

Penghitungan Orang dengan Metode *Gaussian Mixture Model* dan *Human Presence Map*

Studi Kasus: Penghitungan Orang di dalam Kelas

Danu Hary Prakoso¹, Bedy Purnama², Febryanti Sthevanie³

^{1,2,3} Departemen Informatika, Telkom Engineering School, Telkom University
Jalan Telekomunikasi No.1, Dayeuh Kolot, Bandung 40257 dharyprakoso@gmail.com¹,
bedy.purnama@telkomuniversity.ac.id², febryantisthevanie@gmail.com³

Abstrak

Penggunaan teknologi dewasa ini telah banyak digunakan dalam kehidupan sehari-hari yang mungkin kehadirannya bahkan tidak terasa oleh masyarakat. Salah satunya adalah penggunaan CCTV yang dapat berfungsi sebagai suatu perangkat pengawasan. Namun disamping fungsi utamanya tersebut, CCTV sebagai kepanjangan dari mata orang bisa digunakan untuk melakukan penghitungan orang.

Penulis dalam Tugas Akhir ini membahas tentang penghitungan orang dengan studi kasus spesifik yakni menghitung orang dalam suatu kelas. Manfaat yang bisa diambil dalam mengambil estimasi jumlah orang dalam suatu kelas antara lain untuk mengecek apakah kehadiran sesuai dengan jumlah orang dalam kelas. Penghitungan orang dapat dilakukan dengan menggunakan *background subtraction* yakni menggunakan metode *Gaussian Mixture Model* (GMM) untuk melakukan ekstraksi antara objek yang ingin diamati, dalam hal ini orang dengan *background*. Selanjutnya proses klasifikasi melalui *Human Presence Map* untuk memfokuskan pendeteksian piksel dimana orang berada sehingga dapat dilakukan penghitungan orang berdasarkan suatu area wilayah yang telah ditentukan. Dengan metode tersebut, penghitungan orang dapat mencapai tingkat akurasi sebesar 91% untuk kasus deteksi satu orang dalam video uji duduk terpisah.

Kata Kunci: *people counting*, ruang kelas, *gaussian mixture model*, *human presence map*, piksel

Abstract

Technology has been widely used in daily life that may not even its presence is felt people. One of it is the use of CCTV which can serve as a monitoring devices. But besides its main function, CCTV as an extended part of eyes can be used to perform people counting.

The author in this Final Project discusses about counting people with specific case studies, that is counting people in a classroom. The benefits that can be taken in making estimates of the number of people in a class, among others, is to be able to check whether the presence form is in accordance with the number of people in the class. Counting people can be done using background subtraction background with Gaussian Mixture Model (GMM) method to perform the extraction of the object to be observed, in this case is the people and the background. Furthermore, the classification process is done through the Human Presence Map to focus detection pixels where people is located so that we can count people based on a region of predetermined area. With the method proposed, people counting could reach an accuracy of 91% for case of 1 person detection in test video with separate sitting.

Keywords: *people counting*, classrooms, Gaussian mixture models, human presence map, pixel

1. PENDAHULUAN

People counting merupakan upaya menghitung orang dalam suatu wilayah tertentu yang saat ini merupakan masalah yang menantang dan krusial dalam pengawasan secara visual. Hal ini dapat digunakan di berbagai tempat umum seperti untuk mengawasi antrian tiket dalam acara publik seperti konser musik, kepadatan dalam pusat perbelanjaan maupun toko untuk keamanan maupun pengoptimalan pemasaran, dan alokasi sumber daya [1][4][5][7][9][12], serta estimasi jumlah mahasiswa di dalam kelas. Secara sederhana tahapan umum

dari people counting antara lain yaitu penginputan video ke dalam sebuah sistem untuk dilakukan analisis, pendeteksian objek berupa orang, dan algoritma penghitungan. Sebelum ini, proses penghitungan orang memang belum handal karena melibatkan SDM secara manual, atau dengan menggunakan berbagai alat-alat elektronik yang berbasis sensor infra merah ataupun melalui analisis video yang ringan [14], namun tentu hal tersebut tidak efisien dan tidak terlepas dari margin kesalahan yang cukup tinggi.

Dalam beberapa tahun belakangan ini, bidang people counting telah berkembang dengan pesat namun solusi yang

ditawarkan masih banyak terdapat keterbatasan, seperti orang yang menjadi objek harus bergerak, latar tempat harus sederhana, atau resolusi dari gambar harus tinggi.

Berbagai penelitian sudah banyak dilakukan dalam people counting, seperti melakukan penghitungan orang secara outdoor dengan menggunakan gambar beresolusi rendah menggunakan metode Expectation Maximization (EM) [4]. Selain itu ada pula yang memfokuskan pada penghitungan orang dalam suatu grup kelompok orang dengan mengedepankan metode pixel layering [3], penghitungan orang dengan SW-HOG [8] untuk menghitung orang yang sedang duduk, dan dengan integral image [16] untuk dapat mengekstrak fitur orang dengan cepat.

Tugas Akhir ini mengambil studi kasus people counting pada area ruang kelas. Manfaat yang bisa diambil antara lain untuk mengecek apakah jumlah kehadiran sesuai dengan jumlah orang yang berada di kelas pada saat itu secara cepat, karena biasanya penghitungan terjadi secara manual menggunakan presensi dengan memanggil nama orang satu persatu. Namun dalam penerapannya, terdapat berbagai kendala. Dalam situasi ruang kelas, objek yang diamati nyaris tidak bergerak (stationary) sehingga bisa keliru diinterpretasikan sebagai background, dan lain sebagainya. Untuk itu penulis menerapkan berbagai metode seperti background subtraction dengan Gaussian Mixture Model (GMM). GMM digunakan untuk mendeskripsikan piksel background, sehingga menyediakan adaptasi yang robust menghadapi gerakan berulang dalam elemen adegan, objek yang bergerak lambat, dan menambahkan atau menghilangkan obyek dari adegan [12]. Lalu digunakan pula Human Presence Map untuk mendapatkan lokasi peta objek orang berada, mencegah orang termasuk ke dalam model background, mengurangi learning frame yang dibutuhkan, serta membuang pendeteksian foreground palsu [2], serta untuk melakukan penghitungan orang. Diharapkan dengan sistem ini, dapat dilakukan penghitungan orang di dalam kelas dengan lebih akurat.

2. KAJIAN LITERATUR

2.1 Video Digital

Video digital merupakan kegiatan menangkap, memanipulasi, dan menyimpan gambar yang dapat bergerak sehingga dapat diamati dalam layar monitor [6]. Video digital terdiri dari serangkaian gambar digital yang ditampilkan secara cepat dengan laju yang konstan. Gambar-gambar digital ini disebut frame. Satuan yang digunakan untuk mengukur kecepatan sebuah video digital adalah frame per second (FPS).

Setiap citra digital ada video digital memiliki lebar dan tinggi yang sama mulai dari citra pertama sampai citra terakhir. Urutan gambar tersebut menjadi sebuah video ketika ada kesinambungan antara satu frame dengan frame yang lain.

2.2 People Counting

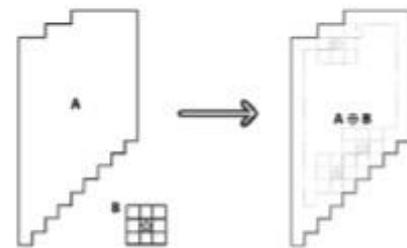
People counting merupakan masalah yang krusial dan menantang dalam pengawasan secara visual [4]. Hal ini bisa digunakan di berbagai tempat umum seperti pusat perbelanjaan, stasiun angkutan umum, maupun teater untuk masalah keamanan, pengoptimalan pelayanan, dan alokasi

sumber daya [5]. People counting umumnya menggunakan media real-time semisal kamera video [1]. Namun sistem yang robust seharusnya tidak bergantung kepada penempatan kamera [13].

Secara sederhana tahapan umum dari people counting antara lain yaitu penginputan video ke dalam sebuah sistem untuk dilakukan analisis, pendeteksian objek yang bergerak, tracking, dan algoritma penghitungan [10].

2.3 Dilasi

Dilasi merupakan bagian dari operasi morfologi pada pengolahan citra digital (image processing), yakni melakukan penambahan piksel pada tepi citra dengan intensitas piksel yang sama berdasarkan pada penggunaan kernel. Kernel dalam hal ini merupakan alat ukur untuk membantu penambahan piksel dengan ukuran minimal 3x3 piksel. Gambar berikut menunjukkan hasil penambahan sebuah citra A dengan menggunakan kernel B berukuran 3x3 piksel.

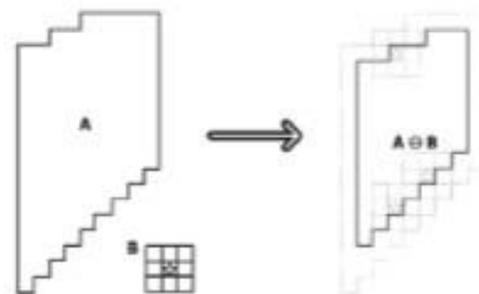


Gambar 2.1 Penambahan citra A dengan kernel B berukuran 3x3 piksel

Dalam sistem ini, dilasi digunakan bersama dengan erosi sehingga dapat dilakukan morfologi open maupun close pada connected component analysis.

2.4 Erosi

Erosi merupakan bagian dari operasi morfologi pada pengolahan citra digital (image processing), yakni dengan melakukan pengurangan piksel pada tepi citra dengan intensitas piksel yang sama berdasarkan kepada penggunaan kernel. Kernel dalam hal ini merupakan alat ukur untuk membantu pengurangan piksel dengan ukuran minimal 3x3 piksel. Gambar berikut menunjukkan hasil pengurangan sebuah citra A dengan menggunakan kernel B berukuran 3x3 piksel.



Gambar 2.2 Pengurangan citra A dengan kernel B berukuran 3x3 piksel

Dalam sistem ini, erosi digunakan untuk bersama dengan dilasi sehingga dapat dilakukan morfologi open maupun close pada connected component analysis.

2.5 Background Subtraction

Sebelum menghitung orang, setiap objek yang bergerak harus terlebih dahulu diekstrak dari background [4]. Background subtraction merupakan proses manipulasi sebuah gambar background untuk mendapatkan objek yang diinginkan, umumnya objek yang bergerak atau dalam people counting merupakan orang yang akan dihitung. Terdapat beberapa metode untuk melakukan hal ini, seperti frame difference yang merupakan metode yang paling sederhana. Metode ini mengekstrak objek yang bergerak sebagai elemen foreground yang diambil dari perbedaan gambar antara setiap frame dan dikenal sebagai background model dari sebuah adegan.[6]

2.5.1 Gaussian Mixture Model (GMM)[9]

Gaussian Mixture Model merupakan salah satu metode dalam background subtraction. Metode ini digunakan untuk mendeskripsikan piksel dari background. Model ini dapat menerima multimodal background, sehingga merupakan model yang robust terhadap gerakan berulang dalam elemen latar, objek yang bergerak lambat, dan memperkenalkan atau menghapus objek dari latar [11].

GMM akan memberikan fungsi-fungsi komponen Gaussian untuk tiap piksel, dengan input adalah warna piksel dimana model-model GMM terbentuk berdasarkan waktu. Model akan membentuk 2 komponen utama, yakni model background dengan model non-background [12] atau bisa disebut sebagai foreground. Model background adalah model yang mencerminkan latar dari area yang diamati, sedangkan model foreground merupakan model yang mencerminkan objek yang bisa diamati.

Pengukuran sebelumnya dari setiap piksel, {Xi, ..., Xt} dapat dimodelkan menggunakan distribusi campuran dari K Gaussian. Kemungkinan untuk mengamati background saat itu dari sebuah piksel Xt adalah jumlah bobot dari distribusi K:

$$p(x_t) = \sum_{k=1}^K \omega_k \cdot \eta(x_t | \mu_k, \Sigma_k) \quad (2.1)[11]$$

Dimana K adalah nomor dari distribusi Gaussian, ω_k adalah perkiraan berat dari distribusi Gaussian ke ith saat waktu t, μ_k adalah rata-rata dari nilai Gaussian ke ith saat waktu t, Σ_k adalah matriks kovariansi dari Gaussian ke ith saat waktu t, dimana η adalah fungsi kemungkinan kepadatan dari Gaussian ke ith, yang diberikan oleh:

$$\eta(x_t | \mu, \Sigma) = \frac{1}{(2\pi)^2 |\Sigma|} \exp(-1/2(x_t - \mu) \Sigma^{-1} (x_t - \mu)) \quad (2.2)[11]$$

K dapat ditentukan oleh komputasi dan kekuatan memori. Contohnya terdapat tiga distribusi Gaussian untuk mendeskripsikan sebuah piksel. Untuk menghindari biaya komputasi diasumsikan variasi dari warna Red, Green, Blue (RGB) sama, sehingga kovariansi matriks dapat didefinisikan sebagai:

$$\Sigma_{RGB} = \sigma^2 I \quad (2.3)[11]$$

Algoritma memasukkan satu GMM ke tiap piksel dari gambar dan memperbaharui parameter dari model (nilai rata-rata dan variansi) secara langsung. Setiap kali sebuah piksel baru didapatkan piksel tersebut dicek dengan distribusi Gaussian K yang sudah ada. Cocok jika:

$$x_t - 2.5 \cdot \sigma < \mu < x_t + 2.5 \cdot \sigma \quad (2.4)[12]$$

Dimana x_t merupakan vektor dari warna piksel (R,G,B) untuk waktu t, μ merupakan vektor nilai mean (R,G,B) dari Gaussian ke k, σ adalah standar deviasi dari Gaussian ke k. Suatu piksel dianggap sebagai foreground apabila piksel tidak cocok dengan semua distribusi yang ada dan dibuat distribusi baru dengan menggantikan distribusi yang distribusi yang paling tidak mencerminkan background.

Jika tidak ada nilai piksel yang cocok dengan distribusi K, distribusi yang paling mungkin digantikan dengan distribusi dengan nilai yang ada sekarang sebagai nilai rata-ratanya. Bobotnya diatur sebagai berikut:

$$\omega_k = (1 - \alpha) \omega_{k-1} + \alpha M_k \quad (2.5)[11]$$

Dimana α merupakan tingkat learning, M_k adalah 1 jika ada kecocokan dan 0 jika kecocokan tidak ditemukan. Parameter μ dan σ untuk distribusi yang tidak cocok tetap sama. Untuk distribusi yang cocok parameter tersebut diperbaharui seperti berikut:

$$\mu_k = (1 - \beta) \mu_{k-1} + \beta x_t \quad (2.6)[11]$$

$$\sigma_k^2 = (1 - \beta) \sigma_{k-1}^2 + \beta (x_t - \mu_k) (x_t - \mu_k)^T \quad (2.7)[11]$$

Dimana

$$\beta = \alpha \eta(x_t | \mu_k, \sigma_k) \quad (2.8)[11]$$

Untuk menentukan Gaussian yang mana dari GMM yang merepresentasikan background, dilakukan penyortiran dengan menggunakan nilai ω / σ . Nilai ini semakin tinggi dengan semakin sedikitnya variansi dari distribusi dan semakin banyaknya distribusi yang digunakan. Hal ini mengarah ke Gaussian dari GMM dimana distribusi yang paling mencerminkan background tetap berada di atas dan distribusi yang kurang mencerminkan background semakin berada di bawah yang nantinya akan digantikan oleh distribusi yang baru. Distribusi B yang pertama dipilih sebagai model background, dimana:

$$B = \arg \max_k (\omega_k / \sigma_k) \quad (2.9)[11]$$

Dimana T adalah threshold tentang proporsi terkecil data yang harus dipertimbangkan sebagai background. Jika T yang terpilih rendah seringkali background yang unimodal akan direpresentasikan tapi dengan T yang lebih tinggi kemungkinan

background yang multi-modal meningkat.

2.6 Human Presence Map [2]

Pada persamaan 2.10, x dan y merupakan koordinat pojok kiri atas dari kotak human region detection, dan w serta h merupakan lebar dan tinggi.

$$\sum_{x=y}^x \sum_{y=y}^y \eta(x, y) \triangleq P \rightarrow (x, y) \triangleq \text{Human Presence Map} \quad (2.10) [2]$$

$$\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \sum_{m=1}^n \sum_{n=1}^n \left(\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \sum_{m=1}^n \sum_{n=1}^n \right) \rightarrow \left(\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \sum_{m=1}^n \sum_{n=1}^n \right) = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \sum_{m=1}^n \sum_{n=1}^n \rightarrow h_{i,j,k,l,m,n} - (\#_{i,j,k,l,m,n} + 1), \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \sum_{m=1}^n \sum_{n=1}^n = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \sum_{m=1}^n \sum_{n=1}^n - (\#_{i,j,k,l,m,n} + 1)$$

$$1) \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \sum_{m=1}^n \sum_{n=1}^n = \frac{\sqrt{\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n \sum_{m=1}^n \sum_{n=1}^n} - 1}{2} \quad (2.11) [2]$$

Human region detector digunakan pada frame utamanya untuk menghilangkan noise dan membuat sistem lebih robust dimana jika tidak terdapat gambar background yang kosong. Analisis komponen yang terhubung dilakukan antara piksel foreground dengan kotak dari deteksi kepala.

Dengan window size sebesar 25x25 piksel (sehingga SIZE = 625 pada persamaan 2.11) dilakukan connected component analysis untuk menghubungkan komponen ke lingkup piksel putih. Komponen yang tidak terhubung dengan human region manapun dihapus, menghasilkan human presence map. Hal ini bertujuan untuk, i) mendapatkan lokasi peta manusia, ii) mencegah manusia untuk tidak terserap ke dalam model background, iii) mengurangi learning frame yang dibutuhkan, dan iv) memfilter deteksi foreground lain yang palsu. Human presence map digunakan untuk, a) mencegah orang untuk terserap ke dalam background, dan b) mencegah non-human-pixels terhitung sebagai human-pixels.



Gambar 2.3 Atas: Image segmentation Bawah: Human Presence Map [2]

Gambar 2.3 memperlihatkan bahwa noise lebih sedikit pada human presence map, yang utamanya merupakan hasil dari connected component analysis yang memfokuskan pendeteksian pada daerah orang saja.

2.6.1 Human Region Detetion

Dengan beberapa contoh, diimplementasikan terlebih dahulu metode human region detection untuk klasifikasi piksel. Kepala dipilih sebagai fitur awal untuk deteksi manusia. Hal ini dikarenakan kepala merupakan bagian yang paling umum terlihat dari manusia [2]. Hal ini dilakukan dengan menggunakan fitur HOG dan melakukan training dengan menggunakan cascade of classifier.

2.7 Histogram of Oriented Gradient

Histogram of Oriented Gradients (HOG) merupakan metode yang digunakan untuk mendapatkan tepian atau

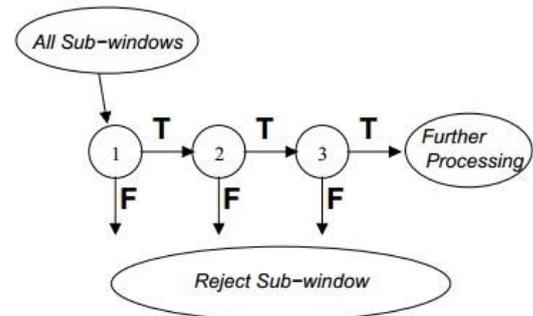
informasi bentuk lokal dari sebuah citra, dimana dimungkinkan adanya perubahan cahaya di dalamnya. Secara sederhana

berikut adalah tahapan dari HOG [17]:

1. Konversi gambar input ke grayscale.
2. Hitung derajat tepi dan kekuatan per piksel.
3. Bagi gambar ke dalam beberapa bagian sel.
4. Bangun histogram dari derajat tepi dan kekuatan piksel.
5. Normalisasi histogram dengan sel tetangga.

2.8 Cascade of Classifier [15]

Cascade of classifier merupakan sebuah metode boosted classifier yang direkosntruksi dengan cara membuang sub-windows negatif ketika mendeteksi hampir seluruh contoh positif. Pengklasifikasian yang lebih sederhana digunakan untuk menolak mayoritas dari sub-windows sebelum pengklasifikasian yang lebih kompleks digunakan untuk mencapai nilai false positive yang rendah.



Gambar 2.4 Skema proses cascade of classifier

Rangkaian pengklasifikasi diaplikasikan kepada seluruh sub-window. Pengklasifikasi awal mengeliminasi contoh negatif dalam jumlah yang banyak dengan pemrosesan yang sangat sedikit. Lapisan pengklasifikasi lainnya mengeliminasi contoh negatif tambahan namun membutuhkan komputasi tambahan pula.

Bentuknya secara umum seperti sebuah degenerate decision tree, yang disebut dengan cascade. Hasil positif dari pengklasifikasi pertama akan memanggil evaluasi dari pengklasifikasi yang kedua. Hasil positif dari pengklasifikasi kedua akan memanggil pengklasifikasi yang ketiga dan seterusnya. Hasil yang negatif pada bagian manapun akan berujung kepada penolakan secara langsung terhadap sub-window yang sedang diproses.

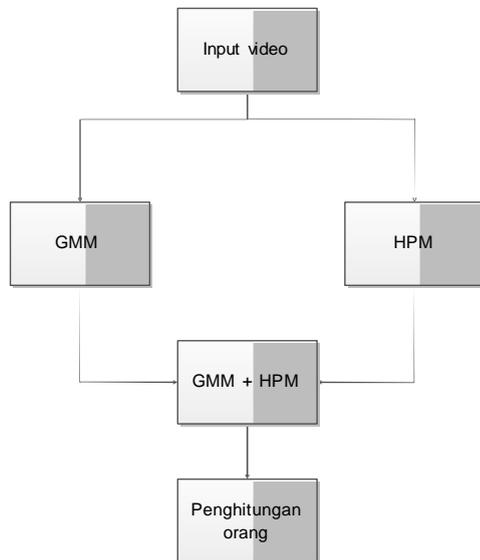
3. METODE/ALGORITMA

3.1 Deskripsi Sistem

Sistem yang dibangun dalam Tugas Akhir ini dibuat melalui beberapa tahap. Tahap yang pertama adalah background subtraction dengan menggunakan metode Gaussian Mixture Model untuk dapat memisahkan antara background dengan foreground pada tiap frame. Disaat bersamaan, digunakanlah Human Presence Map untuk

melakukan ekstraksi ciri sehingga orang dapat terdeteksi. Lalu kedua proses tersebut digabung, dengan Human Presence Map menghilangkan noise dan memfokuskan pendeteksian piksel hanya kepada orang pada frame yang sudah terpisah antara foreground dengan background. Setelah hasil deteksi didapatkan maka langkah selanjutnya adalah melakukan penghitungan orang.

3.1.1 Bagan Alur Sistem



Gambar 3.1 Block process sistem

Input video yang digunakan sama, namun untuk prosesnya terpisah untuk proses GMM dan HPM., yang nantinya akan kembali digabungkan hasil deteksi dari masing-masing proses ke dalam 1 video. Gaussian Mixture Model (GMM) mengekstraksi video uji yang memiliki format Red, Green, dan Blue (RGB) dengan hanya menyisakan 2 komponen utama, yakni foreground yang berwarna putih yakni objek yang terdeteksi bergerak, yang dalam hal ini merupakan orang dan background yang berwarna hitam yang merupakan latar tempat dimana obyek berada. Untuk GMM, pendeteksian dilakukan dengan memberikan bounding box kepada blob foreground. Bounding box ini akan mengikuti kemanapun blob foreground tersebut bergerak.

Human Presence Map (HPM) menggunakan input video yang sama, yakni video uji yang memiliki format Red, Green, dan Blue (RGB). HPM memberikan bounding box kepada bagian piksel yang terdeteksi sebagai orang, dengan fitur yang difokuskan adalah di bagian kepala dan wajah. Fitur ini didapatkan dengan terlebih dahulu melakukan training kepada HPM yang akan dijelaskan lebih rinci di sub bab selanjutnya. Hasil dari training ini akan digunakan langsung ke video uji.

Tahap selanjutnya adalah penggabungan metode antara GMM dengan HPM. Frame yang akan diperlihatkan oleh sistem merupakan video uji RGB dengan hasil deteksi. Oleh sistem, deteksi dilakukan dengan menggunakan frame hasil dari GMM yakni yang terdapat background dengan foreground, dengan mendeteksi blob foreground yang terbentuk pada frame. GMM

akan mendeteksi orang ketika masih bergerak menuju posisi duduk, dan HPM akan memberikan penguatan kepada foreground dengan memberikan warna putih kepada bagian yang terdeteksi oleh HPM, yang dalam hal ini merupakan bagian kepala dari orang. Blob inilah yang akan dihitung oleh sistem sebagai orang.

3.2 Perancangan Sistem

3.2.2 Input Video

Sumber data yang menjadi awal dari tahapan sistem yakni merupakan citra berbentuk video. Proses pengambilan video dilakukan terhadap studi kasus yang akan diamati, yakni ruang kelas dengan beberapa skenario pengujian. Video yang diambil memiliki ukuran 592 x 800 piksel dengan warna RGB untuk menjangkau area yang luas dan menghindari adanya efek blur pada video yang diambil. Video terbagi menjadi 2 bagian, yakni bagian awal yang digunakan sebagai learning pada sistem, dan selain itu merupakan gambar testing.



Gambar 3.2 Data learning



Gambar 3.3 Data testing

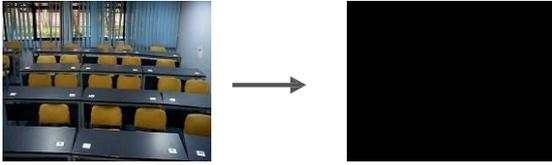
Data learning merupakan gambar background murni, dimana tidak terdapat objek-objek yang bergerak. Data learning ini akan digunakan oleh sistem untuk mendapatkan model background, yang akan lebih detail dijelaskan di sub bab berikutnya. Sedangkan data testing merupakan gambar dimana dalam background terdapat objek-objek yang akan diamati, yang akan digunakan untuk menguji sistem.

3.2.2 Background Subtraction dengan GMM

Untuk memudahkan dalam proses penghitungan orang, langkah awal yang dilakukan dalam sistem adalah dengan melakukan proses segmentasi terhadap input video, yakni proses pemisahan yang bertujuan untuk memisahkan antara background dengan foreground. Pada sistem, proses

segmentasi ini menggunakan Gaussian Mixture Model (GMM).

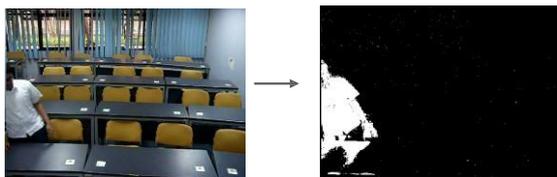
Dalam proses segmentasi ini, pertama-tama sistem akan melakukan fase learning dengan mengambil frame-frame dari input video yang merupakan data learning satu-persatu, dengan jumlah yang bisa ditentukan namun dengan syarat frame-frame tersebut masih dapat dikategorikan sebagai data learning. Untuk setiap framenya, informasi warna dari tiap piksel pada frame akan digunakan untuk membangun model GMM.



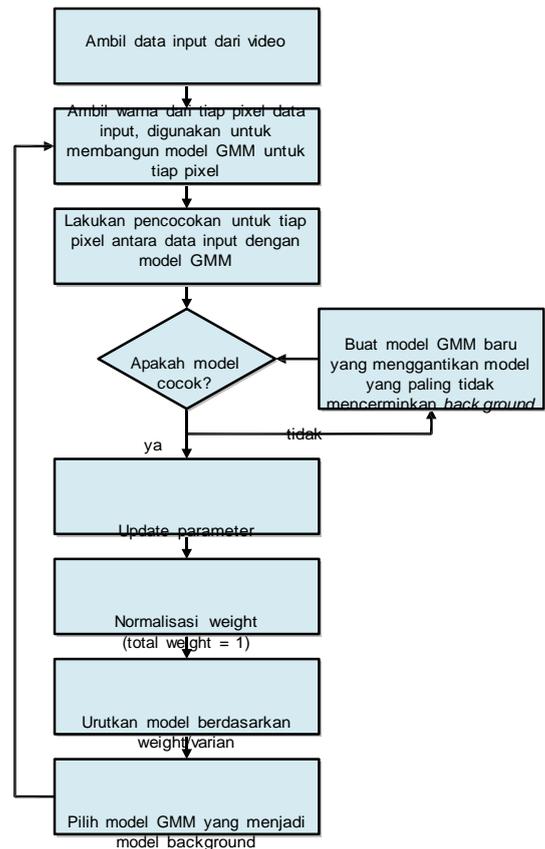
Gambar 3.4 Hasil learning GMM dari 150 frame pertama pada video uji 1 dimana semua piksel diklasifikasikan sebagai background (hitam)

Setelah frame pertama untuk tiap pikselnya telah mendapat model GMM, lalu model GMM tersebut dicocokkan dengan model GMM dari tiap piksel pada frame berikutnya. Jika tidak cocok, maka dibuat model GMM baru untuk menggantikan model GMM yang paling tidak mencerminkan background, yang berarti input tersebut dianggap foreground. Jika terdapat piksel yang cocok, maka dilakukan update terhadap parameter dalam model GMM di piksel tersebut. Diantaranya parameternya adalah weight. Untuk menghindari berlebihnya nilai weight setelah dilakukan proses update parameter, maka dilakukan normalisasi nilai weight sehingga nilainya tidak lebih dari 1. Namun apabila dikatakan suatu piksel cocok dengan model yang bukan model GMM background, maka piksel dinyatakan pula sebagai foreground. Tahap terakhir yakni dilakukannya pemilihan model GMM yang mencerminkan model background.

Keluaran dari tahapan ini adalah frame yang terdiri dari model background yang berwarna hitam, dan model foreground yang berwarna putih, yang disebut dengan blob.



Gambar 3.5 Hasil testing GMM yang didapatkan dari frame ke-150 video uji 1 dimana terdeteksi foreground (putih) pada background (hitam)



Gambar 3.6 Diagram alur Gaussian Mixture Model

3.4 Human Presence Map

Digunakan Metode Human Presence Map digunakan untuk melakukan pendeteksian orang sehingga dapat mencegah agar foreground orang yang tidak bergerak ikut masuk ke dalam background, serta menghilangkan noise agar terminimalisirnya foreground palsu. Untuk dapat memfokuskan pendeteksian piksel hanya pada orang, digunakan ekstraksi fitur HOG dan cascade classifier. Berikut merupakan langkah-langkahnya:

1. Pertama-tama dilakukan training untuk melatih sistem agar dapat mengenali orang. Dataset yang digunakan yakni berasal dari dataset INRIA person, dengan rincian digunakan 460 contoh positif dan 270 contoh negatif berukuran 640x480 piksel dan 480x640 piksel. Contoh positif merupakan gambar yang terdapat orang dan contoh negatif merupakan gambar yang tidak terdapat orang. Untuk orang, difokuskan kepada pendeteksian kepala karena terdapat fitur wajah untuk bisa dideteksi. Dari 460 contoh data positif tersebut, didapatkan 1029 kepala untuk dijadikan acuan deteksi oleh sistem. Bagian yang dijadikan acuan ini setidaknya harus memiliki fitur wajah, seperti misalkan orang yang memalingkan wajah ke samping masih dianggap dapat terdeteksi oleh sistem. Lain halnya dengan orang yang membelakangi kamera dimana hanya terlihat rambut dari orang tersebut, dimana fitur wajah sama sekali tidak terlihat. Pada kasus ini meskipun gambar merupakan contoh positif karena terdapat orang, namun tidak dijadikan acuan deteksi untuk sistem karena justru dapat merusak fitur untuk training secara keseluruhan.



Gambar 3.7 Contoh positif

Contoh negatif yang diberikan untuk sistem merupakan gambar yang tidak terdapat orang, contohnya antara lain adalah gambar jalanan, ruang perkantoran, dan kebun.

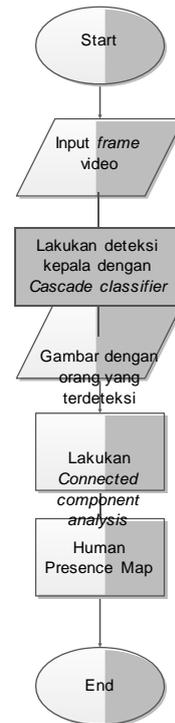


Gambar 3.8 Contoh negatif

2. Dijalankan fungsi ekstraksi fitur HOG sehingga mendapatkan vektor ciri dari fitur wajah dan kepala orang. Vektor ciri ini akan disimpan dan akan digunakan untuk proses pendeteksian.

3. Dengan menggunakan vektor ciri yang telah disimpan sebelumnya, cascade of classifier yang berbasiskan kepada metode Viola-Jones digunakan untuk melakukan training sehingga dapat melakukan pendeteksian orang pada frame.

Setelah orang berhasil terdeteksi, maka dilakukan analisis komponen yang terhubung dengan menggunakan sliding window berukuran 25x25 piksel yang akan memeriksa seluruh frame. Foreground yang tidak terhubung dengan deteksi orang akan dihapus, sehingga hanya akan tersisa foreground dari orang yang terdeteksi.



Gambar 3.9 Diagram alur Human Presence Map pada sistem

3.5 Penghitungan Orang

Proses penghitungan orang merupakan proses terakhir dari sistem ini. Penghitungan orang dilakukan dengan menggabungkan hasil pendeteksian oleh GMM dan HPM, dengan acuan deteksi utamanya adalah dengan menghitung jumlah blob yang terdeteksi pada frame video. Blob ini mendeteksi foreground yang ada pada frame video hasil ekstraksi dari GMM. Penggabungan ini dilakukan untuk mengatasi kelemahan dari masing-masing metode, sehingga tercapai hasil deteksi yang lebih akurat karena kedua metode tersebut dapat saling melengkapi.

Pendeteksian awal menggunakan GMM. Untuk GMM, pendeteksian dilakukan untuk objek yang bergerak, dimana dalam studi kasus ini yang mengalami pergerakan secara radikal adalah orang yang bergerak memasuki wilayah deteksi sampai dengan orang tersebut duduk. Wilayah yang terdeteksi oleh GMM adalah seluruh bagian dari objek yang bergerak, dalam hal ini seluruh tubuh orang yang terlihat pada gambar berikut.



Gambar 3.10 Pendeteksian oleh GMM pada video uji 2

Dengan deteksi awal menggunakan GMM yang mendeteksi objek yang bergerak secara diskriminatif, maka kelemahan HPM dapat tertutupi. GMM secara tepat terus mendeteksi objek yang bergerak tanpa adanya deteksi yang hilang. Hal ini berbeda hal dengan HPM yang masih kurang handal untuk deteksi frame ke frame karena fitur kepala dan wajah yang terdapat pada HPM sebagai hasil training dari dataset tentu tidak semuanya memuat fitur yang ada pada video uji sehingga masih seringkali terdapat detection lost apabila mengandalkan HPM saja.



Gambar 3.11 Pendeteksian oleh HPM pada video uji 2

Ketika orang mulai duduk, secara otomatis gerakannya terbatas sehingga GMM sudah tidak dapat lagi diandalkan. Hal ini dikarenakan orang yang awalnya bergerak memasuki wilayah deteksi ke tempat posisinya duduk masih dianggap sebagai foreground, namun saat orang tersebut duduk maka GMM akan menganggapnya sebagai objek statis karena gerakannya meskipun masih ada tetapi tidak signifikan sehingga perlahan-lahan orang tersebut akan menjadi background. Dalam situasi ini, HPM kemudian mengambil alih pendeteksian dengan hanya mendeteksi bagian kepala dimana terdapat fitur wajah.

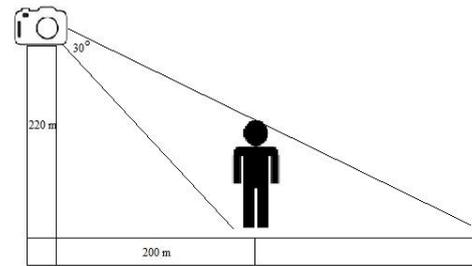
Berbeda dengan hasil pendeteksian GMM yang meliputi seluruh tubuh objek, HPM hanya akan menghasilkan deteksi pada daerah kepala saja. Hal ini mengatasi masalah adanya oklusi yang bisa disebabkan oleh blob yang bergabung ketika ada orang baru masuk ke wilayah deteksi dan posisinya saling berdekatan satu sama lain.frame.

4. HASIL DAN PEMBAHASAN

4.1 Data Uji

Video yang akan digunakan sebagai data uji diambil sendiri oleh penulis sebanyak 3 kali di tempat yang berbeda, yakni Multimedia LC, ruangan kelas B107 dan B109. Setelah dipertimbangkan, diputuskan dari segi kecocokan sebagai tempat skenario seperti intensitas cahaya, kontras dengan latar, posisi kursi bahwa ruangan kelas B107 dan B109 merupakan tempat yang ideal untuk menjadi tempat pengujian.

Pengambilan data dilakukan dengan mengambil video dari arah yang berhadapan dengan orang. Penempatan kamera setinggi $\pm 2,2$ meter dan sudut $\pm 30^\circ$. Video yang didapatkan memiliki ukuran 592x800 piksel dengan 20 fps.



Gambar 4.1 Jarak pengambilan video

Terdapat dua video uji yang digunakan untuk menguji sistem, yakni video uji 1 duduk terpisah dan video uji 2 duduk berdekatan.

Tabel 4.1 Rincian data uji

No	Kasus	Jumlah orang
1	Video uji 1 Duduk terpisah	1
2		2
3		3
4		4
5		5
6	Video uji 2 Duduk berdekatan	1
7		2
8		3
9		4
10		5

4.2 Skenario Uji

Pada penelitian ini terdapat 3 skenario uji utama untuk menghitung performansi sistem, yakni pengujian GMM, pengujian Human Presence Map, lalu pengujian sistem secara keseluruhan untuk menghitung orang. Selain itu terdapat pengujian terpisah antara lain pengujian sistem secara keseluruhan namun dengan ukuran video uji yang diperkecil setengah kalinya dan pengujian GMM untuk perubahan intensitas cahaya.

4.2.1 Pengujian GMM

Proses deteksi dengan GMM merupakan langkah pertama dalam sistem sebelum dapat menghitung orang. Pengujian dilakukan untuk menemukan parameter terbaik untuk GMM yang akan digunakan untuk pengujian berikutnya. GMM untuk sistem perlu berjalan sebagaimana mestinya yaitu hanya dapat mendeteksi obyek yang bergerak maka untuk penghitungannya yang disebut dengan deteksi benar yakni deteksi yang berjumlah satu untuk jumlah orang berapapun. Video yang akan dijadikan untuk pengujian adalah video uji 1.

4.2.2 Pengujian Human Presence Map

Pengujian Human Presence Map dilakukan dengan melihat hasil deteksi orang yang tepat. Sama seperti sebelumnya, untuk HPM pengujian juga dilakukan untuk mencari parameter terbaik yang nantinya akan digunakan pada sistem.

Tabel 4.2 Skenario pengujian Human Presence Map

No.	Jumlah orang
1	1
2	2
3	3
4	4
5	5

Skenario pengujian HPM dilakukan dengan melihat hasil deteksi terhadap video data uji 1 dengan jumlah orang yang terus bertambah seperti terlihat pada tabel 4.2.

4.2.3 Pengujian Sistem dengan Ukuran Video Uji Diperkecil Setengah Kalinya

Pengujian ini sama dengan pengujian sistem namun dengan perbedaan dimana ukuran frame yang digunakan merupakan 1/2 kali ukuran frame dari video uji 1. Hal ini dibarengi pula dengan nilai parameter yang berhubungan dengan jumlah piksel dibuat menjadi setengah kali dari nilai semula.

4.2.4 Pengujian GMM dengan Perubahan Intensitas Cahaya

GMM sangat sensitif terhadap perubahan cahaya, sehingga pengujian ini dilakukan untuk menguji ketangguhan GMM menggunakan video uji 1 dengan intensitas cahayanya diubah, dengan menghitung akurasi orang yang terdeteksi. Perubahan intensitas cahaya dibagi dua, yakni dibuat lebih terang dan dibuat lebih gelap. Masing-masing perubahan intensitas cahaya memiliki nilai 25% dan 50%.

4.2.4 Penghitungan Orang

Penghitungan orang dilakukan dengan menghitung jumlah deteksi benar yang dihasilkan oleh sistem. Untuk menghitung akurasi perhitungan orang ini, digunakan pengklasifikasian biner seperti berikut.

$$Accuracy = 1 - \frac{D}{F} \times 100\% \quad (4.1)[16]$$

Dimana:

1. D adalah jumlah frame dengan deteksi salah untuk jumlah orang ke-n
2. F adalah jumlah frame untuk jumlah orang ke-n. Nilai ini diambil dari selisih frame kemunculan orang ke-n+1 dengan frame kemunculan orang ke-n.

4.3 Hasil Pengujian

4.3.1 Pengujian GMM

Seperti yang sudah dijelaskan sebelumnya, untuk pengujian ini deteksi dianggap benar apabila jumlah deteksi adalah satu untuk jumlah orang berapapun. Hal ini dikarenakan untuk pengujian GMM, GMM diperlukan untuk mempunyai efek yang normal dimana untuk objek yang sebelumnya dianggap sebagai foreground karena bergerak lama kelamaan jika tidak bergerak untuk beberapa lama maka akan dianggap sebagai background.

Berikut merupakan hasil pengujian GMM parameter yang diujikan. Pertama-tama yang diujikan adalah parameter

MinimumBackgroundRatio. Untuk pengujian ini, parameter lain dibuat konstan dimana NumGaussians bernilai 3 dan MinimumBlobArea bernilai 4736.

Tabel 4.3 Hasil pengujian GMM untuk parameter MinimumBackgroundRatio

Jumlah orang	Akurasi Deteksi (%)	
	0.3	0.7
1	93	86.8
2	0	69.6
3	0	69.9
4	0	50.9
5	0	88

Dari pengujian yang telah dilakukan untuk parameter MinimumBackgroundRatio, didapatkan parameter terbaik bernilai 0.7. Nilai ini akan digunakan untuk pengujian berikutnya, yakni pengujian NumGaussians.

Parameter yang bernilai 0.3 memiliki nilai akurasi deteksi yang buruk mencapai 0%. Hal ini dikarenakan dengan menggunakan nilai parameter tersebut, objek foreground meski dalam keadaan tidak bergerak tetap dianggap sebagai foreground, dimana seharusnya lama kelamaan menjadi background. Kesimpulannya adalah semakin kecil nilai parameter MinimumBackgroundRatio, maka semakin lama pula obyek yang sebelumnya dianggap sebagai foreground untuk dapat dianggap sebagai background apabila sudah tidak terjadi gerakan oleh obyek tersebut.

Pengujian selanjutnya adalah pengujian parameter NumGaussians. Untuk pengujian ini, digunakan nilai konstan untuk parameter MinimumBlobArea yakni bernilai 4736, sedangkan MinimumBackgroundRatio menggunakan nilai terbaik dari pengujian sebelumnya yakni 0.7. Berikut adalah tabel pengujian GMM untuk parameter NumGaussians.

Tabel 4.4 Hasil pengujian GMM untuk parameter NumGaussians

Jumlah orang	Akurasi deteksi (%)	
	3	5
1	86.84	86.84
2	69.6	58.5
3	69.9	75.5
4	50.9	49
5	88	59.6

Dari pengujian ini didapatkan parameter terbaik untuk NumGaussians adalah 3. Nilai ini bersama dengan nilai MinimumBackgroundRatio yang sudah diujikan sebelumnya bernilai 0.7 digunakan untuk menguji parameter berikutnya yakni MinimumBlobArea. Berikut adalah tabel pengujian GMM untuk parameter MinimumBlobArea.

Tabel 4.5 Hasil pengujian GMM untuk parameter MinimumBlobArea

Jumlah orang	Akurasi deteksi (%)	
	2368	4736
1	92.1	86.8
2	52.6	69.6
3	48.6	69.9
4	55.7	50.9
5	97.2	88

Jika dilihat dari pengujian-pengujian sebelumnya dan pengujian ini, bisa dilihat bahwa untuk rentang deteksi saat ada 2 hingga 4 orang rata-rata deteksi menurun dibandingkan dengan saat deteksi pada rentang untuk 1 orang dan 5 orang. Hal ini dikarenakan pada video uji 1, pada rentang 2 hingga 4 orang banyak terdapat gerakan dari orang-orang yang sudah duduk semisal meminjam barang dan membuka buku.

Dari pengujian tersebut, didapatkan parameter terbaik untuk parameter MinimumBlobArea adalah 2368 karena unggul dalam 3 segmen jumlah orang. Hasil pengujian tersebut juga dapat dianggap sebagai hasil pengujian untuk GMM dengan parameter terbaik. Dari pengujian ini, didapatkan kesimpulan bahwa parameter terbaik untuk GMM adalah seperti pada tabel berikut.

Tabel 4.6 Parameter terbaik GMM

No.	Parameter	Nilai parameter terbaik
1	Numgaussians	3
2	MinimumBackgroundRatio	0.7
3	MinimumBlobArea	2386

Nilai parameter ini kemudian akan digabungkan dengan parameter terbaik dari HPM untuk menguji sistem.

4.3.2 Pengujian Human Presence Map

Berikut merupakan tabel hasil pengujian menggunakan Human Presence Map untuk mencari parameter terbaik dari HPM.

Tabel 4.7 Hasil pengujian parameter HPM

Jumlah orang	Akurasi deteksi (%)			
	39 32, 10	39 32, 15	49 42, 10	49 42, 15
1	61.2	73.7	72.4	77
2	32.6	43	54.1	55.5
3	23.5	36	44.8	26.6
4	19.4	35.9	51.3	52
5	12.8	20.2	18.3	19.3

Dari hasil pengujian tersebut, didapatkan parameter terbaik HPM seperti pada tabel berikut.

Tabel 4.8 Parameter terbaik HPM

No.	Parameter	Nilai parameter terbaik
1	MinSize	49 42
2	MergeThreshold	15

Nilai parameter terbaik HPM ini kemudian akan digabungkan dengan parameter terbaik GMM yang sudah diujikan sebelumnya untuk menguji sistem.

4.3.3 Pengujian Sistem

Pengujian Pengujian sistem dilakukan setelah adanya hasil dari pengujian GMM dan HPM. Berikut adalah tabel hasil pengujian sistem.

Tabel 4.9 Hasil pengujian sistem

Jumlah orang	Akurasi (%)	
	Video uji 1	Video uji 2
1	91.4	81.7
2	60.7	23.6
3	41.4	3.23
4	27.8	3.2
5	1.8	0

Seerti disajikan pada tabel 4.9, Dapat dilihat bahwa penggabungan metode GMM dengan HPM menghasilkan keseluruhan nilai akurasi yang lebih baik dibandingkan dengan pengujian-pengujian sebelumnya, dimana nilai akurasi tertinggi dicapai untuk pendeteksian 1 orang yang memiliki akurasi sebesar 91%. Hasil pengujian ini masih tidak terlepas dari false detection yang bisa mengurangi maupun menambah tingkat akurasi.terlihat.

4.3.4 Pengujian Sistem dengan Ukuran Video Uji Diperkecil Setengah Kalinya

Pengujian ini dilakukan untuk menguji ketahanan sistem terhadap ukuran frame yang berbeda. Pengujian dilakukan dengan menggunakan video uji 1 namun dengan ukuran ½ kali frame normal dari video uji 1. Video uji 1 berukuran 592x800 piksel diubah menjadi berukuran setengah kalinya yakni 296x400 piksel. Parameter yang berubah adalah MinimumBlobArea menjadi 1193 dan MinSize menjadi 39 32. Berikut adalah tabel hasil pengujian sistem terhadap video uji 1 dengan ukuran ½ kali frame video uji 1.

Tabel 4.10 Tabel pengujian sistem video uji 1 dengan ukuran ½ kali frame video uji 1

Jumlah orang	Akurasi (%)
1	99.3
2	72.6
3	21.9
4	0
5	0

Dengan menggunakan ukuran ½ kali dari frame normal video uji 1, dibandingkan dengan hasil deteksi untuk ukuran frame normal dari video uji 1, akurasi untuk deteksi satu dan dua orang meningkat dengan deteksi 1 orang mencapai nilai 99%.. Namun lain halnya dengan yang didapatkan untuk deteksi tiga, empat, dan lima orang dimana

deteksi 4 dan 5 orang menjadi yang terburuk dengan akurasi 0%.

4.3.5 Pengujian GMM dengan Perubahan Intensitas Cahaya

Berikut merupakan hasil pengujian GMM terhadap masing-masing perubahan intensitas cahaya yang diberikan pada video uji 1. Parameter yang digunakan merujuk kepada parameter terbaik GMM yang telah diperoleh dari pengujian sebelumnya. Perhitungan dilakukan dengan menghitung jumlah orang sesuai dengan jumlah orang sebenarnya seperti pada pengujian sistem. Berikut adalah hasil pengujian GMM dengan perubahan intensitas cahaya.

Tabel 4.11 Hasil pengujian GMM dengan perubahan intensitas cahaya

Perubahan intensitas cahaya	Akurasi jumlah orang (%)				
	1	2	3	4	5
+50%	72.4	36.3	12.5	0	0
+25%	86.8	32.6	11.9	0	0
0%	91	25.9	12.5	0.7	0
-25%	86.8	32.6	11.9	0	0
-50%	72.4	36.3	12.5	0	0

Dari pengujian yang telah dilakukan, tidak terdapat perbedaan yang terlalu signifikan walaupun terjadi perbedaan intensitas cahaya, yang berarti GMM dalam pengujian cukup tangguh terhadap perubahan intensitas cahaya. Berkurangnya akurasi dengan drastis dikarenakan GMM memang kurang handal dalam menangani objek yang statis, maka akurasi terus berkurang seiring jumlah orang yang seharusnya terdeteksi.

5. KESIMPULAN DAN SARAN

Berdasarkan percobaan yang telah dilakukan, maka bisa ditarik beberapa kesimpulan seperti berikut.

1. Metode Gaussian Mixture Model digunakan pada sistem untuk memisahkan background dengan foreground, dan didapatkan parameter terbaik antara lain NumGaussians adalah 3, MinimumBackgroundRatio bernilai 0.7, dan MinimuBlobArea bernilai 2386. Selain itu GMM cukup tangguh dengan perubahan intensitas cahaya, dengan akurasi yang didapatkan mencapai 86% untuk kasus deteksi satu orang dengan penambahan intensitas cahaya sebesar 25%.
2. Human Presence Map dapat membersihkan noise yang ada pada data uji dan dapat mendeteksi keberadaan kepala orang. Parameter terbaik yang didapatkan adalah MinSize bernilai 49 42 dan MergeThreshold bernilai 15. Akurasi bervariasi untuk jumlah orang berbeda dengan hasil terbaik mencapai 77% untuk kasus deteksi 1 orang.
3. Akurasi sistem secara keseluruhan yang didapatkan mencapai 91% untuk kasus deteksi 1 orang untuk video uji 1 yakni duduk terpisah. Dengan menggunakan ukuran $\frac{1}{2}$ frame dari ukuran normal

video uji 1, akurasi untuk deteksi 1 orang mencapai 99%.

Saran yang dapat diberikan penulis berdasarkan hasil dari penelitian yang diperoleh adalah seperti berikut.

1. Sebaiknya penggunaan data latih yang digunakan lebih mirip dengan studi kasus yakni pada ruang kelas sehingga akurasi sistem dapat lebih baik. Selain itu variasi sudut kepala untuk dalam training bisa lebih banyak agar sistem semakin handal dalam mendeteksi kepala orang dengan berbagai posisi.
2. Perlunya metode pengenalan orang yang lebih handal untuk mengenali orang secara tepat sehingga objek yang terdeteksi benar-benar sesuai dengan jumlah orang yang sebenarnya.

6. REFERENSI

- [1] Celik, H., A. Hanjalic, dan E. A. Hendriks, 2006, *Towards a robust solution to people counting*, Proc. IEEE International Conference Image Processing pp. 2401–2404.
- [2] Choudri, S., James M. Ferryman, dan Atta Badii, 2009, *Robust Background Model for Pixel Based People Counting using a Single Uncalibrated Camera*, Twelfth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS- Winter). Fehr, D., Ravishankar Sivalingam, Vassilios Morellas,
- [3] Nikolaos Panikolopoulos, Osama Lotfallah, dan Youngchoon Park, 2009, *Counting People in Groups*, Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance.
- [4] Hou, YL, G.K.H. Pang, 2011, *People Counting and Human Detection in a Challenging Situation*, IEEE Transactions volume 41 issue 1.
- [5] Hsieh, JW, CH Chuang, SY Chen, CS Peng, dan KC Fan, 2007, *Grid-based Template Matching for People Counting*, IEEE 9th Workshop.
- [6] Hsiung, Yu-Lu, 1999, *Introduction to Digital Video*, University of Texas.
- [7] Kumar, R., Tapesh Parashar, dan Gopal Verma, 2012, *Background Modeling and Subtraction Based People Counting for Real Time Video Surveillance*, International Journal of Soft Computing and Engineering.
- [8] Liang, H., Jinchun Wu, dan Kaiqi Huang, 2012, *People in Seats Counting via Seat Detection for Meeting Surveillance*, Springer Pattern Recognition Communications in Computer and Information Science Vol. 321, pp. 202-210.
- [9] Malawski, F., 2014, *Top-view People Counting in Public Transportation using Kinect*, Young Scientist Towards Challenges of Modern Technology pp. 39-40.
- [10] Raheja, J.L., Sishir Kalita, Pallab Jyoti Dutta, dan Solanki Lovendra, 2012, *An Insight into the Algorithms*

- on Real-Time People Tracking and Counting System, International Journal of Computer Applications, volume 46 – number 5.
- [11] Raheja, J.L., Sishir Kalita, Pallab Jyoti Dutta, dan Solanki Lovendra, 2012, *A Robust Real Time People tracking and Counting incorporating shadow detection and removal*, International Journal of Computer Applications.
- [12] Rostianingsih, S., Rudy Adipranata, dan Fredy Setiawan Wibisono, 2008, *Adaptive Background Dengan Metode Gaussian Mixture Models Untuk Real-Time Tracking*, Jurnal Informatika Vol.9 No.1.
- [13] Stauffer, C., Grimson, W. E. L., 1999, *Adaptive Background Mixture Models for Real-Time Tracking*, IEEE Computer Society Vol. 2, pp. 246-25.
- [14] Valle, J.D., L.E.S. Oliveira, A.L. Koerich, dan A.S. Britto Jr., 2007, *People Counting in Low Density Video Sequences*, Springer on Advances in Image and Video Technology Vol. 1872 pp. 737-748.
- [15] Viola, P, M. Jones, 2001, *Rapid Object Detection using a Boosted Cascad of Simple Features*, International Journal of Computer Vision and Pattern Recognition.
- [16] Wang, JM, S. W. Chen, dan C. S. Fuh, 2009, *People Counting System Based on Particle Filter with Memory States for Improvement*, Proc. of the 22th IPPR Conf. on CVGIP.
- [17] Y. Takayanagi and J. Katto, *Human Body Detection using HOG with Additional Color Features*, International Workshop on Advanced Image Technology, 2010.