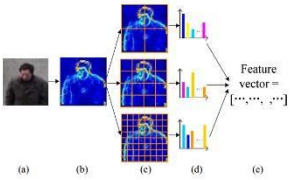
Citra merupakan representasi dari sebuah objek. Citra terdiri dari citra digital dan citra analog. Citra analog memiliki sifat kontinu seperti pada monitor televisi, sinar X, CT *scan* sedangkan citra digital adalah citra yang dapat diolah oleh komputer seperti kamera, *scanne*r, dan alat optik lainnya. Citra digital merupakan fungsi (x,y) yang digambarkan pada bidang 2 dimensi dimana x dan y merupakan koordinat piksel dan merupakan fungsi derajat intensitas piksel tersebut. Citra digital berbentuk matriks 2 dimensi berukuran M x N dimana M merupakan baris dan N merupakan kolom. Secara umum citra digital dibagi menjadi 2 jenis yaitu citra diam atau dikenal dengan *still image* dan citra yang bergerak atau *moving image*. Tugas akhir ini memakai citra yang bergerak (*moving image*) sebagai bahan citranya. [5,6]

2.2. *Frame Difference*

Background substraction digunakan untuk mendeteksi objek bergerak di dalam video dari kamera statis. *Background substraction* merupakan proses memisahkan *foreground* dari *background* yang dapat diimplementasikan dengan menggunakan variasi kompleksitas yang bermacam-macam. *Low complexity* menggunakan metode *frame difference*. Konsep dasarnya adalah dengan cara membadingkan *frame* saat ini dengan frame acuan atau biasa dikenal dengan *background model* atau *background image*. Algoritma *frame difference* adalah mengubah *frame* ke dalam bentuk *grayscale*, mengurangi *frame* saat ini dengan *background model*, kemudian untuk setiap *pixel* jika hasil dari pengurangan lebih besar dari *threshold* maka *pixel* tersebut merupakan *foreground*. Metode yang digunakan ini melihat perbedaan *frame* sekarang dengan *frame* pertama yang diasumsikan bahwa *frame* pertama tidak berisi objek yang sedang bergerak. Dengan kata lain, *frame* pertama dijadikan sebagai bahan acuan dan selalu dibandingkan dengan *frame-frame* selanjutnya. [2,14]

2.3. Multilevel HOG

Multilevel HOG adalah sebuah tahapan pengembangan dari performansi HOG untuk mengekstraksi ciri dari kepala dan bahu.

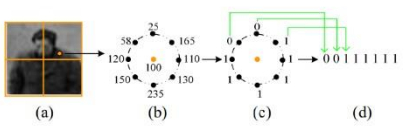


Gambar 2. 1 Prosedur Ekstraksi *Multi Level* HOG [3]

- Input citra, berupa citra berukuran 48x64 *pixel*
- Menghitung hasil *gradient magnitude* dari input citra
- Citra tersebut dibagi pada 3 level yang berbeda. Level pertama terdiri dari 2x2 sel, sel pada level ini berukuran 24x32 *pixel*. Level kedua terdiri dari 4x4 sel, sel pada level ini berukuran 12x16 *pixel*. Level ketiga terdiri dari 8x8 sel, sel pada level ini berukuran 6x8 *pixel*.
- Histogram pada masing-masing level
- Proses konkatenasi dari 3 level tersebut dan didapatkan *feature vector* akhir dari semua level

2.4. Multilevel LBP

Multilevel LBP feature berfungsi sebagai *edge descriptor* akan tetapi belum terlalu baik untuk deteksi kepala dan bahu. Kombinasi dari *edge shape information* dan *texture information* akan mempebaiki hasil deteksi. Tahapannya hampir sama dengan *Multilevel HOG* tetapi hanya saja perbedaannya terletak di *input image* yang dirubah ke dalam bentuk *grayscale* kemudian gambar ditandai dengan 8 *point radius* dengan masing-masing *pixel* transisi nomor biner 0-1. Untuk setiap *block* dalam satu level *pixel* sejenis di dalam *block* dirubah ke dalam bins yang berbeda-beda sedangkan *pixel* yang tidak sejenis dibuat menjadi satu bin. Disini L2-Hys *normalization scheme for the histograms of the blocks* dipakai karena dirasa lebih baik daripada *L1-sqrt normalization scheme* yang d



Gambar 2. 2 Prosedur Ekstraksi *MultiLevel* LBP [19]

Berikut ini merupakan keterangan dari gambar di atas:

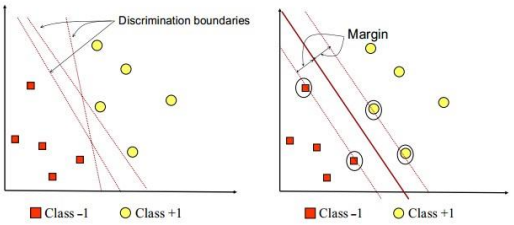
a. Gambar diubah ke dalam *grayscale image*
b. 8 *sampling points* dengan radius 1 *of the pixel in the block.*
c. *The binary value of the sampling point* menganggap *pixel* sebagai *threshold*.
d. Pola biner dari *LBP pattern of the pixel*.

Gambar ini merupakan ilustrasi dari metode ekstraksi ciri *Local Binary Pattern* pada orang di ruangan kelas yang ada di dalam sistem.

Gambar 2. 1 Ilustrasi LBP

2.5. Support Vector Machine

Support Vector Machine adalah metode *learning machine* yang bekerja atas prinsip *Structural Risk Minimization* (SRM) dengan tujuan menemukan *hyperplane* terbaik yang memisahkan dua buah *class* pada input *space*. Prinsip dasar dari SVM adalah *linear classifier* dan selanjutnya dikembangkan agar dapat bekerja pada *problem non-linear* dengan cara memasukkan konsep *kernel trick* pada ruang kerja berdimensi tinggi. SVM *linear classifier* berfungsi untuk mencari nilai optimal dari sebuah *hyperplane* sehingga bertujuan untuk mendeteksi ada atau tidaknya sebuah objek.



Gambar 2. 3 SVM berusaha mencari *hyperplane* terbaik dengan cara memisahkan kelas -1 dan +1. [13]

Gambar pada ruas sebelah kiri memperlihatkan beberapa *pattern* yang merupakan anggota dari dua buah *class*: +1 dan -1. Pattern yang tergabung pada *class* -1 disimbolkan dengan warna merah (bentuk kotak), sedangkan pattern class -1 disimbolkan dengan warna kuning (bentuk lingkaran). Permasalahan pada klasifikasi objek dapat diterjemahkan dengan cara menemukan garis (*hyperplane*) yang memisahkan antara kedua kelompok tersebut. *Hyperplane* pemisah terbaik diantara kedua *class* didapat dengan cara mengukur margin *hyperplane* tersebut dan mencari titik maksimalnya. Margin adalah jarak antara *hyperplane* tersebut dengan pattern terdekat dari masing-masing *class*. *Pattern* terdekat inilah yang disebut sebagai support vector. Gambar pada ruas kanan menunjukkan *hyperplane* terbaik yang terletak tepat pada tengah-tengah kedua *class*, sedangkan titik merah dan kuning yang berada dalam lingkaran hitam merupakan *support vector*. Usaha untuk mencari lokasi *hyperplane* ini yang merupakan inti dari proses pembelajaran pada SVM. [13]

2.6. Pengujian Sistem

Untuk skenario pengujian disesuaikan dengan data uji yang telah diambil. Rumus perhitungan akurasi yang diterapkan dari rumus sebagai berikut [22]:

$$ak\text{[illegible]} = \frac{TP + TN}{T\text{[illegible]} + T\text{[illegible]} + F\text{[illegible]} + FN} X\ 100\ \%$$

(2.1)

Keterangan:

1. TP (*true positive*) adalah jumlah orang yang terhitung oleh sistem.
2. TN(*true negative*) adalah jumlah yang bukan orang dan tidak dihitung oleh sistem.
3. FP(*false positive*) adalah jumlah yang bukan orang tetapi terhitung sebagai orang oleh sistem.
4. FN(*false negative*) adalah jumlah orang yang terhitung sebagai orang namun tidak terhitung oleh sistem.

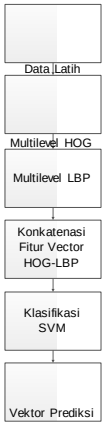
3. Analisis dan Perancangan Sistem

3.1. Deskripsi Sistem

Sistem yang dibangun dalam Tugas Akhir ini dibuat dengan menggunakan kombinasi metode Multilevel *HOG-LBP* untuk fitur ekstraksi (deteksi individu), untuk bagian *background substraction* mengunakan perbandingan *frame* awal dengan *frame-frame* selanjutnya atau biasa dikenal dengan istilah *frame difference*. dan klasifikasi menggunakkan *Support Vector Machine* (SVM). Setelah didapatkan hasil deteksi maka tahapan berikutnya adalah sistem akan melakukan perhitungan orang. Dalam tugas akhir ini terdapat tahapan *Learning* yang bertujuan untuk proses *training* dan tahapan *Detection* dan *Counting* yang bertujuan untuk mendeteksi dan menghitung orang.

3.1.1 Tahapan *Learning*

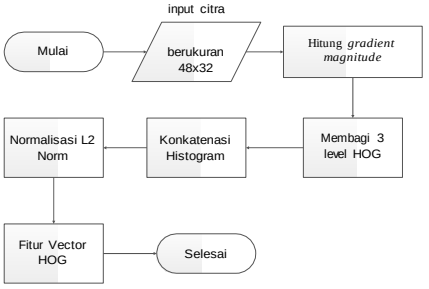
Sistem ini dibangun dengan menggunakan beberapa tahap yang saling terkait dari proses awal sampai akhir. Sistem dapat digambarkan melalui fase-fase berikut:



Gambar 3. 1 Tahapan Fase Learning Umum

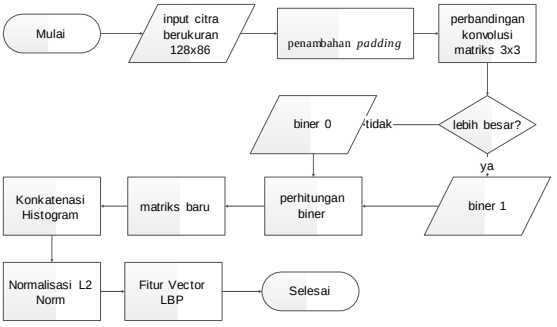
kepala-bahu. Kumpulan data-data latih bersumber dari internet, beberapa data training Gender Dataset Person, INRIA Dataset Person. Kemudian setiap data latih di-*training* dengan HOG dan LBP kemudian diklasifikasi dengan SVM. Data latih yang akan dipakai adalah citra berukuran 256x256 *pixel*.

Data latih yang dipakai di dalam Tahapan Fase Learning HOG adalah citra berukuran 48x64 *pixel*. Lalu citra tersebut akan dibagi ke dalam 3 level yang berbeda dimana setiap level memiliki ukuran citra yang berbeda satu sama lain. Level 1 berukuran 24x32 *pixel* dengan ukuran sel 2x2, Level 2 berukuran 12x16 *pixel* dengan ukuran sel 4x4, Level 3 berukuran 6x8 *pixel* dengan ukuran sel 8x8. Langkah terakhir dari Learning HOG ini adalah proses normalisasi dan konkatenasi histogram setiap level.



Gambar 3. 2 Diagram alir Learning HOG

Secara umum diagram alir antara Fase *learning* HOG dan fase *Learning* LBP sama. Perbedaannya adalah pada bagian *resize* citra dan teknik dalam mengolah citra tersebut.



Gambar 3. 3 Diagram Learning LBP

3.1.2 Tahapan *Detection* dan *Counting*

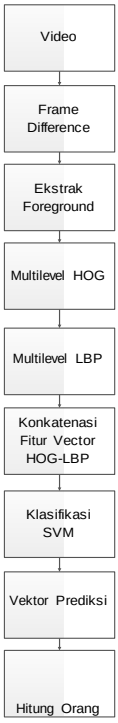
Tahapan *Detection* dan *Counting* merupakan tahapan uji video dimana inputan data bukan merupakan citra namun merupakan video kemudian dilakukan proses *background substraction*, ekstraksi ciri, klasifikasi SVM, dan proses perhitungan orang.

Data latih yang digunakan menggunakan 2 jenis data latih yang berbeda yaitu terdapat data latih positif dimana

berisi kumpulan gambar-gambar kepala-bahu dan data latih negatif dimana berisi kumpulan gambar bukan

Inputan berupa sebuah video kemudian dilakukan proses pemisahan *foreground* dan background. Ekstraksi ciri Multilevel HOG-LBP menghasilkan vektor prediksi dan diproses oleh SVM jika objek terdeteksi sebagai orang

maka sistem akan menghitung. Hasil akhir perhitungan orang berupa jumlah berupa *integer*.



Gambar 3. 4 Tahapan Detection dan Counting

3.2. Perancangan Aplikasi

3.2.1 Tahapan *Learning*

Tahapan *learning* merupakan tahapan awal proses training data latih positif dan negatif, metode ekstraksi ciri, dan klasifikasi menggunakan SVM yang dijadikan acuan pada tahapan *detection* dan *counting*. Berikut ini adalah langkah awal untuk memulai tahapan *learning*:

1. Dalam proses *training*, gambar yang digunakan untuk data *training* menggunakan 2 jenis data latih yang berbeda yaitu terdapat data latih positif dimana berisi kumpulan gambar-gambar kepala-bahu dan data latih negatif dimana berisi kumpulan gambar bukan kepala-bahu. Kumpulan data-data latih ini bersumber dari *web* internet, beberapa data *training Gender Dataset Person*, INRIA *Dataset Person*. Berikut ini adalah detailnya:

Tabel 3. 1 Kumpulan Dataset

Sumber\Jenis Data Latih	Positif	Negatif
INRIA Dataset Person	614	1218
Genre Dataset Person	2044	-
Web internet	36	28

Data latih positif dan negatif dibuat sama ukurannya yaitu 128x86 mengingat resolusi gambar yang mempengaruhi ekstraksi ciri. Jumlah data latih positif adalah 167 dan negative adalah 66 dan total data latih positif dan negatif yang dipakai adalah 233.

Tabel 3. 2 Data Latih Positif dan Negatif

Sumber\Jenis Data Latih	Positif	Negatif
INRIA Dataset Person	-	38
Genre Dataset Person	131	-
Web internet	36	28
Total	167	66

Komposisi data latih positif dengan data latih negatif 70:30 alasannya adalah fokus dari sistem ini adalah objek yang terdeteksi sebagai manusia saja yang akan langsung masuk ke dalam ekstraksi ciri sehingga data latih positif yang diperbanyak. Selain itu di dalam video tidak terdapat gerakan gerakan objek lain atau gerakan semu sehingga data latih negatif dibuat lebih kecil daripada data positif. Pada sistem ini, dilakukan proses kombinasi data training data positif dan negatif yaitu ada yang memiliki *background* selain putih dan yang memiliki *background* berwarna putih polos.

Gambar 3. 5 Sampel positif (INRIA,Gender,web internet)

Dataset yang digunakan untuk sampel negatif merupakan gambar yang tidak menunjukkan ciri kepala dan bahu contohnya adalah kaki manusia atau gambar pemandangan.

Gambar 3. 6 Sampel negatif (INRIA ,web internet)

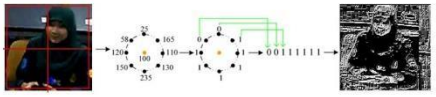
2. Setelah melakukan proses training maka selanjutnya tahapan ekstraksi ciri HOG dan LBP. Dalam sistem ini digunakan metode Multilevel HOG dan Multilevel LBP dengan masing-masing metode 3 tahapan level.

Outputnya berupa hasil konkatenasi *feature vector setiap level HOG dan LBP*.

Gambar 3. 7 Visualisasi HOG

Pada multilevel HOG setiap level pada masing-masing gambar diubah ke dalam bentuk *grayscale* .Setiap level HOG dilakukan proses *resize* gambar dengan variasi ukuran 24x32, 12x16, dan 6x8 piksel. Kemudian pada tahap akhir proses Multilevel HOG dilakukan tahapan penggabungan histogram setiap level yang menghasilkan satu *feature vector* yang panjang yang disebut dengan konkatenasi.

Pada awalnya objek yang terdeteksi sebagai manusia akan diberi *bounding box* dan kemudian masuk ke dalam tahapan ekstraksi ciri Multilevel HOG dan Multilevel LBP.



Gambar 3. 8 Visualisasi LBP

Pada multilevel LBP setiap level pada masing-masing gambar diubah ke dalam bentuk *grayscale* . Setiap level LBP dilakukan proses pembagian area gambar dengan variasi pembagian gambar pembagian 4,16, dan 64. Kemudian pada tahap akhir proses Multlevel LBP dilakukan tahapan penggabungan histogram setiap level yang menghasilkan satu *feature vector* yang panjang.

3. Setelah melalui tahapan ekstraksi ciri Multilevel HOG-LBP maka selanjutnya dilakukan proses konkatenasi antara *feature vector* Multilevel HOG dengan *feature vector* Multilevel LBP yang menghasilkan *feature vector* yang baru.
4. Tahapan klasifikasi dengan menggunakan fungsi *Support Vector Machine classify. Support Vector Machine* merupakan sebuah metode learning yang berfungsi untuk menyelesaikan masalah klasifikasi banyak *class* dari data yang sangat banyak dengan waktu yang singkat. Dalam sistem ini proses klasifikasi data training akan memberikan nilai 1 untuk data positif dan 0 untuk data negatif.
5. Proses akhir dari tahapan *learning* ini adalah berupa vektor prediksi yang berisi hasil klasifikasi data yang akan digunakan sebagai acuan dalam proses deteksi dan *counting*.

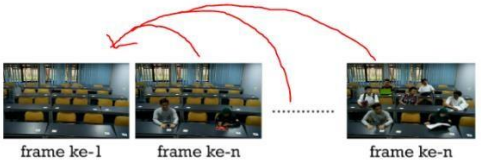
3.2.2 Tahapan *Detection* Dan *Counting*

Tahapan *detection* dan *counting* merupakan tahapan akhir dalam sistem ini. Tahapan ini memiliki tujuan untuk mendeteksi objek manusia yang bergerak di dalam video dan memisahkan bagian *foreground* dan *background* melalui proses *background substraction*. Kemudian bagian yang terdeteksi sebagai f*oreground* masuk ke dalam tahapan ekstraksi ciri dan melalui proses klasifikasi.

Pada akhir tahapan didapatkan hasil perhitungan jumlah manusia berupa *integer*. Berikut ini merupakan pemaparan yang lebih spesifik dari tahapan ini:

1. Proses *Background Substraction*

   Proses *Background Substraction* ini menggunakan teknik *frame difference* dimana *frame* awal dijadikan *role model* dalam keseluruhan *frame* yang ada di dalam video tersebut dengan anggapan *frame* awal tidak berisi objek yang bergerak.Cara kerjanya adalah membandingkan *frame* ke-1 dengan *frame* ke-n..n. Setiap *frame* dirubah ke bentuk *grayscale* jika area *blob* memungkinkan itu orang yang bergerak maka *blob* tersebut diberi bounding dengan ratio *width* dan *height*nya kemudian membuat rasio koordinat x dan y untuk melakukan *bounding* kepala-bahu pada objek yang bergerak. Berikut ini adalah penjelasan ilustrasi *frame difference* yang akan digunakan dalam sistem ini:



Gambar 3. 9 Ilustrasi Background Substraction

2. Hasil Ekstraksi Foreground

   Setelah didapatkan bagian mana yang dianggap *foreground* maka tahapan selanjutnya adalah melakukan proses ekstraksi ciri MultilevelHOG-Multilevel LBP pada setiap objek yang terdeteksi bounding.

3. Hitung orang

   Tahapan akhirnya adalah setelah melewati proses ekstraksi ciri, dianggap sebagai objek manusia, dan lolos klasifikasi SVM dengan nilai 1 maka objek tersebut akan dihitung sebagai manusia. Proses ini akan berlangsung sampai frame terakhir.

4. Hasil dan Analisis

4.1. Implementasi Sistem

Sistem *people counting* ini diimplementasikan dalam lingkungan perangkat keras dan perangkat lunak dengan spesifikasi berikut ini.

4.1.1 Spesifikasi Perangkat Keras

Spesifikasi perangkat keras yang digunakan untuk membangun sistem ini adalah :

1. Processor Intel®Core™ i5-4200U CPU @ 2.40 GHz 2.40 GHz
2. Memory 4.00 GB RAM
3. NVIDIA GT740 2 GB
4. Hardisk 750 GB
5. Kamera Digital Samsung Titanium NV4

4.1.2 Spesifikasi Perangkat Lunak

Perangkat lunak yang digunakan untuk membangun sistem ini adalah sebagai berikut :

1. Sistem Operasi Windows 8.1 Pro 64-Bit
2. RAM : 2 GB
3. Matlab R2014a

4.2. Pengujian Sistem

Pengujian sistem ini dilakukan dengan cara pengujian setiap tahapan dalam sistem. Tahapan pertama adalah pengujian parameter *threshold background* substraction dan pengujian ekstraksi ciri. Pengujian sistem ekstraksi ciri menggunakan dua data uji video dengan dua skenario yang berbeda satu sama lain. Skenario pertama adalah menguji sistem dengan Multilevel HOG-LBP dengan modifikasi ekstraksi ciri LBP dengan 4 *negihbourhood* dan skenario kedua menguji keseluruhan sistem dengan multilevel HOG-LBP tanpa modifikasi. Sehingga akurasi untuk setiap data uji berbeda-beda dalam segi waktu dan persentasi hasilnya.

4.2.1. Data Uji

Dalam proses pengambilan data video sebanyak 3 kali yaitu pada ruangan multimedia LC dengan tinggi kamera ± 2 meter, dan sudut ± 30º, ruangan B109, dan ruangan B107. Ruangan yang paling ideal dalam segi suasana, posisi kamera, intensitas cahaya, dan *noise* adalah ruangan B107. Pengambilan data berupa video dari arah depan kamera,arah masuk objek dari sebelah kiri gambar dicoba, dengan tinggi kamera ± 2.2 meter, menggunakan

ruang kelas B 107, dan sudut ± 30º). Data video tersebut diambil dengan kamera digital dengan resolusi 800x592 piksel dengan kecepatan 20 fps.

Gambar 4. 1 Ruangan Kelas B107

Objek-objek di kursi berisi berbagai macam kemungkinan ada kursi kosong, kursi yang diisi barang seperti tas, dan berisi orang.

Gambar 4. 2 Objek-objek di ruangan kelas B107

4.2.2. Data Testing

Pengujian sistem ini menggunakan dua jenis data uji yang berbeda dimana data uji 1 yang berisi 7 video dan data uji 2 yang berisi 5 video. Data uji 1 adalah data uji orang duduk berjarak satu kursi di setiap baris sedangkan data uji 2 adalah data uji orang duduk berdampingan di setiap baris. Dalam pengujian data testing ini lingkunan (*environment*) yang berada dalam video merupakan *environment* yang telah dibuat ideal dalam segi pencahayaan yang homogen dan jumlah orang yang duduk di kursi, dan sudut pengambilan video adalah sudut ± 30º.

.

Tabel 4. 1 Data Uji 1 dan Data Uji 2

No	Video	Keterangan
Data Uji 1	Video 1	Berisi 1 orang
	Video 2	Berisi 2 orang
	Video 3	Berisi 3 orang
	Video 4	Berisi 4 orang
	Video 5	Berisi 5 orang
	Video 6	Berisi 6 orang
	Video 7	Berisi 7 orang
Data Uji 2	Video 8	Berisi 1 orang
	Video 9	Berisi 2 orang
	Video 10	Berisi 3 orang
	Video 11	Berisi 4 orang
	Video 12	Berisi 5 orang

4.2.3. Skenario Pengujian

Pada penelitian ini terdapat beberapa tahapan skenario pengujian. Berikut ini adalah skenarionya:

4.2.3.1 Pengujian *Background Substraction*

Langkah awal dalam tahapan *detection* dan *counting* adalah *background substraction*. Dalam sistem ini diterapkan *frame difference* untuk memisahkan bagian *foreground* dan *background*. Ide dari *frame difference* ini adalah dengan menjadikan *frame* awal sebagai acuan (*role model*) untuk perbandingan dengan keseluruhan isi *frame*

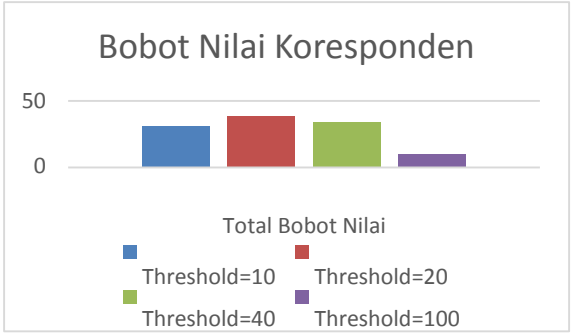
seperti yang telah dijelaskan pada bab 3 pada langkah *background substraction.*
Adapun beberapa parameter yang mempengaruhi proses *background substraction* ini diantaranya adalah:

1. Parameter *thresh* di dalam sistem berfungsi sebagai nilai ambang batas dalam mengubah citra berderajat keabuan menjadi citra biner (berwarna hitam=0 dan putih=1). Jika mengubah nilai dari parameter *thresh* ini maka dapat diketahui daerah mana yang termasuk objek (*foreground*) dan *background* dari *frame* secara jelas. Dari percobaan berbagai macam variasi nilai *thresh* 10,20,40,100, nilai *thresh* yang paling tinggi bobot nilainya adalah *thresh* = 20 karena banyak diberi bobot nilai 4 oleh koresponden yang berjumlah 10 orang.

Tabel 4. 2 Tabel Pengujian Threshold

No.	Nilai Threshold	Total Bobot Nilai
1	Thresh = 10	31
2	Thresh = 20	39
3	Thresh = 40	34
4	Thresh = 100	10

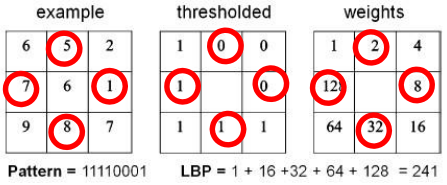
Keterangan:
Pengamatan tabel hasil dilakukan dari *frame* ke-1 sampai *frame* ke-700 kemudian dapat ditarik kesimpulan bahwa semakin kecil nilai *thresh* maka bagian putih semakin dominan jika semakin besar nilai t*hresh* semakin kecil bagian putih. (bagian putih terdeteksi sebagai objek yang bergerak). Dari hasil koresponden, bobot nilai *threshold* yang memiliki akurasi tertinggi sebesar 97.5% dan bobot nilai terbaik adalah *threshold* = 20. Pengujian *threshold* ini menggunakan skala Likert dalam pengujiannya karena variabel tersebut tidak dapat diukur secara langsung tetapi menggunakan kuisioner dan jawaban yang tersedia dengan skala ordinal (skala Likert) dengan menggunakan 4 tingkat skala alternatif jawaban. Contohnya 1=tidak baik, 2=cukup, 3=baik, 4=sangat baik, disebut ordinal karena pernyataan sangat baik mempunyai tingkat yang lebih tinggi terhadap baik dan baik lebih tinggi terhadap cukup, dan seterusnya.
Dari 10 koresponden yang dipilih didapatkan hasil *threshold*=10 bobot nilai=31, *threshold*=20 bobot nilai=39, *threshold*=40 bobot nilai=34, *threshold*=100 bobot nilai=10.



Gambar 4. 3 Koresponden Bobot Nilai Threshold

4.2.3.2 Pengujian Ekstraksi Ciri

Berikut ini adalah penjelasan singkat mengenai metode *Local Binary Pattern*



Gambar 4. 4 Ilustrasi LBP *neighbourhood* 4

Berikut ini merupakan keterangan dari gambar di atas:

a. Setelah gambar diubah ke dalam *grayscale image*
b. Diambil 4 *sampling points* dengan radius 1 *of the pixel in the block.*
c. Pola biner dari *LBP pattern of the pixel*.

4.2.3.3 Pengujian Sistem

Pengujian sistem ini terbagi menjadi beberapa bagian diantaranya adalah pengujian sistem secara keseluruhan dengan modifikasi metode LBP, tanpa modifikasi, pengujian selain ruangan multimedia LC, dan pengujian di koridor.

Tabel 4. 3 Pengujian ekstraksi ciri LBP dengan *neighbourhood* 4

Data uji	Kasus Orang Masuk	TP	TN	FP	FN	Akurasi (%)	Rata-rata Waktu eksekusi / frame
Data Uji 1	1 Orang	1	0	0	0	100%	1.18s
	2 Orang	2	0	0	0	100%	2.21s
	3 Orang	3	0	0	0	100%	3.30s
	4 Orang	4	0	0	0	100%	4.44s
	5 Orang	5	0	0	0	100%	5.53s
	6 Orang	5	0	0	1	83,33%	6.68s
	7 Orang	5	0	0	2	71,42%	7.76s
Data Uji 2	1 Orang	1	0	0	0	100%	1.73s
	2 Orang	2	0	0	0	100%	4.18s
	3 Orang	2	0	0	1	66,66%	4.63s
	4 Orang	3	0	0	1	75%	4.51s
	5 Orang	3	0	0	2	60%	6.82s

Dari hasil analisis pengujian data dengan menggunakan metode Multilevel HOG-LBP *neighbourhood 4* didapatkan hasil pengujian untuk data Uji 1 = 100% untuk kasus masuk 1-5 orang (duduk seling satu) namun akurasi menurun ketika kasus masuk 6-7 orang dan pengujian Data Uji 2 = 100% untuk kasus masuk 1-2 orang(duduk berhimpitan) dan terjadi penurunan akurasi ketika kasus masuk 3-5 orang. Penurunan akurasi terjadi karena ketika objek saling berhimpitan yang menyebabkan *blob* semakin besar sehingga untuk proses pemisahan objek menjadi lebih kompleks. Rata-rata akurasi data Uji 1 = 93.53%, rata-rata akurasi data uji 2= 80.33%.

Tabel 4. 4 Tabel Pengujian Keseluruhan Sistem

Data uji	Kasus Orang Masuk	TP	TN	FP	FN	Akurasi (%)	Rata-rata Waktu eksekusi / frame
Data Uji 1	1 Orang	1	0	0	0	100%	1.16s
	2 Orang	2	0	0	0	100%	2.20s
	3 Orang	3	0	0	0	100%	3.21s
	4 Orang	4	0	0	0	100%	4.39s
	5 Orang	5	0	0	0	100%	5.40s
	6 Orang	5	0	0	1	83,33%	6.58s
	7 Orang	5	0	0	2	71,42%	7.60s
Data Uji 2	1 Orang	1	0	0	0	100%	1.49s
	2 Orang	2	0	0	0	100%	2.48s
	3 Orang	2	0	0	1	66,66%	4.00s
	4 Orang	3	0	0	1	75%	5.31s
	5 Orang	3	0	0	2	60%	6.67s

Dari hasil analisis pengujian data dengan menggunakan metode Multilevel HOG-LBP didapatkan hasil pengujian untuk data Uji 1 = 100% untuk kasus 1-5 orang (duduk seling satu) namun menurun ketika kasus 6-7 orang dan pengujian Data Uji 2 = 100% untuk kasus 1-2 orang(duduk berhimpitan) dan menurun ketika kasus 3-5 orang. Rata-rata akurasi data Uji 1 = 93.53%, rata-rata akurasi data uji 2= 80.33%. Untuk hasil akurasi antara pengujian sistem modifikasi LBP dengan pengujian sistem tanpa modifikasi LBP hasil akurasinya sama namun yang menjadi perbedaannya adalah di bagian waktu akses orang setiap kasus.

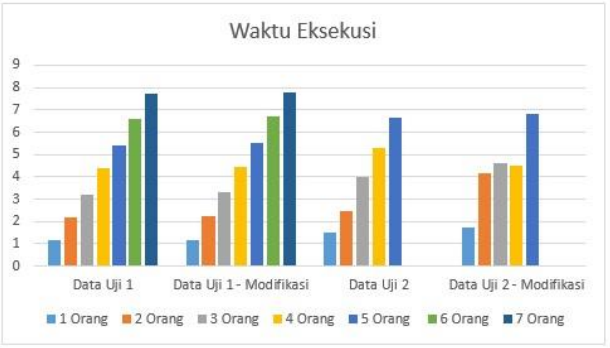
Waktu eksekusi setiap *frame* pada pengujian sistem dengan modifikasi LBP lebih lama daripada pengujian sistem tanpa modifikasi LBP ini disebakan karena metode ekstraksi ciri LBP lebih sedikit menangani *negihbourhood*nya sehingga menyebabkan fitur *vector* setiap level menjadi berkurang. Inilah yang menyebabkan sistem sedikit lebih lambat dalam menentukan apakah objek ini merupakan objek manusia atau bukan. Sehingga kesimpulan akhirnya adalah untuk metode ekstraksi ciri yang baik dalam segi akurasi dan segi waktu eksekusi setiap *frame* adalah metode ekstraksi ciri *Histogram Of Oriented Gradients – Local Binary Pattern* (HOG-LBP) tanpa modifikasi.

Untuk melihat perbedaan secara terperinci perbandingan waktu dapat dilihat melalui gambar di bawah ini:

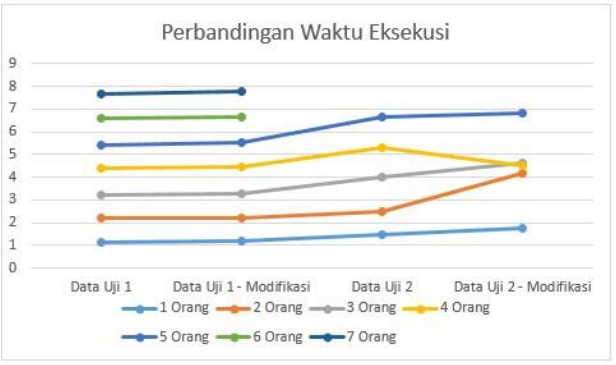
Tabel 4. 5 Perbandingan waktu eksekusi frame

Data uji	Kasus Masuk	Akurasi (%)	Rata-rata Waktu eksekusi / frame (tanpa modifikasi)	Rata-rata Waktu eksekusi / frame (dengan modifikasi)
Data Uji 1	1 Orang	100%	1.16s	1.18s
	2 Orang	100%	2.20s	2.21s
	3 Orang	100%	3.21s	3.30s
	4 Orang	100%	4.39s	4.44s
	5 Orang	100%	5.40s	5.53s
	6 Orang	83,33%	6.58s	6.68s
	7 Orang	71,42%	7.60s	7.76s
Data Uji 2	1 Orang	100%	1.49s	1.73s
	2 Orang	100%	2.48s	4.18s
	3 Orang	66,66%	4.00s	4.63s
	4 Orang	75%	5.31s	4.51s
	5 Orang	60%	6.67s	6.82s

Berikut ini adalah gambar diagram batang pengujian LBP tanpa modifikasi dan pengujian LBP dengan modifikasi untuk data uji 1 dan data uji 2.



Gambar 4. 5 Diagram Batang waktu eksekusi



Gambar 4. 6 Diagram Garis waktu eksekusi

Dapat ditarik kesimpulan bahwa untuk pengujian sistem dengan modifikasi LBP *neighbourhood* 4 waktu akses sistem yang dibutuhkan untuk mendeteksi objek lebih lama daripada pengujian sistem tanpa modifikasi.

Tabel 4. 6 Pengujian Sistem Masuk dan Keluar Tanpa Modifikasi

Data uji	Kasus Orang Masuk-Keluar	TP	TN	FP	FN	Akurasi (%)
Data Uji 1	1 Orang	1	0	0	0	100%
	2 Orang	2	0	0	0	100%
	3 Orang	3	0	0	0	100%
	4 Orang	4	0	0	0	100%
	5 Orang	4	0	0	1	80%
	6 Orang	5	0	0	1	83,33%
	7 Orang	5	0	0	2	71,42%
Data Uji 2	1 Orang	1	0	0	0	100%
	2 Orang	2	0	0	0	100%
	3 Orang	2	0	0	1	66,66%
	4 Orang	3	0	0	1	75%
	5 Orang	3	0	0	2	60%

Dari tabel di atas akurasi untuk setiap data uji 1 dan 2 sama semua hasilnya dengan pengujian sistem orang masuk saja namun pada data uji 1 kasus orang masuk 5

dan keluar 5 terdapat satu FN dimana seharusnya ada objek manusia namun tidak terdeteksi orang.

Studi kasus yang diangkat dalam tugas akhir ini adalah studi kasus perhitungan orang di dalam ruangan kelas. Alasannya adalah *environment* yang terdapat di dalam ruangan kelas adalah *environment* yang ideal dimana pencahayaannya homogen dan tidak tedapat gerakan semu dari objek bukan manusia(contohnya pohon bergerak). Selain menggunakan dan melakukan pengujian terhadap data uji yang memiliki *environment* yang ideal, penulis juga akan melakukan pengujian terhadap *environtment* yang tidak ideal.

Pengujian ini menggunakan data uji di ruangan Multimedia LC dimana pencahayaannya kurang terang sehingga membuat noise di dalam video semakin terlihat jelas. Video ini diambil dengan tinggi kamera ± 2 meter, dan sudut ± 30º

Gambar 4. 7 Ruangan Multimedia LC Pengujian

dilakukan dengan menggunakan data uji duduk
seling satu dimana setiap objek yang terdeteksi ada yang mengenakan jilbab, jaket, dan kemeja putih

Tabel 4. 7 Pengujian Ruangan Multimedia LC

Data uji	Kasus Orang Masuk	TP	TN	FP	FN	Akurasi (%)
Data Uji 1	1 Orang	1	0	0	0	100%
	2 Orang	2	0	0	0	100%
	3 Orang	3	0	0	0	100%
	4 Orang	3	0	0	1	75%
	5 Orang	3	0	0	2	60%
	6 Orang	3	0	0	3	50%
	7 Orang	3	0	0	4	42,85%
	8 Orang	3	0	0	5	37,5%

Pengujian untuk kasus orang 1 sampai orang 3 masuk ke dalam area kursi tingkat akurasi berada di angka 100% namun ketika masuk orang 4 dan selanjutnya maka tingkat akurasi semakin menurun. Ini disebabkan karena objek manusia yang terdapat di dalam video sama gelapnya dengan *background* sehingga jika terjadi gerakan pasif ketika objek yang telah terhitung tersebut duduk maka gerakan pasif itu tidak akan terlihat secara jelas dan akan dianggap sebagai *background*.

Sebenarnya ketika orang tersebut masuk dan belum memasuki kawasan kursi baik orang yang mengenakan jilbab, jaket, dan kemeja putih biasa pada dasarnya sistem sudah mampu mengenali bahwa mereka adalah objek manusia yang terdeteksi sebagai manusia.

Namun permasalahannya muncul ketika objek tersebut memasuki area kursi karena muncul permasalahan oklusi, *foreground* yang bergerak pasif akan dianggap *background* kembali. Dalam hal ini oklusi terjadi karena filter *Gaussian* dan *morphological* erosi dan dilasi belum terlihat sempurna karena banyaknya *noise* dalam data uji video sehingga hasil akhir filter video menjadi kurang bersih dan mempengaruhi hasil akhir akurasi.

Gambar 4. 8 Hasil Uji Multimedia LC

Pada gambar di atas orang yang menggunakan jilbab tidak terdeteksi sebagai objek sedangkan pada saat objek tersebut masuk dan duduk sebelum ada objek lain masuk, objek tersebut terdeteksi sebagai objek.
Oleh karena itulah konklusi ruangan yang paling ideal adalah ruangan B107 dimana kursinya memiliki level ketinggian yang berbeda, pencahayaan yang baik dan homogen, ventilasi udara yang memadai sehingga meminimalisir angin masuk kedalam ruangan kelas dan menggerakan tirai di kelas yang dapaat dianggap sebagai gerakan semu.

Pengujian selanjutnya adalah pengujian di koridor. Untuk menambah hasil pengujian yang lebih variatif maka ditambahkan data uji kasus bukan di ruangan kelas. Data uji ini diambil di koridor gedung F lantai 2 dengan tinggi ±2,3 meter dan sudut pengambilan video ± 30º

Gambar 4. 9 Pengujian Koridor

Berikut ini adalah hasil pengujian di koridor:

Tabel 4. 8 Pengujian Koridor

Data Uji	Kasus Orang masuk bersamaan	TP	TN	FP	FN	Akurasi (%)
Data Uji 1	1 Orang	1	0	0	0	100%
	2 Orang	1	0	0	1	50%
	3 Orang	1	0	1	2	25%

Data uji tertinggi didapat ketika jumlah orang yang berjalan di video hanya satu orang satu orang namun akurasi menurun ketika orang yang masuk dan jalan secara bersamaan. Pada kasus 3 orang masuk secara bersamaan terdapat 1 FP dimana bayangan orang di dalam ruangan terdeteksi sebagai manusia. Ada begitu banyak perubahan *environment* yang terjadi ketika pengambilan video selain di ruangan kelas contohnya seperti bayangan orang, perubahan cahaya di dalam koridor dan arah gerak objek manusia yang bervariatif karena itulah sistem ini cocok diterapkan pada studi kasus di dalam ruangan kelas.

5. Kesimpulan dan Saran

5.1. Kesimpulan

Berdasarkan percobaan yang telah dilakukan, maka dapat ditarik beberapa kesimpulan:

1. Dari skenario pengujian, akurasi tertinggi dan waktu akses setiap data uji terbilang cepat jika manusia yang berada di ruangan kelas sedikit. Untuk data Uji 1 akurasi mencapai 100% untuk kasus 1-5 orang (duduk seling satu) namun akurasi menurun ketika kasus 6-7 orang dan pengujian Data Uji 2 akurasi mencapai 100% untuk kasus 1-2 orang (duduk berhimpitan) dan akurasi menurun ketika kasus 3-5 orang.
2. Semakin banyak manusia di dalam ruangan kelas maka semakin lama waktu rata-rata eksekusi untuk akses per *frame*. Untuk pengujian sistem dengan modifikasi LBP *neighbourhood* 4 waktu akses sistem yang dibutuhkan untuk mendeteksi objek lebih lama daripada pengujian sistem tanpa modifikasi.

5.2. Saran

1. Untuk proses *background substraction* diperlukan metode yang lebih adaptif untuk menangani kasus oklusi sehingga dapat menangani kasus-kasus dimana terdeteksi objek terlihat berhimpitan satu sama lain atau terdeteksi gerakan semu dari latar video. Salah satu metode *background substraction* yang cukup adaptif adalah *Gaussian Mixture Model.*
2. Menggunakan data latih yang memiliki *background* yang mirip dengan orang-orang yang berada di ruangan kelas.

.

6. Daftar Pustaka

[1] Bradsky, G., & Kaehler, A. 2008. *Learning OpenCV Computer Vision with the OpenCV Library.* O'Reilly Media.

[2] C. Srinivas Rao, P.Darwin. Frame Difference And Kalman Filter Techniques For Detection Of Moving Vehicles In Video Surveillance. In IJERA,2012. Vol. 2, Issue 6, November-December 2012,pp.1168-1170.

[3] C. Zeng and H. Ma, "Robust head-shoulder detection by pcabased multilevel hog-lbp detector for people counting," in Pattern Recognition (ICPR), 2010 20th International Conference on. IEEE, 2010, pp. 2069–2072.

[4] Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Volume 1. (2005) 886–89

[5] Gonzales, R. C., & Woods, R. E. (2002). Digital Image Processing Second Edition. Pearson Pretince Hall

[6] Gonzales, R. C., & Woods, R. E. (2002). Digital Image Processing Using Matlab. Pearson Pretince Hall

[7] Hoin, B. K. & Schunck, B. G. (1980). Determining Optical Flow Artificial Intelligence. (pp. 185-203)

[8] Hongyu Liang,Jinchen Wu and Kaiqi Huang, "People in Seats Counting via Seat Detection for Meeting Surveillance" Proceedings of Chinese Conference on Pattern Recognition (CCPR), pp. 202-210, 2012.

[9] Johnson, S. (2006). Stephen Johnson on Digital Photography. O'Reilly

[10] J.S.C. Yuk, K.K.Y. Wong, R.H.Y. Chung, FYL Chin, and KP Chow, "Real-time multiple head shape detection and track-ing system with decentralized trackers," in Intelligent Systems Design and Applications, 2006. ISDA'06. Sixth International Conference on. IEEE, 2006, vol. 2, pp. 384–389

[11] Kumar, Rakesh, Tapesh Parashar, Gopal Verma, 2012, *Background Modeling and Subtraction Based People Counting for Real Time Video Surveillance*, in: International Journal of Soft Computing and Engineering.

[12] M. Li, Z. Zhang, K. Huang, and T. Tan, "Estimating the number of people in crowded scenes by mid based foreground segmentation and head-shoulder detection," in Pattern Recognition, 2008. ICPR 2008. 19th International Conference on. IEEE, 2008, pp. 1–4.

[13] Nugroho, Anto Satriyo, Witarto, Arif Budi, Handoko,Dwi. Support Vector Machine Teori dan Aplikasinya dalam Bioinformatika. http://asnugroho.net/papers/ikcsvm.pdf

[14] Piccardi, Massimo. 2004. Background substraction techniques: a review. IEEE International Conference on Systems, Man and Cybernetics. University of Technology, Sydney (UTS), Australia.

[15] Pietikäinen, Matti, Abdenour Hadid, Guoying Zhao, and Timo Ahonen. Computer vision using local binary patterns. Vol. 40. Springer, 2011.

[16] Raja David Hasugian. *Tugas Akhir : Perhitungan Aliran Orang secara Waktu Nyata dengan Pendekatan Deteksi Kepala (People Counting)*. Bandung : IT Telkom, 2013

[17] T. Ojala, M. Pietikäinen and T. Mäenpää. Multiresolution gray scale and rotation invariant texture analysis with local binary patterns. In PAMI, 2002.

[18] Tu. Jinhui, Zhang Cao and Hao Pengwei. Robust Real-Time Attention-Based Head-Soulder Detection For Video Surveillance. Beijing: Peking University

[19] X. Y. Wang, X. Han and S. C. Yan. An HOG-LBP human detector with partial occlusion handling. In ICCV, 2009.

[20] Y. Takayanagi and J. Katto, "Human body detection using HOG with additional color features, International Workshop on Advanced Image Technology, 2010.

[21] Zheng Yongbin, Shen Chunhua, and Hartley Richard, Fellow. Effective Pedestrian Detection Using Center-symmetric Local Binary/Trinary Patterns. In IEEE, September 2010.

[22] Zhu, Wen, Nancy Zeng, and Ning Wang, 2010. "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS® implementations." NESUG proceedings: health care and life sciences, Baltimore, Maryland (2010).