

## ANALISIS DAN IMPLEMENTASI DETEKSI KEPALA-BAHU UNTUK PEOPLE COUNTING DENGAN MULTILEVEL HOG DETECTOR

Dian Hadiana<sup>1</sup>, Bedy Purnama <sup>2</sup>, Febryanti Hevanie<sup>3</sup>

<sup>1</sup>Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

---

### Abstrak

Pengamatan melalui citra menjadi alternatif untuk melakukan perhitungan jumlah pengunjung. Sistem akan melakukan tracking objek, sekaligus counting pada objek tersebut. Salah satu algoritma yang digunakan untuk membangun sistem counting tersebut adalah Histograms of Oriented Gradients (HOG). Dengan penggunaan multilevel HOG diharapkan mampu meningkatkan kinerja deteksi pada objek pengamatan. Tracking objek yang dilakukan adalah dengan mengamati bagian atas tubuh manusia, yaitu kepala dan bahu. Data latih terdiri dari 100 citra bagian atas tubuh manusia, serta 100 citra acak, diantaranya berupa bagian bawah tubuh manusia. Pada tahap pengujian, dilakukan pengujian terhadap beberapa citra yang jumlah objeknya bervariasi. Hasil pengamatan menunjukkan citra tanpa latar belakang menghasilkan akurasi rata-rata 82.54%.

**Kata Kunci :** image processing, head-shoulder detection, multilevel HOG detector, SVM, people counting.

---

### Abstract

Observations through the image became an alternative to calculating the number of visitors. The system will do object tracking, as well as counting on that object. One of the algorithms used to construct the counting system is the Histograms of Oriented Gradients (HOG). With the use of multilevel HOG is expected to improve detection performance on the object of observation. Object tracking is performed by observing the upper part of the human body, namely the head and shoulders. Training data consists of 100 image of the top of the human body, as well as the 100 random images, including a lower part of the human body. In the testing phase, carried out tests on some image object number varies. The results showed no background images produces better accuracy, ie 82.54%.

**Keywords :** image processing, head-shoulder detection, multilevel HOG detector, SVM, people counting.

---

Telkom  
University

## 2. Dasar Teori

### 2.1 Citra

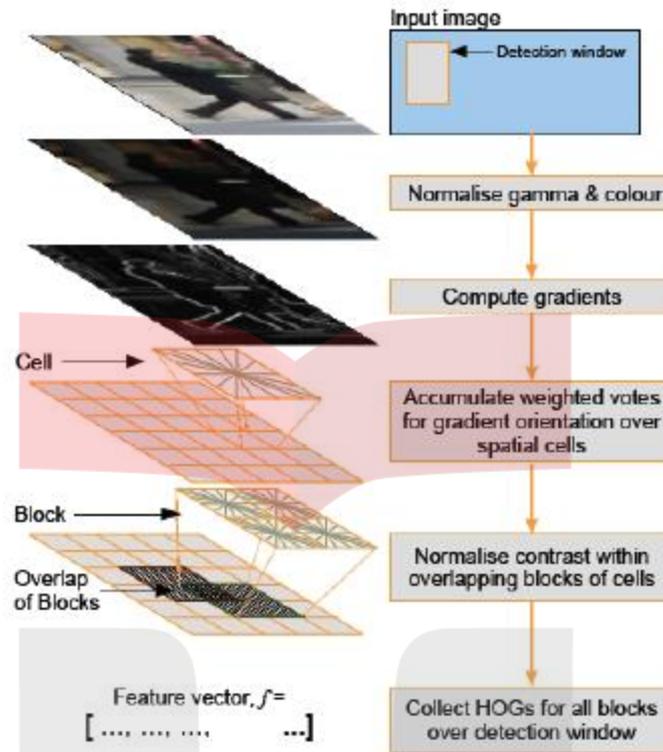
Citra pada dasarnya merupakan *array* dua dimensi. Dua dimensi tersebut menggambarkan ruang pergerakan citra (*spatial*). Citra terdiri dari beberapa *channel* warna. *Channel* tersebut terdiri dari merah (*red*), hijau (*green*), dan biru (*blue*). Ada pun citra yang terdiri dari satu *channel* sehingga hanya menghasilkan warna dengan intensitas dari hitam ke putih, atau lebih dikenal sebagai *grayscale*. Suatu citra digital direpresentasikan dengan sebuah matriks yang masing-masing elemennya merepresentasikan nilai intensitas atau kedalaman warna [3].

### 2.2 People Counting

*People counting* adalah perhitungan objek yang bergerak berupa manusia yang melintasi bagian tertentu di sebuah ruangan pada waktu tertentu. Akurasi *people counting* digunakan untuk memastikan bahwa bangunan tersebut berada di bawah tingkat aman hunian. Sistem menawarkan cara yang relatif akurat untuk mengelola kapasitas suatu ruangan. Counter berbasis citra dapat rentan terhadap perubahan tingkat cahaya dan bayangan, yang dapat menyebabkan penghitungan tidak akurat [11]. Akhir-akhir ini, algoritma yang kuat dan adaptif telah dikembangkan yang dapat mengimbangi perilaku ini dan akurasi penghitungan yang sangat baik saat ini dapat diperoleh untuk menghitung baik outdoor dan indoor.

### 2.3 Histogram of Oriented Gradients

*Histogram of Oriented Gradients* (HOG) adalah fitur deskriptor yang digunakan dalam pengolahan citra untuk mendeteksi obyek [3][11]. Teknik ini menghitung kejadian orientasi gradien dalam porsi lokal dari suatu gambar. Metode ini mirip dengan *edge orientation histograms*, fitur skala invarian mengubah deskripsi, dan bentuk konteks, tetapi berbeda dalam hal itu dihitung pada grid padat sel seragam spasi dan menggunakan tumpang tindih normalisasi kontras lokal untuk meningkatkan akurasi [3].



**Gambar 2.1 Histogram of Oriented Gradients (HOG) [5]**

Penjelasan untuk gambar di atas adalah sebagai berikut:

- 1) Normalisasi Gamma, adalah hasil dari akar kuadrat setiap channel (merah, hijau dan biru)
- 2) Komputasi Gradien, citra grayscale difiter dengan [-1, 0, 1], seperti yang tertera pada persamaan (2.1) dan persamaan (2.2). Untuk citra RGB, masing-masing channel akan dihitung, lalu dicari nilai maksimumnya.

$$D_x = [-1 \quad 0 \quad 1] \quad \dots \quad (2.1)$$

$$D_y = [-1 \quad 0 \quad 1]^T \quad \dots \quad (2.2)$$

$$|G| = \sqrt{I_x^2 + I_y^2} \quad \dots \quad (2.3)$$

$$\theta = \text{arc tan} \frac{I_y}{I_x} \quad \dots \quad (2.4)$$

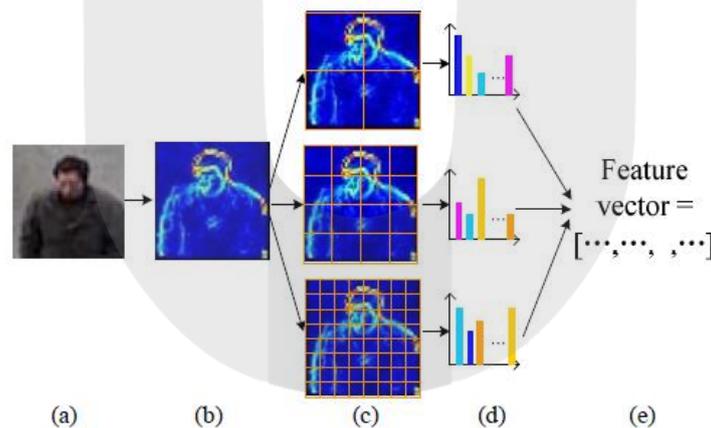
$D_x$  serta  $D_y$  adalah matriks konvolusi yang digunakan untuk filter citra asli.  $I_x$  merupakan matrix hasil dari filter oleh  $D_x$ , sedangkan  $I_y$  merupakan matrix hasil dari filter oleh  $D_y$ . Hasil dari komputasi gradien ini ada 2 buah matriks, yaitu  $|G|$  yang

merupakan matriks gradient magnitude (scalar), yang kedua  $\theta$  adalah orientasi dari gradient tersebut.

- 3) Sel dan Blok, citra akan dibagi ke dalam beberapa sel. Sedangkan beberapa sel dikelompokkan menjadi sebuah blok.
- 4) Normalisasi Blok, dari beberapa normalisasi blok [3], dipilih L2-Hys untuk mengurangi pengaruh variasi lokal dalam pencahayaan dan kontras *foreground-background*.

### 2.3.1 Multi HOG Feature

HOG adalah descriptor yang sangat baik untuk menangkap distribusi gradien intensitas lokal benda. Ini telah berhasil digunakan untuk mendeteksi tubuh holistik manusia. Dalam rangka meningkatkan kinerja HOG, digunakan *multilevel* HOG fitur untuk menggambarkan kepala-bahu [11].



**Gambar 2.2 Multi HOG**

Penjelasan untuk gambar 2.2 adalah sebagai berikut:

- a) Input citra, berupa citra berukuran 48x64 pixel
- b) Hasil gradient magnitude dari input citra
- c) Sel-sel pada 3 level yang berbeda. Level pertama terdiri dari 2x2 sel, sel pada level ini berukuran 24x32 pixel. Level kedua terdiri dari 4x4 sel, sel pada level ini berukuran 12x16 pixel. Level ketiga terdiri dari 8x8 sel, sel pada level ini berukuran 6x8 pixel.
- d) Histogram pada masing-masing level
- e) Feature vector akhir dari semua level

## 2.4 Support Vector Machine

*Support vector machine* (SVM) adalah metode klasifikasi yang bekerja dengan cara mencari *hyperplane* dengan margin terbesar. *Hyperplane* adalah garis batas pemisah data antar-kelas, sedangkan margin adalah jarak antara *hyperplane* dengan data terdekat pada masing-masing kelas. Adapun data terdekat dengan *hyperplane* pada masing-masing kelas inilah yang disebut *support vector* [7].

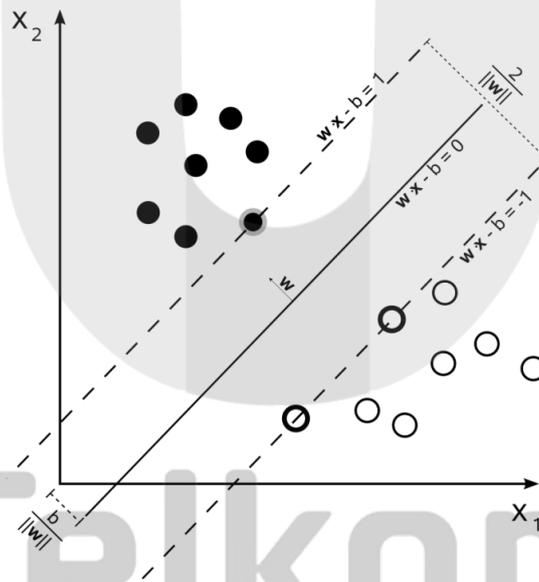
Suatu permasalahan klasifikasi linier dengan data latih  $x_i \in R^n$  dan kelas label  $y_i \in \{-1, +1\}$  dirumuskan pada:

$$D = \{(x_1, y_1), \dots, (x_l, y_l)\}, x \in R^n, i = 1, \dots, l \quad \dots \quad (2.5)$$

Fungsi *hyperplane* SVM didefinisikan sebagai berikut:

$$f(x) = w^T \cdot x_i + b = 0 \quad \dots \quad (2.6)$$

Nilai  $w$  adalah koefisien vektor dan  $b$  adalah nilai bias.



**Gambar 2.3 Support Vector Machine (SVM)**

*Hyperplane* yang dicari adalah *hyperplane* dengan nilai margin terbesar. Memaksimalkan margin dapat dilakukan dengan menyelesaikan persamaan berikut:

$$\min_w \frac{1}{2} \|w\|^2 \quad \dots \quad (2.7)$$

dengan konstrain  $y_i((w^T \cdot x_i) + b) \geq 1, i = 1, \dots, l$  dengan  $x_i$  adalah data latih,  $y_i$  adalah kelas label,  $w$  dan  $b$  adalah parameter yang dicari nilainya.

Pada contoh gambar 2.3 dapat dilihat, dari sekumpulan data tersebut terdiri dari 2 kelas, titik-titik berwarna hitam (kelas pertama), dan titik-titik berwarna putih (kelas kedua). Pada kelas pertama ada satu *support vector*. Sedangkan pada kelas kedua ada dua *support vector*.



## 5. Kesimpulan dan Saran

### 5.1 Kesimpulan

Setelah melakukan implementasi, pengujian dan analisis, maka dapat diambil beberapa kesimpulan, yaitu:

- 1) Multi HOG *detector* bisa mendeteksi objek kepala dan bahu dengan cukup baik dengan akurasi rata-rata di atas 80%.
- 2) Perbedaan ragam pose pada data latih, mempengaruhi hasil akurasi. Penggunaan data latih beragam pose hasilnya lebih baik untuk data uji seragam maupun beragam pose jika dibandingkan dengan data latih seragam.
- 3) Multi HOG *detector* masih kurang baik dalam mendeteksi objek yang jumlahnya banyak dan berhimpitan. Akan tetapi jika jumlah objek banyak dengan pose berdiri biasa dan tidak berhimpitan menghasilkan akurasi lebih baik.

### 5.2 Saran

- 1) Untuk meningkatkan akurasi, dapat dioptimasi dengan mengubah metode counting.
- 2) Penambahan data latih, terutama data positif untuk objek yang berhimpitan, tidak hanya objek tunggal pada satu data latih positif. Untuk data negative, menambahkan yang lebih spesifik, seperti bagian tubuh lain, diharapkan bisa meningkatkan akurasi
- 3) Untuk deteksi berbasis video, akan lebih baik jika menggunakan *background subtraction* sehingga lebih mudah memproses *frame per frame*

## Daftar pustaka

- [1] Aziz, K., Merad, D., Fertil, B. and Thome, N., 2010, *Pedestrian Head Detection and Tracking Using Skeleton Graph for People Counting in Crowded Environments*, Marseille, LSIS – IM.
- [2] Chan, C. 2008, *Multi-scale Local Binary Pattern Histogram for Face Recognition*, Surrey, University of Surrey.
- [3] Dalal, N. and Triggs, B., 2005, *Histograms of Oriented Gradients for Human Detection*, Montbonnot, INRIA Rhone-Alps.
- [4] Guraya, F., Bayle, P. and Cheikh, F., 2009, *People Tracking via a Modified CAMSHIFT Algorithm*, Gjovik, Gjovik University College, Dijon, Universite de Bourgogne.
- [5] Lin, D. S., Liu, Y. M., Jhu, Y. R., 2013, *A Robust Image Descriptor for Human Detection Based on HOG and Weber's Law*, Taiwan, National Dong Hwa University.
- [6] Maturana, D., Mery, D. and Soto, A., 2009, *Face Recognition with Local Binary Patterns, Spatial Pyramid Histograms and Naive Bayes Nearest Neighbor classification*, Santiago, Pontificia Universidad Católica.
- [7] Nugroho, A. S., Witarto, A. B., Handoko, D., 2003, *Support Vector Machine: Teori dan Aplikasinya dalam Bioinformatika*, Kuliah Umum Ilmu Komputer.
- [8] Pishchulin, L., Jain, A., Wojek, C., Thormählen, T. and Schiele, B., 2011, *In Good Shape: Robust People Detection based on Appearance and Shape*, Saarbrücken, MPI Informatics.
- [9] Prastianto, J. A., 2012, *Analisis Modified CAMShift Algorithm untuk People Counting berbasis Video Processing*, Bandung, IT Telkom.
- [10] Spinello, L. and Arras K., 2011, *People Detection in RGB-D Data*, Freiburg, University of Freiburg.
- [11] Zeng, C. and Ma, Ha., 2010, *Robust Head-shoulder Detection by PCA-Based Multilevel HOG-LBP Detector for People Counting*, Beijing, Beijing University of Posts and Telecommunications.
- [12] Zhu, Q., Avidan, S., Yeh, M. and Cheng, K., 2006, *Fast Human Detection Using a Cascade of Histograms of Oriented Gradients*, Cambridge, Mitsubishi Electric Research Laboratories, Inc.