

**PENGENALAN GENDER MEMANFAATKAN WAJAH
MANUSIA DENGAN MENGGUNAKAN METODE
KLASIFIKASI NEAREST NEIGHBOR**

Ardilla Ayu Dewanti Ridwan¹, Ivanna K. Timotius², Iwan Setyawan³

Program Studi Teknik Elektro

Fakultas Teknik Elektronika dan Komputer

Universitas Kristen SatyaWacana

ardilladewanti@gmail.com¹, ivanna.timotius@staff.uksw.edu²,

iwana.setyawan@ieee.org³

INTISARI

Pada tulisan ini, dirancang sebuah sistem pengenalan *gender* memanfaatkan citra wajah manusia. Sistem pengenalan *gender* yang dirancang terdiri dari tahap pemrosesan data awal yang terdiri atas tahap *cropping* citra, konversi citra RGB ke *grayscale*, *histogram equalization*, dan *resizing* citra, serta metode pengenalan *gender*. Pada tulisan ini, dipakai klasifikasi *nearest neighbor* dengan jarak Euclidean sebagai metode klasifikasi, sedangkan performa sistem dinilai menggunakan metode *cross validation*. Selain itu, penilaian performa sistem pada tulisan ini dilakukan dengan skenario terhadap dua jenis *dataset*, yaitu *dataset* yang berisi citra yang tidak dilakukan *cropping* dan *dataset* yang berisi citra yang dilakukan *cropping* pada bagian wajahnya. Dari hasil pengujian sistem didapatkan bahwa sistem yang dirancang memberikan hasil rata-rata keakuratan terbaik untuk sistem yang diujikan menggunakan *10-fold cross validation* yaitu sebesar 78,92% untuk citra dalam *dataset* yang dilakukan *cropping* dan 58,96% untuk citra dalam *dataset* yang tidak dilakukan *cropping*.

Kata kunci: pengenalan *gender*, klasifikasi *nearest neighbor*, *cross validation*.

1. PENDAHULUAN

Sistem pengenalan *gender* merupakan sebuah sistem yang bertujuan untuk mengklasifikasikan seseorang sebagai pria atau wanita. Sistem pengenalan *gender* yang dirancang pada tulisan ini dibuat dengan menggunakan metode klasifikasi *nearest neighbor* yang memanfaatkan citra wajah manusia, terdiri atas tahap pemrosesan data

awal yaitu *cropping* citra, konversi citra RGB ke *grayscale*, *histogram equalization*, dan *resizing* citra. Tahap selanjutnya adalah metode klasifikasi *nearest neighbor* dengan menggunakan perhitungan jarak Euclidean.

Beberapa sistem pengenalan *gender* telah dibuat dengan menerapkan metode yang berbeda-beda. Cottrell [1] dengan menerapkan metode *neural network* menghasilkan rata-rata keakuratan sebesar 63% dengan *dataset* berjumlah 64 buah. Yen [2] dengan menerapkan metode yang sama dengan Cottrell yaitu *neural network* dengan total *dataset* yang lebih banyak yaitu sebesar 1400 buah menghasilkan rata-rata keakuratan sebesar 90%. Lyons [3] menerapkan metode Gabor wavelet dengan menggunakan PCA dan LDA menghasilkan keakuratan sebesar 91,3%. Castrillon-Santana dan Vuong [4] menggunakan PCA dan *Support Vector Machine* sebagai metode pengenalannya menghasilkan rata-rata keakuratan sebesar 70% - 75%.

Tulisan ini berisi hal-hal sebagai berikut. Pada bagian 2 dari tulisan ini dibahas mengenai klasifikasi *nearest neighbor*, bagian 3 dibahas mengenai deskripsi sistem, bagian 4 berisi eksperimen, dan bagian 5 berisi kesimpulan dari tulisan ini.

2. KLASIFIKASI NEAREST NEIGHBOR

Nearest neighbor adalah metode klasifikasi yang memilih objek latih yang memiliki sifat ketetanggaan (*neighborhood*) yang paling dekat [5]. Dalam klasifikasi *nearest neighbor* untuk mengetahui kedekatan dari obyek yang diteliti maka dihitung jarak antara dua buah obyek tersebut. Pada skripsi ini, digunakan persamaan *Euclidean distance* dalam klasifikasi *nearest neighbor*. Persamaan untuk menghitung *Euclidean distance* dapat dilihat pada persamaan (1) [5]. *Euclidean distance* menghitung jarak antara dua buah data yang dibandingkan. Nilai d yang semakin kecil menandakan data pertama dan kedua semakin menyerupai satu sama lain.

Pengenalan Gender Memanfaatkan Wajah Manusia Dengan Menggunakan Metode Klasifikasi Nearest Neighbor

Ardilla Ayu Dewanti Ridwan, Ivanna K. Timotius, Iwan Setyawan

$$d(\mathbf{g}, \mathbf{h}) = \|\mathbf{g} - \mathbf{h}\| = \sqrt{\sum_{i=1}^r (g_i - h_i)^2} \tag{1}$$

$d(\mathbf{g}, \mathbf{h})$ = jarak Euclidean dari vektor \mathbf{g} ke \mathbf{h}

\mathbf{g} = vektor pertama

\mathbf{h} = vektor kedua

g_i = nilai pada baris ke- i

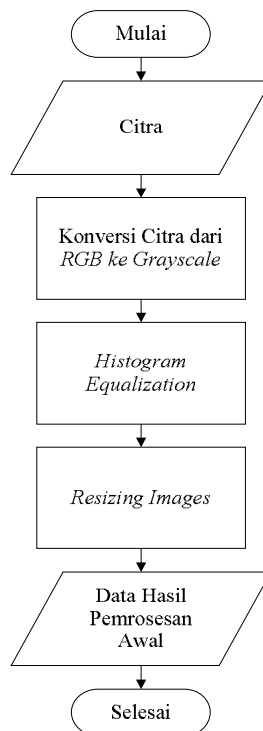
h_i = nilai pada baris ke- i

r = jumlah baris dari vektor pertama atau kedua

3. DESKRIPSI SISTEM

3.1. Pemrosesan Data Awal

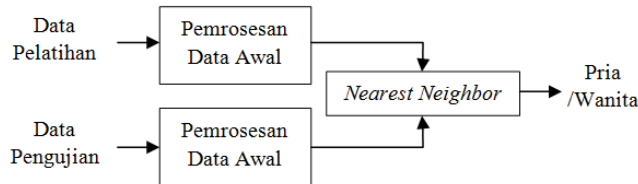
Pada bagian pemrosesan data awal untuk sistem pengenalan *gender* ini terdiri atas proses konversi citra RGB ke *grayscale*, *histogram equalization*, dan *resizing* citra. Gambar 1 merupakan *flowchart* untuk sistem pengenalan *gender* yang dibuat pada tulisan ini.



Gambar 1. *Flowchart* Sistem Pengenalan Gender

3.2. Gambaran Sistem

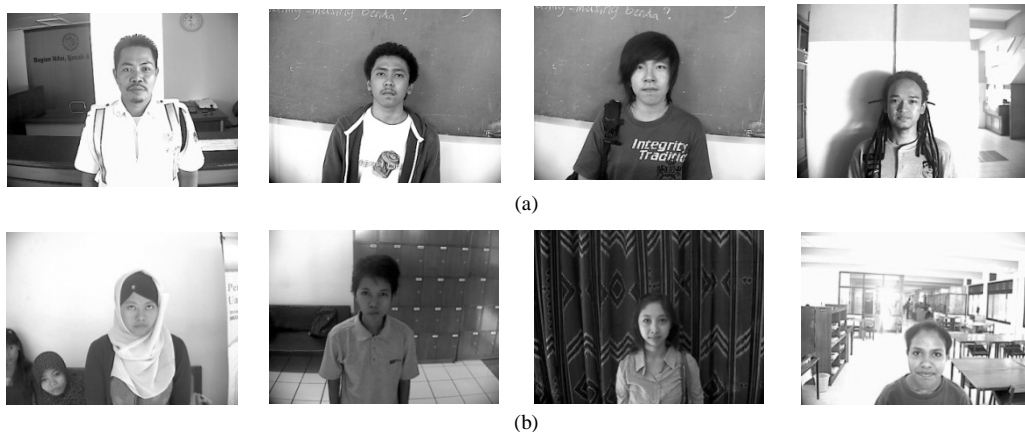
Pada tulisan ini, dirancang sebuah sistem pengenalan *gender* dengan menggunakan metode klasifikasi *nearest neighbor*. Data pelatihan dan data pengujian diproses menggunakan pemrosesan data awal seperti dijelaskan pada bagian 3.1. Gambar 2 merupakan blok diagram sistem pengenalan *gender* pada tulisan ini.



Gambar 2. Blok Diagram Sistem Pengenalan *Gender*

4. EKSPERIMEN

Dataset untuk sistem pengenalan *gender* pada tulisan ini berupa citra wajah yang didapatkan dari 500 orang responden yang terdiri dari 250 pria dan 250 wanita yang diambil gambar wajahnya dengan ekspresi netral. Pembuatan *dataset* dilakukan dalam kondisi pencahayaan yang cukup. Data yang digunakan dalam perancangan adalah citra yang diambil menggunakan *web camera* yang dihubungkan pada *port* USB komputer. Ukuran awal citra yang didapatkan adalah sebesar 640×480 piksel seperti terlihat pada Gambar 3. Untuk jenis *dataset* yang pertama, citra ini kemudian di-*crop* pada bagian wajah menggunakan perangkat lunak *editor* gambar, sehingga dihasilkan citra yang berukuran 96×128 piksel seperti terlihat pada Gambar 4. Citra yang telah dilakukan *cropping* maupun tidak, akan melewati tahap pemrosesan data awal, kemudian citra yang telah dilakukan proses tersebut akan diubah menjadi berukuran 32× 24 piksel. Contoh citra dalam *dataset* adalah sebagai berikut.



Gambar 3. Contoh Citra yang Tidak Dilakukan *Cropping*

Pengenalan Gender Memanfaatkan Wajah Manusia Dengan Menggunakan Metode Klasifikasi Nearest Neighbor

Ardilla Ayu Dewanti Ridwan, Ivanna K. Timotius, Iwan Setyawan



Gambar 3. Contoh Citra yang Dilakukan *Cropping*

Pada tulisan ini, citra diubah menjadi vektor kolom terlebih dahulu sebelum dilakukan proses pengenalan *gender* menggunakan metode klasifikasi *nearest neighbor* dengan jarak Euclidean.

Untuk menghitung keakuratan dalam sistem pengenalan *gender* ini digunakan metode *k-fold cross validation* [6]. Rata-rata keakuratan dari sistem pengenalan *gender* pada tulisan ini ditunjukkan pada Tabel 1. Dari Tabel 1 diperoleh rata-rata keakuratan terbaik yang dihasilkan oleh sistem yang diuji menggunakan *10-fold cross validation* yaitu sebesar 78,92% untuk citra yang dilakukan *cropping* dan 58,96% untuk citra yang tidak dilakukan *cropping*. Dari hasil eksperimen tersebut terlihat bahwa jumlah data pelatihan berpengaruh terhadap performa sistem. Sistem dengan data pelatihan yang lebih banyak menghasilkan rata-rata keakuratan yang lebih baik. Selain itu, citra yang dilakukan *cropping* menghasilkan rata-rata keakuratan yang lebih baik dibandingkan dengan yang tidak dilakukan *cropping*. Hal ini terjadi karena ekstraksi fitur yang dilakukan oleh sistem tidak hanya pada bagian wajah, namun juga pada *background* dari citra tersebut.

Tabel 1. Hasil Rata-Rata Keakuratan Sistem Pengenalan *Gender*

Jenis Dataset	2-fold	5-fold	10-fold
<i>Cropped</i>	78,46%	78,54%	78,92%
<i>Non-Cropped</i>	58,68%	59,08%	58,96%

5. KESIMPULAN

Dari hasil eksperimen yang telah dilakukan, metode klasifikasi *nearest neighbor* untuk sistem pengenalan *gender* ini menghasilkan rata-rata keakuratan terbaik untuk sistem yang diujikan dengan menggunakan *10-fold cross validation* yaitu sebesar 78,92% untuk citra yang dilakukan *cropping* dan 58,96% untuk citra yang tidak dilakukan *cropping*. Citra yang dilakukan *cropping* menghasilkan rata-rata keakuratan yang lebih baik dibandingkan dengan yang tidak dilakukan *cropping*.

DAFTAR PUSTAKA

- [1] G. Cottrell dan M. Fleming, "Face Recognition Using Unsupervised Feature Extraction," *International Conference on Neural Network*, pp. 322–325, 1990.
- [2] S. Yen, P. Sajda, dan L. Finkel, "Comparison Of Gender Recognition by Pdp and Radial Basis Function Networks," *The Neurobiology of Computation*, pp. 433–438, 1994.
- [3] M. Lyons, J. Budynek, A. Plante, dan S. Akamatsu, "Classifying Facial Attributes using a 2-D Gabor Wavelet Representation and Discriminant Analysis," *Proceedings International Conference on Automatic Face and Gesture Recognition*, IEEE, Grenoble, France, pp. 202–207, 2000.
- [4] M. Castrillon-Santana dan Q.C. Vuong, "an Analysis of Automatic Gender Classification," *Proceedings of Conference on Progress in Pattern Recognition, Image Analysis and Applications*, pp. 271–280, 2007.
- [5] D. S. Dhaliwal, P. S. Sandhu, dan S. N. Panda, "Enhanced K-Nearest Neighbor Algorithm," *Journal World Academy of Science, Engineering and Technology*, vol. 73, pp. 681–685, 2011.
- [6] A. F. Putra, "Pengenalan Wajah dengan Menggunakan Metode Kernel Fisher's Discriminant Analysis (KFDA) dan Nearest Neighbor (NN)," FTEK, UKSW, Salatiga, 2011 (Skripsi untuk meraih gelar sarjana S1 Teknik Elektro).