

Evaluasi Metode Ekstraksi Fitur *Hu Moment Invariants* untuk Pengenalan Aktivitas Manusia

Hans Christian Kurniawan^{#1}, Kevin Suryajaya Soemarto^{#2}, Bernardo Nugroho Yahya^{*3}

[#]*Departemen Informatika, Institut Teknologi Harapan Bangsa
Jl. Dipati Ukur no. 80-84, Bandung, Indonesia*

¹*hans_christian@ithb.ac.id*

²*kevinsuryajaya08@gmail.com*

^{*}*Department of Industrial and Management Engineering, Hankuk University of Foreign Studies
Global Campus, 81, Oedae-ro, Mohyeon-eup, Cheoin-gu, Yongin, South Korea*

³*bernardo@hufs.ac.kr*

Abstract— *Vision-based Human Activity Recognition has been widely used due to a bunch of video data availability in the present days through CCTV and another mechanism which contains some human activities. This data is going to be very useful to improve and automate decision-making in several fields including security surveillance. In this field, it is important to achieve a good performance (i.e., accuracy) inefficient computational time. While there are many approaches in this field, most complex approaches require high computational time. In this work, we are evaluating Hu Moments performance, as well as being compared to other methods (i.e., Zernike Moment and Histogram of Oriented Gradient) by its accuracy and computational time. We also improved HAR flow by adding image denoising which has proven effective in increasing accuracy. The testing process includes videos that contain human activities such as walking, jogging, and running. The result shows that Hu Moments is superior among other methods, however there's also some room for improvements found through this experiment.*

Keywords— *Human Activity Recognition (HAR), Hu Moment Invariants, Zernike Moment, HOG, Gait Pattern, SVM.*

Abstrak— Dalam era di mana terdapat banyak data video yang berisi aktivitas manusia, baik melalui rekaman CCTV maupun mekanisme lain, data tersebut menjadi sangat berharga untuk dapat diproses untuk pengenalan aktivitas manusia, atau *Human Activity Recognition (HAR)* yang dapat membantu pengambilan keputusan, di antaranya *security surveillance*. Untuk itu, diperlukan akurasi yang tinggi dan waktu komputasi yang efisien. Meskipun telah banyak metode di ranah ini, suatu teknik yang kompleks pada umumnya membutuhkan waktu komputasi yang tinggi. Dalam penelitian ini, dilakukan evaluasi dengan menggunakan metode *Hu Moments* yang akan dibandingkan dengan metode lainnya, yaitu *Zernike Moment* dan *Histogram of Oriented Gradient (HOG)*, untuk segi akurasi dan waktu komputasinya. Ditambahkan juga tahap *image denoising* yang mampu meningkatkan akurasi. Proses pengujian menggunakan berbagai data video aktivitas manusia yang meliputi: berjalan, jogging, dan berlari. Hasil riset menunjukkan bahwa metode *Hu Moments* memiliki performa yang lebih unggul dibandingkan metode ekstraksi fitur lainnya. Berdasarkan eksperimen yang dilakukan, terdapat beberapa area yang masih dapat ditingkatkan, untuk penelitian selanjutnya.

Kata Kunci— *Pengenalan Aktivitas Manusia, Hu Moment Invariants, Zernike Moment, HOG, Gait Pattern, SVM.*

I. PENDAHULUAN

Saat ini video sudah menjadi sebuah media rekaman yang umum. Ketersediaan video dalam jumlah besar, seperti melalui rekaman CCTV, membuka peluang yang luas dalam bidang pengolahan data agar dapat dimanfaatkan untuk mendukung pengambilan keputusan secara cepat. Data video yang berisi aktivitas manusia, seperti berjalan, berlari, jogging, duduk, merangkak, dan lain sebagainya, dapat direkognisi dan digunakan untuk berbagai keperluan. Contohnya antara lain: *security surveillance* yang dapat mendeteksi aktivitas manusia yang mengendap-ngendap secara mencurigakan sehingga dapat membunyikan alarm keamanan dan pengawasan aktivitas lansia di panti jompo yang dapat mendeteksi bila terdapat lansia yang terjatuh dan membutuhkan pertolongan, maka sistem dapat memberikan notifikasi kepada perawat.

Proses pengolahan data berbasis video ini dikenal dengan sebutan *vision-based Human Activity Recognition (HAR)* yakni sebuah metode untuk mengenali beragam aktivitas manusia dari sebuah video yang mengandung objek manusia [1]. Dalam sebuah sistem HAR, akurasi atau ketepatan pengenalan aktivitas manusia serta waktu komputasi yang cepat adalah hal yang sangat krusial, mengingat dalam aplikasinya seharusnya HAR dapat dilakukan dengan cepat, bahkan *real-time*.

Terdapat beberapa metode HAR yang umum digunakan, seperti *posture feature extraction* yang menggunakan fitur *skeleton* dalam HAR yang cukup efektif dan memiliki akurasi hingga 98,2%. Namun, algoritme tersebut tidak bisa membedakan aktivitas yang memiliki bentuk *skeleton* yang serupa dengan fitur yang diekstraksi [2]. Selain itu, terdapat juga metode *Hidden Markov Model (HMM)* yang memiliki akurasi cukup baik, hingga 91,3%, apabila citra masukannya memiliki kualitas tinggi, tetapi waktu komputasinya relatif lambat [3]. Terdapat juga metode lainnya, yaitu *Zernike Moments (ZMI)* dan *Histogram of Oriented Gradients (HOG)*. ZMI dianggap dapat memberikan hasil yang lebih baik, hingga 95% [4], sedangkan HOG hingga 99,7% [5]. Namun, proses komputasi keduanya cukup kompleks. Perlu dievaluasi cocok atau tidaknya bila digunakan dalam sistem yang

membutuhkan waktu komputasi yang cepat. Selain metode-metode tersebut, terdapat juga satu metode lain yang banyak digunakan, yakni *Hu Moment Invariants*. Metode ini digunakan untuk ekstraksi fitur dan dapat digabungkan dengan metode lain, seperti SVM untuk proses klasifikasinya. *Hu Moments* merupakan metode ekstraksi fitur yang populer karena diklaim tahan terhadap manipulasi citra, seperti translasi, skala, rotasi, dan *skew*, serta waktu komputasi yang relatif cepat dengan akurasi yang baik. Fitur yang paling umum digunakan dalam pengenalan aktivitas adalah fitur rupa atau bentuk, fitur gerak, dan gabungan keduanya [6].

Dalam penelitian ini dilakukan beberapa percobaan untuk menguji akurasi, waktu komputasi, dan ketahanan dari metode *Hu Moments*. Juga dilakukan perbandingan dengan metode lain, seperti ZMI dan HOG, agar dapat diketahui suatu metode yang cocok untuk digunakan dalam sistem HAR.

Berdasarkan berbagai evaluasi yang telah dilakukan dalam penelitian ini diketahui bahwa metode *Hu Moments* lebih unggul dari sisi akurasi dan waktu komputasi, serta memiliki ketahanan yang baik terhadap manipulasi data masukan video yang digunakan. Selanjutnya, dilakukan juga pengembangan metode dengan menambahkan proses *Image Denoising Salt and Pepper* yang dapat meningkatkan akurasi model dan analisis kesalahan terhadap beberapa data masukan. Hal ini diharapkan dapat dieksplorasi lebih lanjut pada penelitian selanjutnya.

II. METODOLOGI

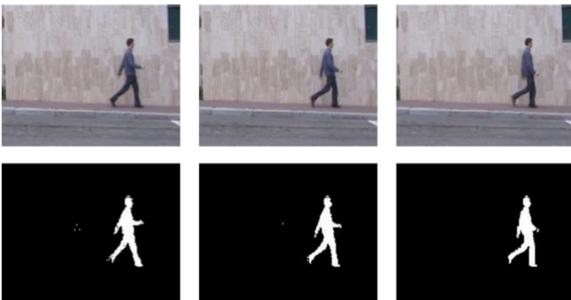
A. Model Human Activity Recognition

1) Gaussian Mixture Model (GMM)

GMM merupakan model probabilistik yang mengasumsikan semua titik data dihasilkan dari campuran sejumlah distribusi Gaussian yang terbatas dengan parameter yang tidak diketahui [7]. GMM digunakan untuk melakukan proses *background subtraction*, atau pemisahan objek, dengan latarnya. Citra yang dihasilkan adalah citra biner dengan nilai 1 untuk objek *foreground* dan 0 untuk *background* [1]. Gambar 1 merupakan contoh citra hasil proses GMM yang dilakukan dalam penelitian ini.

2) Image Denoising

Berdasarkan hasil pengujian dengan berbagai video masukan dalam penelitian ini, kumpulan citra yang dihasilkan



Gambar 1 Contoh *background subtraction* pada video frames dengan GMM, citra atas adalah citra asli dan citra bawah setelah dilakukan GMM

dari proses *background subtraction* pada umumnya akan memiliki *salt and pepper noise* dengan intensitas yang berbeda-beda. Untuk itu, diperlukan teknik *image denoising* untuk membersihkan *noise*. Dalam penelitian ini ditambahkan algoritme *denoising* menggunakan *Median Filter* yang terbukti mampu membersihkan citra dari *noise* tersebut [8]. Gambar 2 merupakan contoh citra hasil *image denoising* yang dilakukan dalam penelitian ini.

3) Hu Moment Invariants

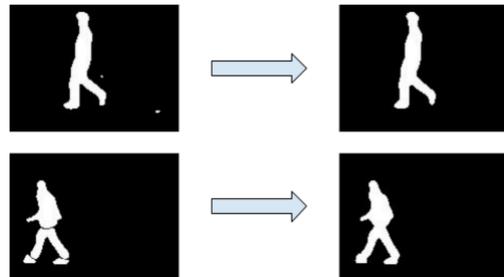
Hu Moment Invariants merupakan metode ekstraksi fitur yang digunakan untuk mendefinisikan 7 buah *moment* (fitur) yang merepresentasikan suatu objek. Pada sistem pengenalan aktivitas manusia diperlukan 4 *moment* pertama yang memiliki daya tahan terhadap skala, translasi, dan rotasi. *Moment* ke-5 sampai ke-7 merupakan deviasi relatif dari *moment* ke-2 dan ke-3 sehingga dapat dihilangkan tanpa mempengaruhi kemampuannya dalam mendeteksi bentuk objek [6] [9].

4) Minimum Enclosing Rectangle

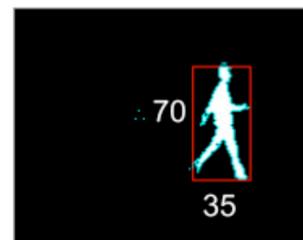
Minimum Enclosing Rectangle merupakan *bounding box*, atau sebuah bentuk persegi panjang terkecil yang mengelilingi wilayah objek target. Proses ini melibatkan beberapa *frame* citra masukan yang kemudian diekstraksi sehingga diperoleh *ratio* tinggi dan lebar dari *bounding box* yang menandai area target mengikuti rumus [1]. Contoh citra hasil dari proses *Minimum Enclosing Rectangle* dapat dilihat pada Gambar 3.

5) Posture Change Velocity

Posture Change Velocity adalah metode untuk menghitung perubahan bentuk postur sebuah objek dari *frame* ke-*i* dan *frame* ke-*(i-1)* dengan membagi rasio yang sudah dihitung dalam persamaan *Minimum Enclosing Ratio* [1].



Gambar 2 Hasil dari proses *Image Denoising* (citra kiri adalah citra asli, citra kanan adalah setelah dilakukan *denoising*)



Gambar 3 Hasil proses *Minimum Enclosing Rectangle*

6) *Motion Rate*

Persamaan *Motion Rate* digunakan untuk menghitung kecepatan perpindahan koordinat. Jika koordinat *centroid*, atau pusat dari gambar, ke-*i* adalah (x_i, y_i) , perpindahan objek antara kedua *frame* yang berdekatan dapat dihitung oleh perpindahan massa pusat [1]. Gambar 4 menunjukkan contoh citra hasil dari proses perhitungan *Motion Rate*.

7) *Behavior Characteristic Representation*

Behavior Characteristic menunjukkan perilaku dari seseorang seperti tanda tangan, cara berjalan, dll. Pada umumnya sebuah perilaku manusia dapat langsung dilabeli sebagai suatu aktivitas, seperti berjalan, jogging, atau berlari hanya dalam waktu 0,4 – 0,8 detik [10]. Dalam penelitian ini digunakan 7 buah fitur untuk merepresentasikan aktivitas manusia, yakni 4 buah fitur *Hu Moments*, *Minimum Enclosing Rectangle*, *Posture Velocity Change*, dan *Motion Rate* [1].

8) *Support Vector Machine (SVM)*

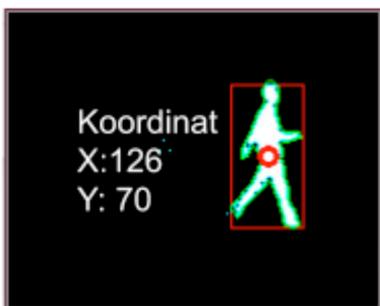
SVM merupakan salah satu algoritme klasifikasi dengan mencari *hyperplane* terbaik yang berfungsi sebagai pemisah dari dua atau lebih kelas. SVM mengidentifikasi beberapa aturan pada data yang mencirikan sebuah properti yang kemudian dapat digunakan untuk menentukan apakah suatu data baru memiliki properti tersebut [10]. Metode SVM memiliki akurasi yang baik dan performansi yang cepat sehingga banyak digunakan dalam pengembangan HAR [1].

9) *Zernike Moment Invariants (ZMI)*

Zernike Moment Invariants merupakan metode alternatif dari *Hu Moments* yang digunakan untuk menghindari permasalahan matematis dengan menggunakan geometri dan *complex moment*. Zernike menghasilkan *3D moment invariant* dari sebuah momen ortogonal. Metode ini memiliki sifat *invariants* sehingga banyak digunakan sebagai deskriptor fitur, sama seperti *Hu Moments*. Informasi untuk ZMI tidak didapatkan dengan mengambil informasi dari sebuah titik tertentu melainkan menggunakan informasi dari citra utuh. Hal ini dilakukan agar sistem dapat mengambil properti global, seperti orientasi gambar yang tidak didapatkan secara lokal [4].

10) *Histogram of Oriented Gradients (HOG)*

Histogram of Oriented Gradient juga merupakan metode alternatif *Hu Moments* untuk mengekstraksi fitur dari sebuah citra dengan mengambil histogram orientasi intensitas tepi di



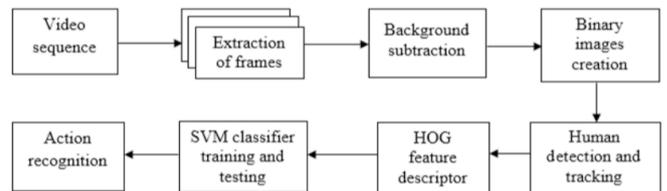
Gambar 4 Contoh hasil penentuan *centroid* (ditandai lingkaran merah)

wilayah lokal. Proses untuk mengenali aktivitas manusia dengan HOG dimulai dengan memecah video menjadi banyak *frame* yang kemudian dilakukan *background subtraction*, sehingga menghasilkan citra biner. Proses berikutnya dilakukan dengan ekstraksi fitur dan klasifikasi SVM [5]. Gambar 5 menunjukkan proses HAR menggunakan metode HOG.

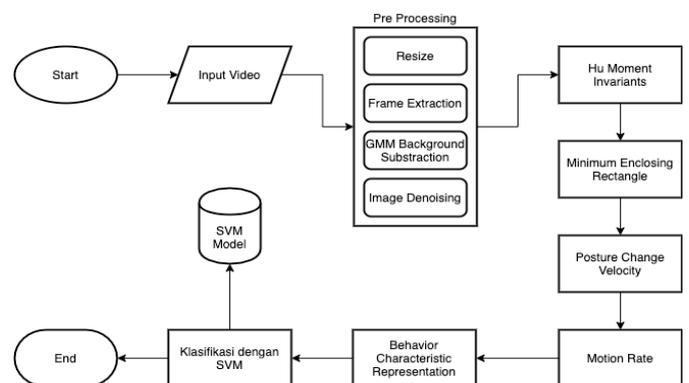
B. Pengembangan Sistem

Berdasarkan landasan teori yang telah dibahas sebelumnya, dikembangkan sebuah Sistem Pengenalan Aktivitas manusia dengan uraian *flowchart* sesuai Gambar 6.

- 1) Pada tahap pertama dimasukkan *video* yang menampilkan salah satu aktivitas manusia, yakni berjalan, jogging, atau berlari. Karakteristik video masukan tersebut dijelaskan pada bagian *dataset*.
- 2) Pada tahap *preprocessing*, dilakukan beberapa proses awal terhadap video masukan yakni :
 - i) Penyeragaman ukuran video (*resizing*).
 - ii) Ekstraksi *frames* dari video masukan. Dalam penelitian ini untuk setiap video masukan yang berdurasi 1-3 detik dengan 25 fps (*frame per second*) cukup diambil 15 *frames* saja, atau setara dengan 0,6 detik, untuk diproses. Hal ini berdasarkan teori *Behavior Characteristic Representation*, di mana waktu tersebut dinilai sudah cukup untuk mengenali aktivitas manusia dalam sebuah video. Kelima belas *frames* yang dipilih adalah 15 *frames* terakhir. Hal ini disebabkan pada beberapa *frames* awal seringkali berupa *frame* kosong, atau belum ada objek yang terekam di video, sehingga dapat mengurangi akurasi. Pengaturan ini mungkin saja berbeda jika kebutuhan dan data yang digunakan



Gambar 5 Diagram blok HAR menggunakan HOG [5]



Gambar 6 *Flowchart* Sistem Pengenalan Aktivitas

memiliki karakteristik yang berbeda pula dengan yang digunakan dalam penelitian ini.

- iii) *Background subtraction* dari video masukan menggunakan GMM untuk memisahkan *foreground objects* dengan *background* sehingga diperoleh keluaran berupa sekumpulan citra biner.
- iv) *Image denoising* untuk menghilangkan *salt and pepper noise* pada video menggunakan *Median Filter* sehingga diperoleh citra yang bersih. Tahapan ini khusus ditambahkan dalam penelitian ini setelah melihat hasil *background subtraction* yang tidak bersih dan mengurangi akurasi modelnya.

- 3) Ekstraksi fitur menggunakan *Hu Moment Invariants* untuk mendapatkan 4 dari 7 *moments* sebagai fitur pemodelan. Fitur ke-5 hingga ke-7 dapat diabaikan karena tidak berpengaruh dalam penentuan aktivitas manusia.
- 4) Mencari area objek manusia dengan *Minimum Enclosing Rectangle* sehingga diperoleh *bounding box* dengan ukuran sesuai objek manusia tersebut.
- 5) Menghitung *Posture Change Velocity* untuk mengukur rasio perpindahan postur objek antar *frame*.
- 6) Menghitung kecepatan perpindahan (*Motion Rate*) objek manusia yang dihitung dari titik pusat (*centroid*).
- 7) Sampai tahap ini telah diperoleh 7 fitur (4 fitur, yaitu *Hu Moments*, *Minimum Enclosing Rectangle*, *Posture Velocity Change*, dan *Motion Rate*) dan masing-masing 15 *frames* sehingga total terdapat 15 buah vektor fitur.
- 8) *Support Vector Machine* (SVM) digunakan untuk melakukan klasifikasi dengan model yang sudah didapat melalui proses *training*. Pada tahap ini masing-masing vektor fitur untuk setiap *frame* diklasifikasikan secara independen. Mayoritas kelas yang terklasifikasi akan menjadi kesimpulan akhir. Sebagai contoh, jika hasil klasifikasi adalah 9 *frame* diidentifikasi sebagai jalan, 2 *frame* jogging, dan 4 *frame* lari, maka kesimpulan akhir aktivitas dalam video ini adalah berjalan.

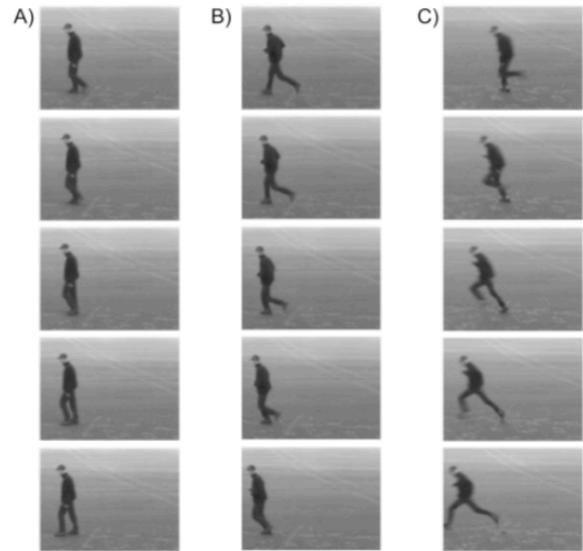
C. Dataset

Penelitian ini menggunakan *dataset* Weizmann, KTH, dan variasi video rekaman sendiri. Video rata-rata berdurasi kurang dari 3 detik dengan 25 fps. Total *dataset* Weizmann dan KTH yang digunakan adalah sebanyak 233 buah yang terdiri dari aktivitas berjalan, jogging, dan berlari yang dilakukan oleh objek manusia berjumlah 34 orang. Pada tahap *training* dipilih secara manual citra-citra yang baik yang tidak mengandung banyak *noise*, tidak tertutupi oleh objek lain, dan terlihat jelas bentuk tubuh dari kepala hingga ujung kaki. Gambar 7 merupakan contoh *dataset* Weizmann untuk ketiga aktivitas manusia.

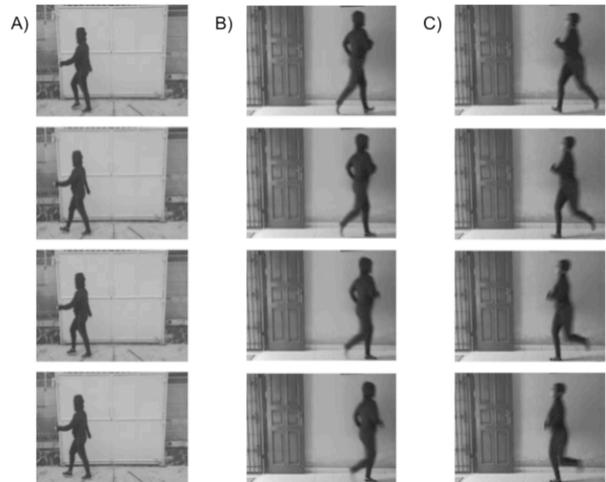
Dilakukan juga perekaman data sendiri yang terdiri atas 27 video dengan 2 objek manusia. Contoh *frames* citra rekaman sendiri dapat dilihat pada Gambar 8. Perekaman ini ditujukan untuk proses pengujian beberapa variasi masukan yang tidak tersedia di *dataset* sebelumnya, seperti perubahan orientasi kamera di tengah perekaman. Komposisi citra *training* berbanding *testing* adalah 146 video (*dataset* Weizmann) dan berbanding 87 video (*dataset* KTH dan rekaman sendiri) yang seluruhnya berbeda.

III. HASIL DAN PEMBAHASAN

Dalam penelitian ini, dilakukan beberapa skenario pengujian untuk mengukur akurasi model, sistem, kondisi spesifik, dan perbandingannya dengan metode lain.



Gambar 7 Contoh *dataset* Weizmann (A) Jalan, (B) Joging, dan (C) Lari



Gambar 8 Contoh *dataset* rekaman sendiri (A) Jalan, (B) Joging, dan (C) Lari

A. Pengujian Akurasi Model dengan *Hu Moment Invariants*

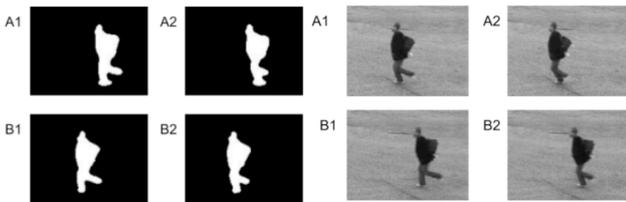
Pada tahap pengujian akurasi model klasifikasi, data *frames* yang digunakan untuk pengujian berjumlah 675 *frames*, terdiri dari kumpulan 15 *frames* untuk masing-masing 45 video *dataset* KTH (15 jalan, 15 jogging, dan 15 lari). Berdasarkan hasil pengujian pada Tabel I dapat dilihat rata-rata akurasi model SVM menggunakan *Hu Moments* dalam mendeteksi aktivitas manusia sebesar 93,19%.

Setelah dilakukan analisis terhadap data yang salah klasifikasi, ditemukan beberapa permasalahan pada *dataset*. Permasalahan pertama diakibatkan oleh *Gait Pattern*, atau cara bergerak manusia, yang memiliki karakteristiknya masing-masing [11]. Pada Gambar 9 gaya jogging dan berlari dari objek tersebut sangat mirip sehingga sulit untuk dibedakan. Hal ini mengakibatkan kesalahan pengenalan. Berbeda dengan Gambar 10, terlihat perbedaan yang jelas pada postur dan cara bergerak objek.

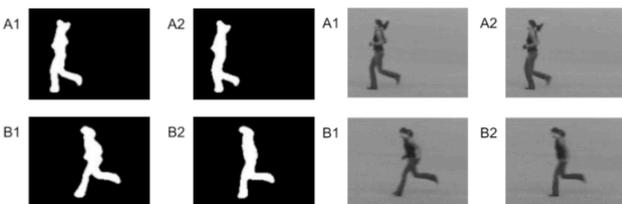
Selain itu, terdapat juga kesalahan klasifikasi yang diakibatkan oleh warna pakaian yang mirip atau sama dengan warna latar, seperti ditunjukkan pada Gambar 11 sehingga timbul *noise*. Ketika melakukan *background subtraction* dan objek *foreground*, menjadi tidak utuh.

TABEL I
HASIL PENGUJIAN AKURASI MODEL DENGAN *HU MOMENT INVARIANTS*

Aktual	Predicted (frame)			Akurasi
	Jalan	Joging	Lari	
Jalan	200	19	6	93,33%
Joging	19	192	14	90,81%
Lari	1	10	214	95,41%
Rata-Rata Akurasi Model				93,19%



Gambar 9 Contoh data dengan *Gait Pattern* yang kurang baik. (A) adalah aktivitas jogging dan (B) adalah aktivitas berlari



Gambar 10 Contoh data dengan *Gait Pattern* yang baik. (A) Aktivitas jogging dan (B) Aktivitas berlari

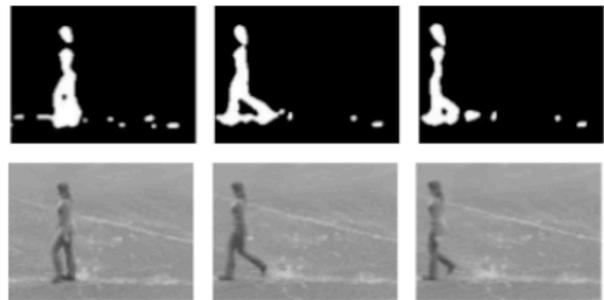
Terakhir, meskipun telah dilakukan tahapan *image denoising*, beberapa data masih tetap terdapat *noise* yang mengakibatkan akurasi berkurang, seperti terlihat pada Gambar 12. Perlu dilakukan analisis lebih lanjut untuk mengkombinasikan beberapa metode *image denoising* sesuai dengan karakteristik data yang digunakan.

B. Pengujian Pengaruh *Image Denoising* Terhadap Akurasi Model

Dalam bagian ini diuji akurasi model dengan data yang sama seperti digunakan pada Pengujian A. Pengujian dilakukan tanpa menerapkan proses *image denoising* karena proses ini merupakan pengembangan dalam penelitian ini. Hasilnya dapat dilihat pada Tabel II. Rata-rata akurasi model yang diperoleh adalah 91,60%. Ini menunjukkan bahwa penambahan tahapan ini berpengaruh dalam meningkatkan akurasi model hingga 1,59% dibandingkan tanpa *image denoising*.



Gambar 11 Contoh data dengan warna pakaian yang mirip dengan latar



Gambar 12 Contoh data yang masih memiliki *noise*

TABEL II
HASIL PENGUJIAN AKURASI MODEL TANPA *IMAGE DENOISING*

Aktual	Prediksi (video)			Akurasi
	Jalan	Joging	Lari	
Jalan	193	21	11	92,15%
Joging	20	185	20	89,19%
Lari	1	12	212	93,48%
Rata-Rata Akurasi Model				91,60%

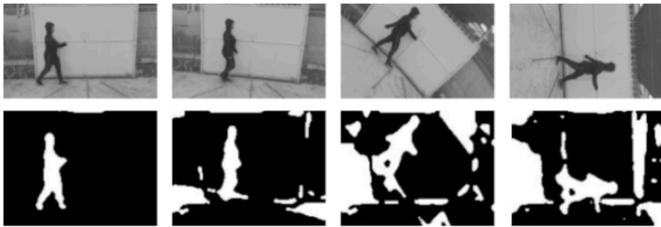
C. Pengujian Akurasi Sistem dengan *Hu Moment Invariants*

Pada tahap pengujian akurasi sistem dianalisis masing-masing video masukan berdasarkan mayoritas hasil klasifikasi setiap *frame*. Diperoleh akurasi model untuk setiap *frame* sebesar 93,19% (Tabel I), namun secara keseluruhan aktivitas dalam video *input* 100% dapat dikenali dengan akurat (Tabel III).

D. Pengujian Akurasi Model dengan Perubahan Orientasi Kamera

Dalam pengujian ini digunakan 9 buah video rekaman sendiri (135 *frames*). Dilakukan perubahan orientasi kamera saat rekaman sedang berjalan, seperti dapat dilihat pada Gambar 13. Proses ini mengakibatkan banyak *noise* yang muncul sehingga mengakibatkan penurunan akurasi model.

Berdasarkan hasil pada Tabel IV, diperoleh rata-rata akurasi 48,15%, sehingga dapat diketahui bahwa metode *Hu Moments* belum mampu untuk menangani perubahan orientasi kamera secara mendadak.



Gambar 13 Contoh *frames* pada perubahan orientasi kamera

TABEL III
HASIL PENGUJIAN AKURASI SISTEM DENGAN *HU MOMENT INVARIANTS*

Aktual	Prediksi (video)			Akurasi
	Jalan	Joging	Lari	
Jalan	15	0	0	100%
Joging	0	15	0	100%
Lari	0	0	15	100%
Rata-Rata Akurasi Sistem				100%

TABEL IV
HASIL PENGUJIAN AKURASI MODEL DENGAN PERUBAHAN ORIENTASI

Aktual	Prediksi (<i>frame</i>)			Akurasi
	Jalan	Joging	Lari	
Jalan	28	3	29	48,89%
Joging	37	2	36	43,70%
Lari	0	0	0	51,85%
Rata-Rata Akurasi Model				48,15%

E. Pengujian Akurasi Model dengan Pergerakan Diagonal

Dalam pengujian ini digunakan *dataset* KTH, di mana objek manusia bergerak secara diagonal dari sisi bawah kiri ke atas kanan. Contoh *dataset* yang digunakan dapat dilihat pada Gambar 14.

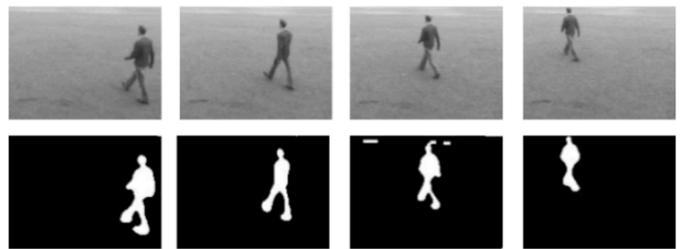
Berdasarkan hasil yang diperlihatkan pada Tabel V, diperoleh rata-rata akurasi 62,96%. Analisis lebih lanjut menunjukkan bahwa masalah utama pada *dataset* ini adalah kecepatan perpindahan (*Motion Rate*) dari objek. Pergerakan diagonal mengakibatkan jarak perpindahan lebih kecil dan kecepatan menjadi lebih rendah dibandingkan pergerakan horizontal sehingga tidak cocok dengan model yang sudah dibuat. Namun, perbedaan kecepatan antar aktivitas masih tetap ada. Apabila model dilatih menggunakan *dataset* diagonal saja, maka masih dimungkinkan untuk mengenali *dataset* ini dengan cukup baik.

F. Pengujian Ketahanan Metode *Hu Moment Invariants* terhadap Rotasi dan Skala

Melalui pengujian ini, ingin diketahui ketahanan model apabila *dataset* diubah rotasinya sebesar 90°, 180°, dan 270°. Selain itu, pada *dataset* juga dilakukan perubahan skala sebesar x0.5 dan x1.5. Hasil yang diperoleh sama dengan Tabel I sehingga terbukti bahwa *Hu Moments* tahan terhadap perubahan rotasi dan skala.

G. Pengujian Perbandingan Akurasi Model dan Waktu Komputasi dengan Berbagai Metode

Pada pengujian ini dilakukan perbandingan hasil akurasi model dan kecepatan komputasinya di antara tiga metode yang umum digunakan dalam HAR, yakni *Hu Moments*, Zernike Moments, dan HOG. Pengujian menggunakan *dataset* yang sama seperti pada Pengujian A. Hasilnya dapat dilihat



Gambar 14 Contoh *dataset* dengan pergerakan diagonal

TABEL V
HASIL PENGUJIAN AKURASI MODEL DENGAN PERGERAKAN DIAGONAL

Aktual	Prediksi (<i>frame</i>)			Akurasi
	Jalan	Joging	Lari	
Jalan	74	1	0	65,78%
Joging	55	20	0	53,78%
Lari	21	48	6	69,33%
Rata-Rata Akurasi Model				62,96%

pada Gambar 15, di mana akurasi *Hu Moments* lebih unggul dengan persentase 93,19% (Tabel I). Akurasi *Zernike Moments* sebesar 93,09%, sedangkan HOG terpaut jauh di 89,14%.

Jika dilihat dari waktu komputasinya, metode *Hu Moments* dapat melakukan pengenalan aktivitas manusia dengan waktu paling cepat, yakni 1,232 detik. Waktu komputasi *Zernike Moments* adalah 1,602 detik dan HOG 2,031 detik, seperti yang dapat dilihat pada Gambar 16.

Melalui pengujian ini dapat diketahui bahwa metode *Hu Moments* lebih unggul, baik dalam akurasi maupun waktu komputasi, sehingga lebih cocok untuk digunakan dalam sebuah sistem HAR dibandingkan kedua metode lainnya.

IV. SIMPULAN

Berdasarkan hasil pengujian yang dilakukan dalam penelitian ini, dihasilkan sistem pengenalan aktivitas manusia dengan menggunakan metode *Hu Moment Invariants* yang memiliki akurasi model hingga 93,19% dan rata-rata komputasi 1,23 detik, serta akurasi sistem dalam penarikan kesimpulan untuk setiap video hingga 100%. Apabila dibandingkan dengan model menggunakan *Zernike Moments*, akurasinya 93,09% dan rata-rata komputasi 1,60 detik, dan model *Histogram of Oriented Gradient*, akurasinya 89,14% dan rata-rata komputasi 2,0 detik. Oleh karena itu, dapat disimpulkan bahwa metode *Hu Moment Invariants* jauh lebih unggul dan cocok untuk digunakan dalam pengembangan sistem deteksi aktivitas manusia.

Penambahan metode *image denoising* juga terbukti dapat meningkatkan akurasi hingga 1,59%. Masih sangat terbuka untuk dikombinasikan dengan metode *denoising* lainnya agar akurasi dapat semakin meningkat. Selain itu, penelitian ini juga membuktikan bahwa metode *Hu Moment Invariants* tahan terhadap perbedaan rotasi dan skala dari input video. Namun, belum mampu menanganinya dengan baik apabila orientasi kamera diubah di tengah perekaman. Metode ini juga belum mampu menangani data campuran antara pergerakan objek secara horizontal dan diagonal. Hal ini dikarenakan adanya perbedaan hasil pengukuran kecepatan dan perpindahan objek yang mengakibatkan kesalahan klasifikasi.

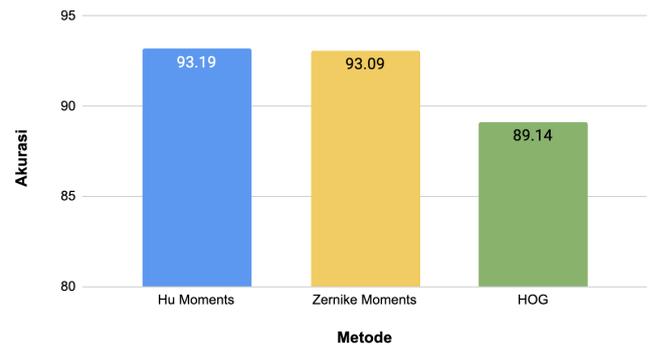
Masih terdapat beberapa keterbatasan dalam sistem pengenalan aktivitas manusia yang dikembangkan dalam penelitian ini, seperti penanganan variasi *gait pattern*, variasi *noise*, dan warna pakaian yang mirip dengan warna latar. Diperlukan penelitian lebih lanjut untuk mencari cara agar dapat menangani *input* data dengan pergerakan objek secara diagonal dan perubahan orientasi kamera saat perekaman.

DAFTAR REFERENSI

[1] S. Jiabin, C. Zhenxin, Z. Zian, dan H. Yongmei, "Research on Recognition of Indoor Fall Behaviors Based on Video Monitoring," dalam *2017 Chinese Automation Congress (CAC)*, Jinan, China, 2017.

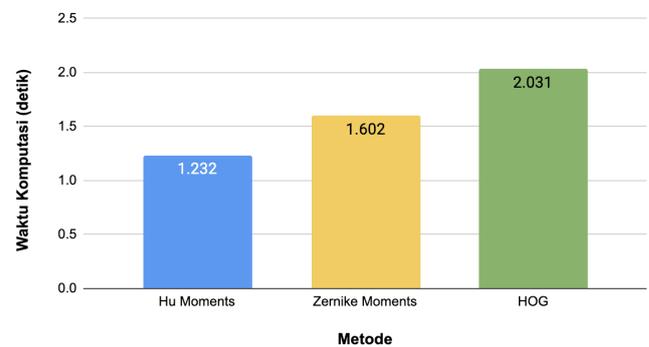
[2] E. Cipitelli, S. Gasparrini, E. Gambi, dan S. Spinsante, "A Human Activity Recognition System using Skeleton Data from RGBD Sensors," *Computational Intelligence and Neuroscience*, vol. 2016, 2016.

Perbandingan Akurasi Model



Gambar 15 Perbandingan hasil uji akurasi model dengan metode *Hu Moments*, *Zernike Moments*, dan HOG

Perbandingan Waktu Komputasi



Gambar 16 Perbandingan hasil uji waktu komputasi

[3] S. Kamal, A. Jalal, dan D. Kim, "Depth Images-based Human Detection, Tracking and Activity Recognition Using Spatiotemporal Features and Modified HMM," *Journal of Electrical Engineering and Technology*, 2016.

[4] C. Dhiman dan D.K. Vishwakarma, "A Robust Framework for Abnormal Human Action Recognition using R-Transform and Zernike Moments in Depth Videos," *IEEE Sensors Journal*, vol. 19, no. 13, 2019.

[5] M.P. Chandrashekar, B. Jagadeesh, dan M.N. Meghana, "An Approach of Understanding Human Activity Recognition and Detection for Video Surveillance using HOG Descriptor and SVM Classifier," dalam *Proceedings of 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (ICCTCEEC-2017)*, India, 2017, hlm. 481-485.

[6] H. Ming-Kuei, "Visual pattern recognition by moment invariants," *Information Theory, IRE Transactions*, vol. 8, 1962.

[7] C.M. Bishop, *Pattern Recognition and Machine Learning*, edisi ke-1, Berlin: Springer Science + Business Media, 2006.

[8] A. Jayaswal, M. Dembrani, dan S. Seragi, *Image Denoising Using Adaptive Filter*, edisi ke-1, Saarbrücken: LAP, 2017.

[9] J. Flusser, T. Suk, dan B. Zitov, *Moments and Moment Invariants in Pattern Recognition*, edisi ke-1, London: Wiley, 2009.

[10] N. Christianini dan E. Ricci, "Support Vector Machines," dalam *Encyclopedia of Algorithms*. M. Y. Kao Ed., New York: Springer, 2016, hlm. 2170-2174.

[11] J.P. Gupta, P. Dixit, N. Singh, dan V.B. Aemwal, "Analysis of Gait Pattern to Recognize the Human Activities", *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 2, hlm. 7-16, 2014.

Hans Christian Kurniawan, menerima gelar Sarjana Teknik Informatika dari Institut Teknologi Harapan Bangsa (ITHB) Bandung pada tahun 2016 dan Magister Informatika dari Institut Teknologi Bandung pada tahun 2019. Saat ini aktif sebagai Pengajar di ITHB serta sebagai *Engineering Lead* di Bukalapak.

Kevin Suryajaya Soemarto, menerima gelar Sarjana Komputer dari Jurusan Informatika, Institut Teknologi Harapan Bangsa (ITHB) Bandung pada tahun 2020 dan saat ini bekerja sebagai IT Engineer.

Bernardo Nugroho Yahya, menerima gelar *Master of Information System Engineering* dari Dongseo University, Korea Selatan pada tahun 2004 dan Ph.D. di jurusan *Industrial Engineering* di Pusan National University, Korea Selatan tahun 2012. Saat ini aktif sebagai Professor di Departemen Teknik dan Manajemen Industri, Hankuk University of Foreign Studies, Korea Selatan.