

Person Reidentification pada Sistem Surveillance Cerdas Menggunakan Metode Bag of Visual Feature

(Person Reidentification on Intelligent Surveillance System Using Bag of Visual Feature Methods)

Muhammad Yusuf Manshur¹, Wahyono^{2#}

Abstract—Camera-based surveillance systems have been widely used to monitor public places for ensuring security and safety. One of the problems in the surveillance process is identifying human objects on different CCTV cameras, which is referred to as person reidentification (Re-ID). Re-ID is the process of identifying whether the images of human objects, captured from two or more images from CCTV cameras with the different viewpoints, are the same person or not. This paper proposes a method based on visual features of the object image, named as Bag of Visual Feature (BOVF). BOVF works by representing image data as a collection of local features that are used with a feature clustering mechanism. BOVF implementation uses the Density Based Spatial Clustering of Applications with Noise (DBSCAN) clustering methods in the Histogram of Oriented Gradient (HOG) features. The results of this study with 70 image frames from iLIDS-VID dataset obtained the best accuracy at R-20 by 88% using DBSCAN with a processing speed of 1.85 seconds.

Intisari—Sistem pengawasan berbasis CCTV telah banyak digunakan untuk memantau tempat umum, sehingga keamanan dan kenyamanan lebih terjamin. Salah satu permasalahan pada proses pengawasan adalah mengidentifikasi objek manusia pada kamera CCTV yang berbeda, yang disebut sebagai *person reidentification* (Re-ID). Re-ID adalah proses untuk mengidentifikasi gambar objek manusia, yang berasal dari dua citra atau lebih yang berada di sudut pandang kamera CCTV berbeda, adalah orang yang sama atau tidak. Pada makalah ini diusulkan metode berbasis fitur visual dari citra objek dengan nama *Bag of Visual Feature* (BOVF). BOVF bekerja dengan merepresentasikan data sebagai kumpulan fitur lokal yang digunakan dengan mekanisme pengelompokan fitur (*feature clustering*). Implementasi BOVF menggunakan algoritme pengelompokan fitur *Density Based Spatial Clustering of Applications with Noise* (DBSCAN) pada proses pelatihan model citra dengan mempelajari fitur *Histogram of Oriented Gradient* (HOG). Hasil dari penelitian yang menggunakan *dataset* sebanyak 70 citra *frame* dari *dataset* iLIDS-VID ini mendapatkan akurasi terbaik pada R-20 sebesar 88% menggunakan DBSCAN dengan kecepatan proses sebesar 1,85 detik.

Kata Kunci— *Person Reidentification, Bag of Visual Feature, Pengelompokan Fitur, DBSCAN, Histogram of Oriented Gradient.*

^{1,2} Departemen Ilmu Komputer dan Elektronika, Fakultas MIPA, Universitas Gadjah Mada, Sekip Utara, Bulaksumur, Yogyakarta 55281 INDONESIA (email: yusuf.manshur@mail.ugm.ac.id; wahyo@ugm.ac.id)

#Corresponding Author

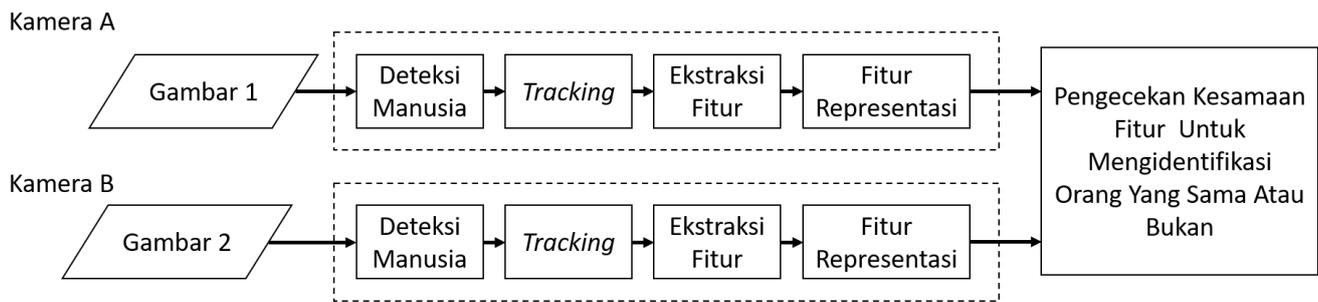
I. PENDAHULUAN

Sistem pengawasan telah banyak digunakan untuk memantau tempat umum, seperti terminal bus, bandara, stasiun, sekolah, perguruan tinggi, dan lobi kantor. Pada umumnya, sistem pengawasan digunakan dengan tujuan meningkatkan sistem keamanan ataupun untuk keperluan pada sistem di bidang lainnya, seperti transportasi, pariwisata, dan pendidikan. Sistem pengawasan biasanya menggunakan beberapa kamera *Closed Circuit Television* (CCTV) yang dipasang untuk memantau beberapa sudut di ruang tertentu.

Kamera-kamera CCTV pada sistem pengawasan terhubung melalui suatu jaringan dan dapat mengirimkan data berupa rekaman video secara *streaming*, sehingga setiap kamera CCTV yang mencakup beberapa area pandang tertentu dapat dipantau secara tidak langsung oleh petugas melalui layar. Sistem ini mengharuskan petugas memantau setiap kejadian yang ditangkap oleh beberapa kamera secara bersamaan. Namun, pengamatan oleh manusia memiliki keterbatasan, antara lain rentan keliru serta membutuhkan banyak waktu dan biaya, sehingga mengurangi efektivitas pengawasan [1]. Diperlukan upaya untuk mengembangkan sistem pengawasan agar pemantauan terhadap setiap kejadian dapat dilakukan dengan lebih optimal. Salah satu solusinya adalah dengan mengembangkan sistem pengawasan secara cerdas (*intelligent surveillance systems*).

Salah satu permasalahan pada proses pengawasan adalah deteksi dan pelacakan objek (*object tracking*). Telah diusulkan suatu sistem untuk mendeteksi bagasi di suatu lorong dan mengidentifikasi pemiliknya melalui rekaman kamera CCTV [2]. Identifikasi dilakukan untuk melacak pemilik bagasi melalui rekaman kamera CCTV di tempat yang berbeda. Sistem ini mengurangi usaha petugas untuk mencari pemilik bagasi tersebut. Solusi ini dinamakan dengan *person reidentification* (Re-ID). Re-ID adalah proses pelacakan (*tracking*) seseorang yang bergerak (disebut *probe*) melintasi beberapa kamera CCTV yang dipasang di berbagai lokasi yang berbeda.

Tugas utama pada proses Re-ID adalah memodelkan objek manusia dari citra yang dihasilkan oleh kamera CCTV, sehingga dapat digunakan untuk mengidentifikasi objek yang sama di antara identitas objek lain dari sudut pandang kamera lainnya (disebut kamera *gallery*). Sistem biometrik barangkali tidak akan memadai, karena resolusi citra yang dihasilkan kamera CCTV cenderung rendah [3]. Selain rendahnya kualitas citra yang dihasilkan kamera CCTV, pada kenyataannya orang



Gbr. 1 Diagram sistem Re-ID.

yang sama akan terlihat sangat berbeda jika diamati dari kamera yang berbeda dan orang yang berbeda bisa jadi akan sama dalam pandangan kamera yang sama. Di antara faktor-faktor yang menyebabkan permasalahan tersebut adalah variasi sudut pandang kamera, pose, pencahayaan, tumpang tindih objek (disebut dengan oklusi), dan latar belakang yang tidak teratur [4].

Oleh sebab itu, makalah ini mengusulkan penggunaan metode berbasis fitur visual yang diharapkan mampu mengatasi permasalahan yang telah disebutkan sebelumnya [5]. Penelitian yang diusulkan berbasis pada metode *Bag of Visual Feature* (BOVF) yang bekerja dengan cara mengelompokkan fitur visual (*feature clustering*) *Histogram of Oriented Gradient* (HOG) dari objek manusia. Makalah ini menggunakan metode pengelompokan fitur, yaitu *Density Based Spatial Clustering of Applications with Noise* (DBSCAN).

II. PERSON RE-IDENTIFICATION (RE-ID)

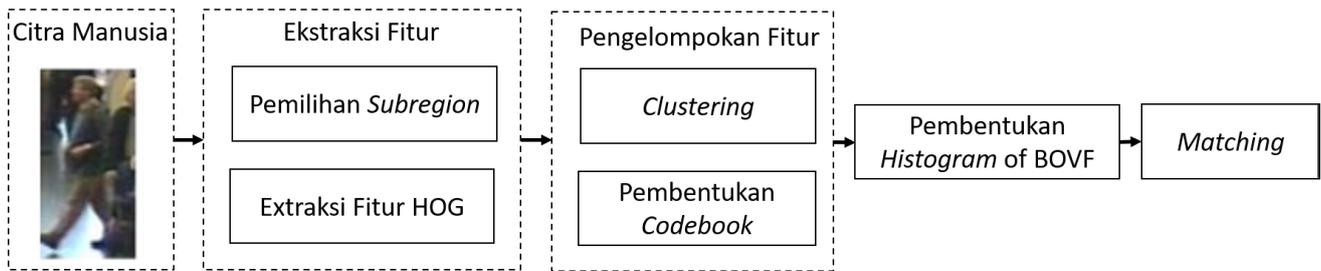
Dalam satu dekade terakhir, Re-ID telah menjadi topik penelitian yang intensif dilakukan. Penelitian memuat cara membuat model yang tepat untuk mendeskripsikan tubuh seseorang dan menentukan metode perhitungan korespondensi yang akurat terhadap dua atau lebih citra dari kamera CCTV yang berbeda [1]. Secara umum, sistem Re-ID ditunjukkan pada Gbr. 1.

Ada beberapa penelitian tentang Re-ID yang telah dilakukan. Sebuah penelitian memperkenalkan sistem Re-ID yang serbaguna dengan pendekatan persepsi simetri dan asimetri dari bagian tubuh seseorang pada citra [3]. Pada awal proses ekstraksi fitur, daerah tubuh yang terdeteksi dibagi dengan pendekatan simetri *silhouette partition*, kemudian fitur dikembangkan dengan *Recurrent Highly Structured Patches* (RHSP). Pada proses pemodelan citra, fitur yang dihasilkan *chromatic-content* menggunakan *HSV histogram* diubah ke dalam bentuk *Weighted Color-Histograms* (WH), kemudian fitur warna pada setiap bagiannya ditingkatkan dengan *Maximally Stable Colour Regions* (MSCR). Metode ini diuji pada *dataset* VIPeR, iLIDS MCTS, dan ETZH dengan hasil skor *Cumulative Matching Characteristic* (CMC) terbaik sekitar 60-70% pada *rank-20*. Namun, metode ini masih menggunakan pendekatan pasif (belum menggunakan teknik *learning*) pada proses ekstraksi fitur maupun *matching*, sehingga menyebabkan metode yang diusulkan membutuhkan data yang sangat besar.

Untuk mengatasi masalah ini, diusulkan perspektif baru untuk Re-ID berdasarkan pembelajaran fitur *unsupervised salience* [4]. Pendekatan ini dilakukan agar proses ekstraksi fitur tidak memerlukan pelabelan identitas (ID) pada prosedur pelatihan. Proses ekstraksi *salience* menggunakan metode *adjacency-patch matching* yang disesuaikan untuk membangun korespondensi yang kuat berdasarkan histogram warna dan SIFT (dColorSIFT) antara pasangan citra. *Salience* yang diekstrak merupakan fitur yang paling menonjol pada citra seseorang (tidak hanya bagian tubuh), antara lain tas, baju, dan celana. Namun, metode yang diusulkan masih menggunakan *dataset* yang belum mencerminkan keberagaman sudut pandang kamera CCTV, sehingga ketika citra diambil dari sudut pandang yang berbeda, akan dihasilkan akurasi yang buruk. Pendekatan *deep learning* mulai digunakan dalam metode metrik pada proses *matching* [6]. Akan tetapi, model dengan pendekatan pembelajaran fitur juga masih berpotensi untuk dikembangkan, termasuk model *Bag of Words* (BOW) yang berkembang menjadi BOVF. Metode *deep-learning* untuk Re-ID dijelaskan lebih detail pada sebuah *paper review* [7].

Penelitian selanjutnya membuat *dataset* berskala besar dan berkualitas tinggi untuk Re-ID dengan nama "Market-1501" [8]. Model yang digunakan adalah BOW yang dapat mengidentifikasi fitur lokal dan memungkinkan dilakukannya *matching* fitur global dengan cepat. Meskipun demikian, oklusi pada gambar masih membawa kendala dalam menghasilkan akurasi yang baik.

Penelitian lainnya mengusulkan sebuah sistem untuk Re-ID dengan penggunaan model representasi BOVF dan *random windows* pada proses ekstraksi fitur lokal [9]. Fungsi *random window* adalah untuk mengatasi masalah oklusi pada objek gambar. Model dibentuk dari fitur HOG [10] yang dilatih menggunakan algoritme *k-means*, sehingga didapatkan *codebook*. Proses *matching* dilakukan dengan menghitung histogram BOVF (HBOVF) yang dibentuk dengan metode *windows quantization*. Namun, penggunaan *k-means* menjadikan setiap model dibentuk tidak cukup merepresentasikan identitas citra yang berbeda, karena pada proses pembentukannya jumlah kelas atau kluster fitur visual adalah sama untuk masing-masing identitas citra. Salah satu algoritme pengelompokan lainnya, yaitu DBSCAN [11], mampu melakukan pengelompokan dengan jumlah kluster berdasarkan kerapatan data. Algoritme DBSCAN juga dapat digunakan untuk mengeliminasi *noise* pada proses



Gbr. 2 Diagram proses sistem Re-ID yang diusulkan.

pengelompokan data dengan jumlah yang banyak dan berdimensi tinggi secara efisien [12].

Oleh sebab itu, terinspirasi dari penelitian sebelumnya [9], pada makalah ini digunakan BOVF dengan *random windows* tetapi untuk mengatasi masalah *k-means*, yang harus sudah menentukan jumlah kelompok *feature*, dan diusulkan untuk memanfaatkan DBSCAN sebagai metode pengelompokan fitur visual.

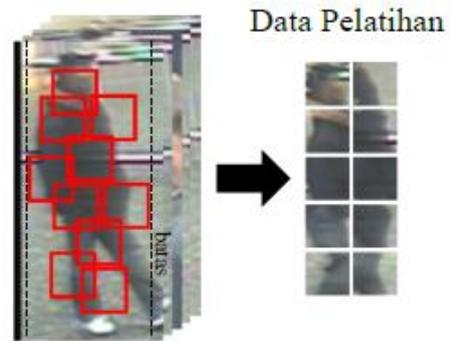
III. METODOLOGI

Meskipun sistem Re-ID biasanya diawali dengan pendeteksian objek manusia, pada makalah ini diasumsikan objek manusia pada citra telah terdeteksi, sehingga penelitian hanya berfokus pada cara melakukan representasi objek manusia untuk kemudian dilakukan *re-identification*. Seperti terlihat pada Gbr. 2, metode yang diusulkan berawal dari sebuah citra masukan berupa objek manusia pada data latih berukuran 124×64 piksel yang kemudian dilakukan proses dengan ekstraksi fitur, pengelompokan fitur, pembentukan HBOVF, dan terakhir proses *matching* dengan objek citra manusia lainnya. Secara detail, proses ini akan dijelaskan satu persatu pada subbagian berikutnya.

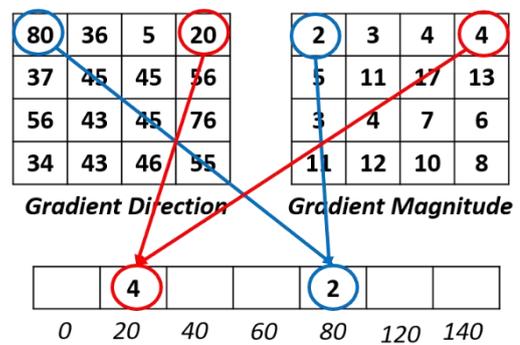
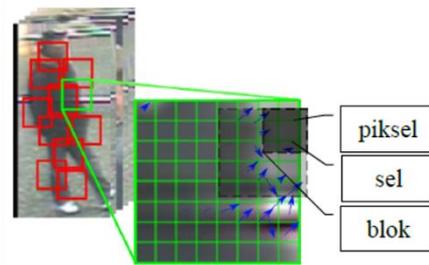
A. Ekstraksi Fitur

Pada setiap data masukan citra objek manusia dilakukan proses ekstraksi fitur-fitur lokal. Namun, terlebih dahulu dilakukan prapemrosesan terhadap citra dengan mengubah seluruh citra ke dalam format *grayscale*. Tahap ekstraksi fitur diawali dengan proses pemilihan *subregion*. *Subregion* adalah bagian-bagian lebih kecil dari citra yang diekstrak menggunakan *window* dengan ukuran tertentu. Untuk mendapatkan *subregion*, digunakan metode pada penelitian sebelumnya, yaitu *random windows* [9]. Metode ini menggunakan sebanyak *n* *window* berbentuk persegi dengan sisi ukuran *s* piksel yang disebar secara acak di atas sebuah citra. *Random windows* dikenakan ke setiap citra yang ada pada semua data citra. Jika data latih memiliki sebanyak *m* sampel citra berurutan, maka pada tahap pelatihan ini didapat $n \times m$ *subregion*. Ilustrasi proses pemilihan *subregion* ditunjukkan pada Gbr. 3.

Pada tahap selanjutnya dilakukan ekstraksi fitur pada *subregion* yang dihasilkan dari tahap sebelumnya. Fitur yang digunakan adalah *9-bins* HOG [10], dengan ukuran sel 8×8 piksel, dan ukuran blok 2×2 sel. Pada Gbr. 4 ditunjukkan hasil ekstraksi *subregion*, kemudian bobot dan orientasi kemiringan setiap piksel dihitung. Fitur sel dihitung dengan melakukan *voting* yang menghasilkan *9-bins* HOG.



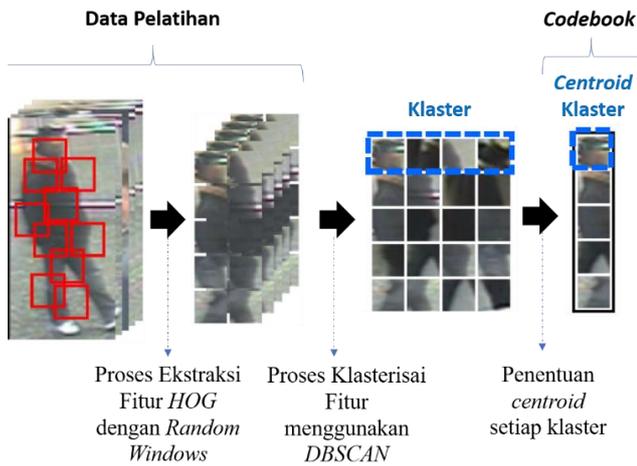
Gbr. 3 Ilustrasi pembentukan *subregion*.



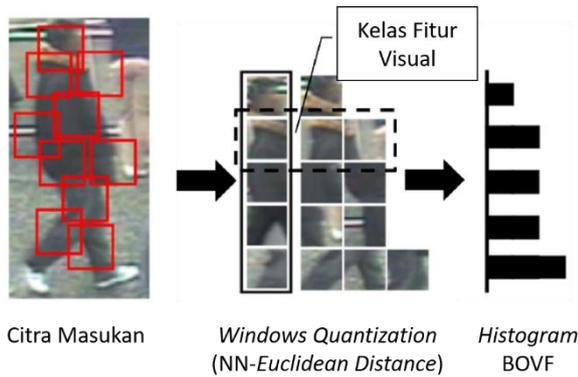
Gbr. 4 Ilustrasi pembentukan *subregion*.

B. Pengelompokan Fitur dan Pembentukan Codebook

Langkah selanjutnya adalah proses pengelompokan fitur untuk pembentukan *codebook*. Proses ini diawali dengan klusterisasi fitur menggunakan algoritme DBSCAN. Fitur-fitur yang dikelompokkan adalah fitur-fitur vektor yang dihasilkan dari proses ekstraksi fitur setiap citra pada data pelatihan. Untuk menjalankan algoritme DBSCAN, dibutuhkan dua parameter, yaitu *Eps* dan *Minpts*, yang didefinisikan sebagai nilai jarak keanggotaan ke titik pusat dan banyaknya titik minimal sebuah kelompok dianggap sebagai sebuah kluster.



Gbr. 5 Ilustrasi pembentukan codebook.



Gbr. 6 Ilustrasi pembentukan HBOVF.

Parameter *Minpts* ditentukan berkisar dengan nilai $\ln(n)$, dengan n adalah banyaknya data latih [13], sedangkan nilai *Eps* ditentukan dengan langkah-langkah yang diusulkan pada [11]. Pada tahap selanjutnya, dibentuk *codebook* yang digunakan sebagai representasi *probe* (*descriptor* citra *probe*). Keseluruhan proses pembentukan *codebook* diilustrasikan pada Gbr. 5. Proses dimulai dari ekstraksi fitur, klasterisasi fitur, dan kemudian penentuan *centroid* klaster.

Seperti pada Gbr. 5, *codebook* terdiri atas k fitur visual yang merupakan *centroid* dari setiap klaster hasil klasterisasi fitur, dengan asumsi titik-titik pada klaster saling berdekatan secara terpusat (*hypersphere*). Misalnya sebuah klaster adalah himpunan titik k dimensi ($A \in \mathbb{R}^k$) dan titik *centroid* dari himpunan tersebut adalah p , maka titik p dicari dengan menghitung rata-rata vektor seluruh titik pada A .

C. Pembentukan Histogram BOVF (HBOVF)

Pada tahap ini dilakukan pembentukan HBOVF yang akan digunakan sebagai representasi citra *query* maupun kandidat, seperti ditunjukkan pada Gbr. 6. Setelah didapatkan sebanyak k klaster dari proses klasterisasi fitur, dilakukan proses pembentukan histogram HBOVF menggunakan metode *window quantization* yang menghasilkan histogram atau vektor k dimensi.

Metode *window quantization* diawali dengan proses klasifikasi masing-masing fitur yang telah didapatkan pada

proses ekstraksi fitur. Masing-masing fitur tersebut diklasifikasikan ke dalam salah satu *code* penyusun *codebook*, yang setiap *code* mewakili kelas fitur visual. Proses klasifikasi menggunakan algoritme 1-*Nearest Neighbors* (1-NN) dengan metrik jarak Euclidean. HBOVF dibentuk dengan menghitung jumlah anggota setiap kelas fitur visual, kemudian histogram dinormalisasi menggunakan l_1 -norm. HBOVF dari sebuah citra a mengembalikan sebuah vektor berdimensi k , $H_{BOVF}(a) = [a_1 \ a_2 \ \dots \ a_k]$, dengan h_1, h_2, \dots, h_k merupakan jumlah anggota masing-masing kelas fitur visual dari hasil *windows quantization*.

D. Menghitung Tingkat Kemiripan (Matching)

Tahapan ini merupakan proses terakhir dari sistem Re-ID yang diusulkan. Tahap ini terdiri atas dua proses, yaitu proses *matching* dan pengurutan. Proses *matching* dilakukan untuk menyelidiki atau membandingkan citra *query* dengan seluruh citra kandidat yang ada pada *gallery*. Proses pembandingan diawali dengan pembentukan HBOVF untuk citra *probe/query* dan citra kandidat yang ada pada *gallery*. Kemudian dihitung jarak antara citra *query* dan citra kandidat berdasarkan HBOVF masing-masing. Nilai jarak dari kedua HBOVF citra menunjukkan tingkat kemiripan dari dua citra tersebut. Nilai jarak dari kedua HBOVF citra dihitung menggunakan metrik jarak Euclidean. Misalnya q adalah citra *query* dan g adalah citra kandidat, maka jarak keduanya, yaitu d , dihitung menggunakan (1). Hasil (1) akan menunjukkan kemiripan antara dua citra. Semakin rendah nilai jarak, semakin tinggi tingkat kemiripan dua citra tersebut.

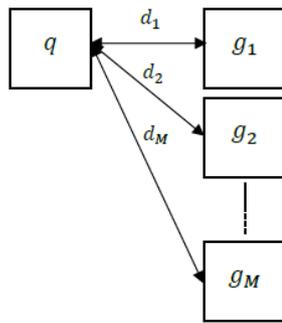
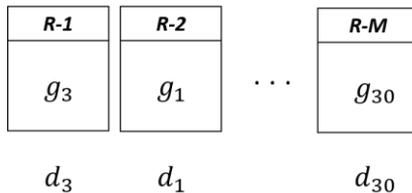
$$d(H_{BOVF}(g), H_{BOVF}(q)) = \sqrt{\sum_{i=1}^k (g_i - q_i)^2} \tag{1}$$

Proses *matching* dilakukan dengan skenario *single-gallery-shot* (MvsS). Perhitungan jarak antara citra *query* (q) dengan setiap data citra pada *gallery* dilakukan menggunakan (1). Misalnya $G = \{g_1, g_2, \dots, g_M\}$ adalah data *gallery* dengan M citra dan $D = \{d_1, d_2, \dots, d_M\}$ adalah himpunan hasil perhitungan jarak M citra kandidat terhadap citra *query* q , maka proses *matching* diilustrasikan pada Gbr. 7.

Setelah proses *matching* selesai, hasil perhitungan jarak, yaitu D , diurutkan menaik (*ascending*), sehingga didapatkan daftar citra sebanyak M dengan posisi citra pada R - n , dengan n adalah peringkat kandidat seperti pada Gbr. 8. Berdasarkan Gbr. 8, citra g_3 berada di posisi $R-1$ (peringkat 1), yang artinya citra tersebut adalah citra yang paling diduga merupakan orang yang sama dengan citra q .

IV. HASIL DAN PEMBAHASAN

Pada bagian ini ditunjukkan hasil dari proses pelatihan dan pengujian dari model sistem Re-ID yang diusulkan. Proses pelatihan pada sistem Re-ID dilakukan menggunakan metode *random windows* dan algoritme DBSCAN. Evaluasi terhadap sistem didasarkan pada nilai akurasi yang dihitung menggunakan nilai *Cumulative Matching Characteristic* (CMC). CMC adalah salah satu ukuran yang digunakan untuk mengevaluasi kinerja sistem identifikasi [14]. Sistem identifikasi menghasilkan peringkat (R) daftar kandidat. CMC

Gbr. 7 Skenario proses *matching*.

Gbr. 8 Proses perhitungan kemiripan.

menunjukkan bahwa identitas *query* terdapat pada daftar kandidat n teratas ($R-n$) yang dihasilkan oleh sistem, dengan n adalah peringkat kandidat, dimulai dari 1 sampai dengan banyaknya kandidat. Dapat dikatakan bahwa semakin tinggi nilai CMC, akurasi sistem dianggap baik, dan sebaliknya.

A. Akuisisi Data

Pada makalah ini, *dataset* iLIDS-VID digunakan sebagai data pelatihan maupun pengujian [15]. Data pelatihan digunakan untuk melatih model atau identitas *visual probe*. Data pelatihan terdiri atas sejumlah citra berurutan dari salah satu kamera yang menampilkan *probe*. Data pengujian terdiri atas citra *query* dan *gallery*. Citra *query* adalah data uji yang merupakan satu sampel citra *probe* (*single-shot matching*) dari sumber kamera yang sama dengan data pelatihan. Kemudian, *gallery* adalah sejumlah citra dari orang-orang yang berbeda yang diambil dari kamera lainnya sebagai citra kandidat. Perancangan data untuk sistem Re-ID disajikan pada Tabel I.

B. Pengaruh *Minpts* pada DBSCAN

Pada bagian ini akan dijelaskan pengaruh *Minpts* terhadap akurasi dan kecepatan sistem. Seperti ditunjukkan pada Tabel II, diketahui bahwa semakin tinggi nilai *Minpts* yang digunakan, akurasi semakin rendah, tetapi kecepatan semakin tinggi untuk semua peringkat R . Dari hasil ini, *MinPts* dengan nilai 4 akan digunakan sebagai parameter optimal untuk pengujian yang nantinya dibandingkan dengan *k-means*.

C. Pengaruh Jumlah dan Ukuran *Window*

Pada bagian ini akan dijelaskan pengaruh jumlah *window*, ukuran *window* terhadap akurasi, dan kecepatan sistem. Seperti ditunjukkan pada Tabel III, diketahui bahwa jumlah *window* yang terlalu banyak atau sedikit akan mengurangi akurasi. Sementara itu, semakin banyak jumlah *window*, kecepatan menjadi semakin rendah. Selanjutnya, dari Tabel IV diketahui bahwa jika ukuran *window* semakin kecil, akurasi dan kecepatan cenderung menjadi semakin tinggi. Akan tetapi,

TABEL I
PERANCANGAN DATA UNTUK SISTEM RE-ID

Jenis Data	Sumber Kamera	Total Frame
Data <i>query/probe</i>	Cam1	1
Data pelatihan	Cam1	20
Data <i>gallery</i>	Cam2	50

TABEL II
PENGARUH NILAI *MINPTS* TERHADAP AKURASI DAN KECEPATAN

<i>MinPts</i>	Akurasi CMC (%) dalam Rank (R)					Waktu (detik)
	R-1	R-5	R-10	R-15	R-20	
4	10,67	42,67	68,33	80,67	88,00	1,85
6	6,67	32,00	55,67	71,33	81,00	1,32
8	4,00	23,33	40,33	50,67	62,33	1,20

TABEL III
PENGARUH NILAI JUMLAH *WINDOW* TERHADAP AKURASI DAN KECEPATAN

Jumlah	Akurasi CMC (%) dalam Rank					Waktu (detik)
	R-1	R-5	R-10	R-15	R-20	
64	13,67	39,33	58,33	69,33	82,00	4,98
32	10,67	42,67	68,33	80,67	88,00	1,85
16	8,33	33,00	60,33	75,00	86,00	0,69

TABEL IV
PENGARUH NILAI UKURAN *WINDOW* TERHADAP AKURASI DAN KECEPATAN

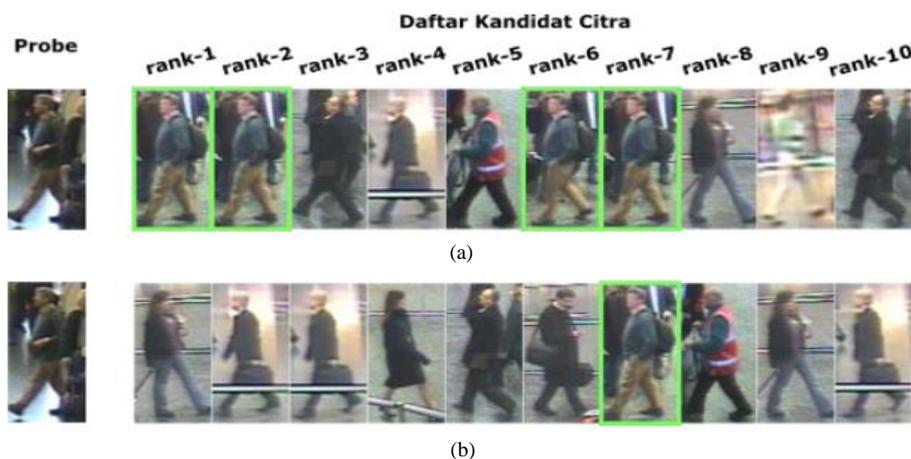
Ukuran	Akurasi CMC (%) Dalam Rank					Waktu (detik)
	R-1	R-5	R-10	R-15	R-20	
32	14,33	41,67	60,67	72,33	82,67	5,35
24	13,67	39,00	59,67	72,67	81,67	2,94
16	10,67	42,67	68,33	80,67	88,00	1,85

khusus $R-1$ memberikan hasil yang berbeda, yaitu semakin rendah ukuran *window*, semakin rendah juga akurasinya. Dari hasil ini, jumlah *window* 32 dengan ukuran 16 digunakan sebagai parameter optimal untuk pengujian yang nantinya dibandingkan dengan *k-means*.

D. Perbandingan dengan Metode Lain

Selanjutnya, model Re-ID menggunakan DBSCAN dengan pengaturan parameter terbaik, yaitu $n-s-m=32-16-4$ (dengan n adalah jumlah *window*, s adalah ukuran *window*, dan m adalah *Minpts*), dibandingkan terhadap model Re-ID menggunakan *k-means* yang diusulkan pada penelitian sebelumnya [9], dengan pengaturan jumlah dan ukuran *window* yang sama. Karena algoritme *k-means* harus menentukan jumlah kluster yang akan diikuti dalam sistem Re-ID, dilakukan pengujian beberapa nilai kluster, yaitu 5, 15, dan 20, yang kemudian didapatkan bahwa nilai kluster 15 pada *k-means* mendapatkan hasil terbaik. Hasil ini kemudian dibandingkan dengan metode yang diusulkan, seperti ditunjukkan pada Tabel V.

Dari data yang ditunjukkan pada Tabel V, dapat diketahui bahwa nilai akurasi metode yang diusulkan dengan menggunakan DBSCAN relatif lebih tinggi dibandingkan dengan yang menggunakan *k-means* [9], dengan pengaturan jumlah dan ukuran *window* yang sama, kecuali pada $R-1$ dan $R-2$. Sedangkan untuk kecepatan, *k-means* lebih cepat daripada DBSCAN sebesar 0,21 detik. Hal tersebut disebabkan penggunaan DBSCAN dalam beberapa kasus *probe*



Gbr. 9 Ilustrasi hasil *ranking* daftar citra kandidat untuk *probe* person229, (a) menggunakan DBSCAN, (b) menggunakan *k-means* [8].

TABEL V
PERBANDINGAN AKURASI DENGAN METODE [8]

Algoritme	Akurasi CMC (%) dalam Rank					Waktu (s)
	R-1	R-5	R-10	R-15	R-20	
Usulan	10,67	42,67	68,33	80,67	88,00	1,85
[9]	13,67	43,33	63,67	76,33	87,00	1,64

memerlukan perhitungan dengan jumlah atribut representasi yang lebih banyak. Untuk menerapkan Re-ID pada sistem *monitoring* secara *real-time*, akurasi lebih diutamakan, karena kecepatan yang lebih lambat dapat diatasi dengan penggunaan *hardware* yang lebih baik dalam hal pemrosesan atau komputasi. Selanjutnya, sebagai ilustrasi, Gbr. 9 menunjukkan ilustrasi visual perbandingan hasil Re-ID antara algoritme DBSCAN dan *k-means clustering*. Kotak berwarna hijau menandakan hasil yang sesuai, yang artinya merupakan orang yang sama dengan gambar pada *probe*.

V. KESIMPULAN

Makalah ini mengusulkan perbaikan metode Re-ID berbasis BOVF dengan memanfaatkan *clustering* DBSCAN sebagai metode pengelompokan citra dibandingkan dengan *k-Means clustering*. Secara umum, metode yang diusulkan berawal dari sebuah masukan citra berupa objek manusia pada data latih berukuran 124×64 piksel yang kemudian dikenai proses ekstraksi fitur, pengelompokan fitur, pembentukan histogram BOVF, dan terakhir proses *matching* dengan objek citra manusia lainnya. Hasil menunjukkan bahwa akurasi metode yang diusulkan pada R-20 sebesar 88%, mengungguli metode yang diusulkan sebelumnya. Namun, metode yang diusulkan masih lebih lambat dibandingkan metode sebelumnya.

REFERENSI

[1] A. Bedagkar-Gala dan S.K. Shah, "A Survey of Approaches and Trends in Person Re-Identification," *Image and Vision Computing*, Vol. 32, No. 4, hal. 270–286, 2014.
 [2] Wahyono, A. Filonenko, dan K.H. Jo, "Unattended Object Identification for Intelligent Surveillance Systems Using Sequence of Dual Background Difference," *IEEE Transactions on Industrial Informatics*, Vol. 12, No. 6, hal. 2247–2255, 2016.

[3] M. Farenzena, L. Bazzani, A. Perina, V. Murino, dan M. Cristani, "Person Re-Identification by Symmetry-Driven Accumulation of Local Features," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, hal. 2360–2367.
 [4] R. Zhao, W. Ouyang, dan X. Wang, "Unsupervised Saliency Learning for Person Re-Identification," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2013, hal. 3586–3593.
 [5] M.Y. Manshur, "Implementasi Bag of Visual Feature Menggunakan Algoritme Pengelompokan DBSCAN untuk Person Re-Identification," B. Comp. thesis, Universitas Gadjah Mada, Yogyakarta, Indonesia, Jan. 2020.
 [6] D. Yi, Z. Lei, dan S. Z. Li, "Deep Metric Learning for Practical Person Re-Identification," *CoRR*, abs/1407.4, 2014.
 [7] D. Wu, S.-J. Zheng, X.-P. Zhang, C.-A. Yuan, F. Cheng, Y. Zhao, Y.-J. Lin, Z.-Q. Zhao, Y.-L. Jiang, dan D.-S. Huang, "Deep Learning-based Methods for Person Re-Identification: A Comprehensive Review," *Neurocomputing*, Vol. 337, hal. 354–371, 2019.
 [8] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, dan Q. Tian, "Scalable Person Re-identification: A Benchmark," *IEEE International Conference on Computer Vision (ICCV)*, 2015, hal. 1116–1124.
 [9] Wahyono dan K.H. Jo, "Person Re-Identification Using Bag of Visual Feature on Random Windows," dipresentasikan pada *International Workshop of Frontiers in Computer Vision*, Seoul, Korea, 2017.
 [10] N. Dalal dan B. Triggs, "Histograms of Oriented Gradients for Human Detection," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR05)*, 2005, hal. 1–8.
 [11] J. Sander, M. Ester, H.-P. Kriegel, dan Xiaowei Xu, "Density-based Clustering in Spatial Databases: The Algorithm GDBSCAN and Its Applications," *Data Mining and Knowledge Discovery*, Vol. 2, No. 2, hal. 169–194, 1998.
 [12] K. Khan, S.U. Rehman, K. Aziz, S. Fong, dan S. Sarasvady, "DBSCAN: Past, Present, and Future," *5th International Conference on the Applications of Digital Information and Web Technologies (ICADIWT)*, 2014, hal. 232–238.
 [13] D. Birant dan A. Kut, "ST-DBSCAN: An Algorithm for Clustering Spatial-temporal Data," *Data & Knowledge Engineering*, Vol. 60, No. 1, hal. 208–221, 2007.
 [14] R.M. Bolle, "The Relation Between the ROC Curve and the CMC," *Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05)*, 2005, hal. 15–20.
 [15] M. Li, X. Zhu dan S. Gong, "Unsupervised Person Re-Identification by Deep Learning Tracklet Association," *Proc. European Conference on Computer Vision*, 2018, hal. 1–17.