

IMPLEMENTASI 2D GABOR WAVELET DAN JARINGAN SARAF TIRUAN LEARNING VECTOR QUANTIZATION UNTUK DETEKSI PORNOGRAFI PADA CITRA DIGITAL

Bagus Radityo Wibowo¹, Tjokorda Agung Budi Wirayuda², Retno Novi Dayawati³

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Pada Tugas Akhir ini dilakukan implementasi metode 2D Gabor Wavelet sebagai metode ekstraksi ciri dan Jaringan Saraf Tiruan Learning Vector Quantization (LVQ) sebagai metode pengklasifikasian citranya. Citra akan diklasifikasikan ke dalam 2 kelas yaitu citra porno atau citra bukan porno. Data yang digunakan pada saat pembangunan knowledge adalah data citra berukuran 40x40 pixel berjumlah 90 data latih, 30 data validasi dan 40 data uji. Dengan menggunakan metode ini, akurasi pada saat pembangunan knowledge adalah 100% untuk data latih, 100% untuk data validasi dan 97.5% untuk data uji. Akurasi terbaik pada saat melakukan deteksi gambar porno sebesar 95,5% dengan menggunakan parameter shifting pixel 5. Sedangkan akurasi terbaik pada saat melakukan deteksi gambar porno dan gambar tidak porno sebesar 83,53% dengan menggunakan parameter shifting pixel 10. Total jumlah gambar pada saat deteksi adalah 45 gambar porno dan 40 gambar bukan porno.

2D Gabor Wavelet adalah ekstraksi ciri yang memiliki kemampuan dalam membedakan berbagai tekstur didasarkan atas berbagai frekuensi dan orientasi spasial. Sedangkan Learning Vector Quantization merupakan JST dengan pembelajaran yang terawasi. Kelebihan LVQ adalah hanya memiliki 1 hidden layer dan 1 linear layer sehingga proses learning / pembangunan knowledge serta proses pendeteksian yang cepat. Akurasi tertinggi pada saat pembangunan knowledge diperoleh dengan parameter Gabor Wavelet frekuensi 2, 3, 4, 5, 6 orientasi 300 serta ukuran kernel 7 dan parameter LVQ adalah 4 hidden neuron serta learning rate 0.04. Pembangunan knowledge dilakukan dengan menggunakan 90 data latih, 30 data validasi dan 40 data uji.

Kata Kunci : klasifikasi, citra porno, Gabor Wavelet, Learning Vector

Abstract

In my final project, implemented 2D Gabor Wavelet as a method of feature extraction and Artificial Neural Network Learning Vector Quantization (LVQ) as a method of images recognition. The images will be classified into two classes which are pornographic images or non-pornographic images. The images size used during the construction of knowledge is a 40x40 pixel with 90 training data, 30 validation data and 40 test data. By using this method, the accuracy at of construction the knowledge is 100% for training data, 100% for validation data and 97.5% for testing data. The best system accuration of detection pornographic images using shifting pixel 5 reaches 95,5%. But the best system accuration of detection pornographic and non pornographic images using shifting pixel 10 reaches 83,53%.

2D Gabor Wavelet is a feature extraction that has the ability to differentiate various textures based on various frequency and spatial orientation. Learning Vector Quantization is a Artificial Neuron Network with supervised learning. LVQ is only has 1 hidden layer and 1 linear layer so that rapid process of learning / knowledge development and detection process. Highest accuracy at the time of construction of knowledge obtained with Gabor Wavelet frequency parameters 2, 3, 4, 5, 6 orientations 300, kernel size 7 and the parameters of LVQ are 4 hidden neurons and learning rate 0.04. Knowledge development done by using 90 training data, 30 data validation and 40 test data.

Keywords : classification pornographic images, Gabor Wavelet, Learning

1. Pendahuluan

1.1 Latar Belakang Masalah

Pada era komputerisasi pornografi menjadi dampak negatif yang muncul seiring dengan perkembangan teknologi dan informasi yang cepat. Pornografi beredar dan berkembang dengan cepat setiap harinya dan harus diimbangi dengan penanggulangan yang cepat pula. Definisi pornografi berdasarkan Kamus Besar Bahasa Indonesia adalah penggambaran tingkah laku secara erotis dengan lukisan atau tulisan untuk membangkitkan nafsu birahi. Pornografi sering disalahgunakan dan dinikmati oleh pihak yang belum cukup umur (dibawah 17 tahun). Ini memicu gairah seksual yang tinggi yang tidak diimbangi dengan kedewasaan serta iman yang kuat kemudian berimbas pada beberapa tindak kejahatan. Salah satu bentuk pornografi yang sering diakses adalah gambar porno. Keberadaan unsur pornografi dalam suatu gambar/citra dapat dideteksi dengan metode tertentu diantaranya dengan memanfaatkan *2D Gabor Wavelet* sebagai ekstraksi ciri dari masing – masing pola gambar serta jaringan saraf tiruan sebagai klasifikasinya.

2D Gabor Wavelet merupakan suatu teknik ekstraksi ciri. Ekstraksi ciri bertujuan untuk mendapatkan informasi dan tekstur penting dari sebuah citra. Teknik ekstraksi ciri ini digunakan untuk mengekstraksi ciri pada citra yang ternormalisasi. Teknik ini mendapatkan ciri dari suatu citra melalui kombinasi frekuensi dan orientasi yang bervariasi. Teknik ini mengekstraksi ciri suatu citra baik citra data latih berupa ciri porno dan citra data uji. Hasil ekstraksi ciri tersebut kemudian digunakan sebagai *input* dalam jaringan saraf tiruan.

Jaringan saraf tiruan merupakan bagian dari *Artificial Intelligence*. Kehandalan jaringan saraf tiruan atau yang lebih dikenal dengan JST adalah pengenalan pola, klasifikasi, peramalan, dll. Pada Tugas Akhir ini digunakan jaringan saraf tiruan *LVQ* untuk mendeteksi adanya unsur pornografi pada citra digital. Pemilihan metode *Learning Vector Quantization (LVQ)* dikarenakan perkembangan pornografi yang cepat sehingga memerlukan proses perhitungan pada saat *learning* dan pendeteksian yang cepat. Pada tahap identifikasi pola, *LVQ* mengelompokkannya ke dalam pola yang mempunyai jarak *Euclidian* paling dekat. Proses pendeteksian dilakukan dengan memecah citra menjadi blok berukuran kecil. Blok tersebut kemudian dilakukan *preprocessing*. Tahap selanjutnya dilakukan proses ekstraksi ciri *gabor wavelet* yang kemudian berdasarkan ciri tersebut dapat diklasifikasikan apakah blok tersebut porno atau tidak.

Dengan kombinasi metode *Gabor Wavelet* dan JST *LVQ*, diharapkan sistem deteksi pornografi akan dapat mengenali citra porno secara akurat dengan waktu yang relatif cepat.

1.2 Perumusan masalah

Berdasarkan latar belakang yang sudah dipaparkan di atas, maka rumusan masalah yang dapat dirumuskan diantaranya adalah sebagai berikut :

1. Bagaimana mendapatkan ciri (citra porno dan citra bukan porno) yang baik menggunakan *2D Gabor Wavelet* ?
2. Bagaimana membangun *knowledge JST LVQ* yang baik untuk mengenali citra porno ?
3. Bagaimana menggunakan *knowledge JST LVQ* terbaik untuk mendeteksi citra dengan unsur porno?
4. Bagaimana keterkaitan antara ciri citra, pembangunan *knowledge*, serta penggunaan *knowledge* tersebut dalam sistem deteksi gambar porno dalam hal akurasi serta waktu proses?

1.3 Tujuan

Tujuan dari Tugas Akhir ini yaitu :

1. Melakukan implementasi sistem deteksi pornografi pada citra menggunakan *2D Gabor Wavelet* dan Jaringan Saraf Tiruan *Learning Vector Quantization*.
2. Menganalisis hasil implementasi sistem dari sisi akurasi dan waktu deteksi citra.

1.4 Batasan masalah

Dalam tugas akhir ini diperlukan batasan masalah agar mencegah luasnya ruang lingkup permasalahan yang harus ditangani. Batasan-batasannya adalah:

1. Definisi pornografi pada tugas akhir ini adalah payudara wanita dewasa yang terlihat putingnya.
2. Pada tahap awal, citra dilakukan *preprocessing* terlebih dahulu diantaranya penghilangan *noise*, *grayscale* dan *image enhancement*.
3. Deteksi warna kulit yang digunakan adalah deteksi warna kulit *peer model*.
4. Ukuran pemecahan blok yang digunakan adalah *40x40 pixel*.
5. Sistem pada Tugas Akhir ini tidak bersifat *real time* dengan menggunakan data secara tidak langsung.
6. *Input* citra yang digunakan berekstensi .JPEG.
7. Resolusi maksimal citra yang dideteksi adalah *800 x 600* dan resolusi minimalnya adalah *40 x 40 pixel*.

1.5 Metodologi penyelesaian masalah

Langkah-langkah yang dilakukan dalam penyelesaian Tugas Akhir ini yaitu :

1. Pengumpulan data

Data yang dikumpulkan pada Tugas Akhir ini berupa data gambar dengan format JPEG. Data terbagi menjadi empat bagian yaitu data latih, data validasi, data uji serta data yang akan dilakukan deteksi. Keempat data ini didapatkan dari internet. Data yang digunakan dalam pembangunan *knowledge* adalah data latih, data validasi, serta data uji. Proses pengumpulan ketiga data tersebut dilakukan dengan melakukan *cropping* secara manual berukuran *40x40 pixel*

pada bagian tertentu. Untuk citra kelas porno, dilakukan *cropping* secara manual pada bagian puting payudara sedangkan untuk kelas bukan porno dilakukan *cropping* secara manual pada anggota tubuh contoh mata, hidung, mulut, telinga.

2. Studi Literatur

Pada tahap studi literatur ini, studi literatur dilakukan dengan mengumpulkan, membaca, serta memahami metode yang terkait yaitu JST *LVQ* dan *gabor wavelet*. Selain kedua metode itu, studi literatur juga mempelajari beberapa metode tambahan seperti deteksi warna kulit untuk meningkatkan keakuratan sistem deteksi. Studi literatur juga mempelajari beberapa pengertian dan pembahasan mengenai definisi, undang-undang, dan aneka ragam pornografi. Sumber dan referensi didapat dari buku, undang-undang, karya ilmiah, dan internet

3. Perancangan sistem dan implementasi

Perancangan sistem disini terdapat beberapa tahap :

a. Tahap *preprocessing*

Pada tahap *preprocessing* ini, citra yang digunakan untuk pembangunan *knowledge* dilakukan *grayscale* terlebih dahulu. *Grayscale* ini bertujuan untuk mengubah citra menjadi derajat keabuan. Citra hasil *grayscale* kemudian dilakukan penghilangan *noise* menggunakan *median filter*. Penghilangan *noise* dilakukan untuk menghilangkan *noise* yang bisa mempengaruhi hasil pembangunan *knowledge*. Tahap selanjutnya adalah dilakukan proses *image enhancement* dengan menggunakan ekualisasi histogram. Citra keluaran dari ekualisasi histogram ini digunakan untuk proses pada tahap selanjutnya.

b. Tahap ekstraksi ciri *gabor wavelet*

Citra masukan pada proses ini adalah citra hasil ekualisasi histogram pada tahap *preprocessing*. Proses pembangunan ciri *gabor wavelet* dilakukan dengan menggunakan kombinasi parameter frekuensi dan orientasinya. Panjang matriks ciri yang dihasilkan adalah perkalian antara frekuensi dan orientasi yang digunakan. Sebagai contoh jika frekuensi yang digunakan adalah 2,3,4 dan orientasi adalah 30, 60, 90, maka panjang matriks ciri yang dihasilkan adalah 9 buah. Proses ekstraksi ciri *gabor wavelet* adalah dengan melakukan konvolusi antara matriks citra input dengan matriks kernel *imaginer* dan kernel *real*. Setelah dilakukan konvolusi, kedua hasil tersebut kemudian dilakukan di-*magnitude*-kan. Dari hasil *magnitude* ini, kemudian dicari cirinya yaitu berupa ciri *mean*, ciri entropi, serta ciri gabungan.

c. Tahap pembangunan *knowledge JST LVQ*

Pada tahap ini, terdapat beberapa tahap diantaranya :

- Tahap pencarian dan observasi parameter *gabor wavelet* terbaik serta ciri terbaik. Pada tahap ini digunakan arsitektur dan parameter JST *LVQ* yang telah ditentukan terlebih dahulu.

- Setelah didapatkan ciri terbaik dan parameter terbaik dari *gabor wavelet*, dilakukan pencarian parameter dan arsitektur JST *LVQ* terbaik.
- Pelatihan JST *LVQ* dengan menggunakan semua parameter terbaik dan menyimpan hasil bobot pelatihan.
- Penentuan *threshold* jarak dengan menggunakan kembali data latih untuk menentukan jarak terjauh data pada kelas porno.
- Menyimpan semua *knowledge* yang ada untuk kemudian digunakan pada proses deteksi citra

d. Tahap deteksi citra

Pada tahap deteksi citra, *knowledge* yang sudah didapatkan pada tahap sebelumnya digunakan untuk melakukan deteksi citra apakah terdapat unsur pornografi atau tidak. Pada tahap ini citra masukan berukuran minimal 40x40 pixel dan maximal berukuran 800x600 pixel. Citra tersebut pertama kali dilakukan proses *resize* dengan menggunakan *rescale factor*. Proses ini bertujuan untuk mengantisipasi blok porno yang melebihi ukuran blok 40x40 pixel. Citra tersebut kemudian dilakukan pemecahan per blok dengan ukuran blok 40x40 pixel. Dengan cara ini, maka sistem akan mendeteksi citra tersebut per blok sehingga akan dapat mendeteksi unsur porno secara akurat. Pemecahan blok citra ini dilakukan dengan menggunakan mekanisme pergeseran pada saat pengambilan blok/ *shifting pixel*. Blok hasil pemecahan tersebut kemudian dilakukan *preprocessing*, ekstraksi ciri, kemudian masuk ke tahap klasifikasi menggunakan *knowledge* yang telah dibangun. Hasil klasifikasi blok tersebut kemudian dilakukan validasi menggunakan deteksi warna kulit jika blok tersebut diklasifikasikan ke dalam kelas porno. Apabila disekitar blok tersebut terdeteksi adanya warna kulit, maka blok tersebut dapat dikatakan sebagai blok porno dan dilakukan proses sensor. Deteksi warna kulit ini merupakan pengembangan dari metode yang digunakan. Proses ini digunakan sebagai penunjang sistem dalam hal akurasi.

e. Implementasi

Implementasi yang digunakan menggunakan bahasa pemrograman *Matlab2009B*. Pemilihan bahasa pemrograman ini dikarenakan tersedianya fungsi-fungsi yang memudahkan pada saat implementasi sebagai contoh fungsi *rgb2gray*, *medfilt2*, *histeq*, *imcrop*, dan sebagainya.

4. Pengujian sistem dan analisis

Skema pengujian dilakukan untuk mendapatkan parameter-parameter terbaik pada *gabor wavelet* dan JST *LVQ*. Parameter tersebut kemudian digunakan pada saat pelatihan. Proses pengujian parameter *gabor wavelet* dan JST *LVQ* dilakukan pada saat pembangunan *knowledge* pada sistem. Selain pengujian parameter-parameter di atas, skema pengujian lain yang dilakukan adalah skema pengujian parameter yang digunakan pada saat tahap deteksi citra. Pada Tugas Akhir ini, pengujian yang dilakukan pada saat deteksi adalah pengujian parameter *shifting pixel* dan pengujian keberadaan deteksi warna kulit. Pengujian parameter *shifting pixel* ini bertujuan untuk mencari parameter

shifting pixel dengan akurasi yang tinggi akan tetapi waktu proses yang relatif cepat. Sedangkan pengujian keberadaan deteksi warna kulit bertujuan untuk melihat akurasi sistem dengan atau tanpa deteksi warna kulit.

5. Pembuatan buku dan laporan

Tahap terakhir pada Tugas Akhir ini adalah pembuatan laporan tugas akhir, yang meliputi dasar teori, perancangan sistem, pengujian dan hasil analisis. Pembuatan laporan bertujuan untuk mendokumentasikan semua tahap yang telah dilakukan.



5. Kesimpulan dan Saran

5.1 Kesimpulan

Berdasarkan hasil implementasi, pengujian dan analisis hasil yang telah dilakukan dapat diambil kesimpulan yaitu :

1. Ciri terbaik yang memberikan akurasi terbaik adalah ciri *mean*. Ciri ini memberikan akurasi lebih baik daripada ciri *entropi* dan ciri gabungan.
2. Pada saat pembangunan *knowledge*, penggunaan frekuensi yang lebih banyak menghasilkan *knowledge* yang relatif lebih baik dari sisi akurasi. Selain frekuensi, ukuran kernel juga mempengaruhi *knowledge* yang dibangun. Semakin kecil ukuran kernel, *knowledge* yang dibangun justru relatif lebih baik.
3. Jumlah *hidden neuron* yang digunakan pada arsitektur *LVQ* adalah *hidden neuron* terkecil (*hidden neuron* 4) agar waktu pada saat deteksi cepat.
4. Pada saat proses pendeteksian citra, parameter *shifting pixel* sangat mempengaruhi akurasi deteksi dan waktu pendeteksian. Nilai *shifting pixel* berbanding lurus dengan akurasi deteksi gambar tidak porno dan berbanding terbalik dengan akurasi deteksi gambar porno serta waktu proses pendeteksian citra.
5. Pada saat melakukan deteksi yang berfokus **hanya** pada kelas porno saja, maka gunakan *shifting pixel* yang kecil untuk mendapatkan hasil yang akurat. Sedangkan untuk deteksi yang berfokus pada kedua kelasnya, gunakan *shifting pixel* yang teruji dan mampu mengenali gambar pada kedua kelasnya dengan akurasi/total gambar benar lebih besar.

5.2 Saran

Berdasarkan penelitian yang telah dilakukan muncul beberapa saran untuk pengembangan lebih lanjut :

1. Batasan pornografi diperluas lagi tidak hanya berkisar pada puting payudara wanita dewasa saja.
2. Modifikasi proses deteksi warna kulit sehingga warna kulit serta pencahayaan relatif tidak mempengaruhi akurasi sistem dan relatif tahan terhadap *noise*.
3. Modifikasi proses deteksi sehingga waktu proses lebih cepat akan tetapi tidak terlalu memakan banyak *resource*.
4. Pengembangan sistem deteksi menjadi *real time detection* dengan menggunakan bahasa pemrograman yang lebih cepat daripada Matlab serta memodifikasi algoritma deteksi porno sehingga waktunya cepat.

Daftar Pustaka

- [1] Agung BW, Tjokorda. Rudy Hermanto, I Gede. Novi D, Retno. 2009. *Pengenalan Huruf Bali Menggunakan Metode Modified Direction Feature (MDF) dan Learning Vector Quantization (LVQ)*. <http://yudiagusta.files.wordpress.com/2009/11/007-012-knsi09-002-pengenalan-huruf-bali-menggunakan-metode-modified-direction-feature-mdf-dan-learning-vector-quantization-lvq.pdf>. Didownload pada tanggal 29 Maret 2010.
- [2] Bandung Putra, Andika. Iwut T, Iwan. Haryatno. 2007. *Speech Recognition Menggunakan Gabor Wavelet dan Jaringan Syaraf Tiruan Backpropagation untuk Sistem Keamanan Berbasis Suara*. <http://yudiagusta.files.wordpress.com/2009/11/120-129-snsi07-022-speech-recognition-menggunakan-gabor-wavelet-dan-jaringan-saraf-tiruan-backpropagation-untuk-sistem-keamanan-berbasis-suara.pdf>. Didownload pada tanggal 24 Maret 2010
- [3] Biehl, Michael. Hammer, Barbara. Schneider, Petra. *Matrix Learning in Learning Vector Quantization*. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.94.486&rep=rep1&type=pdf>. Didownload pada tanggal 12 Mei 2010
- [4] Citraningrum, Nancy. 2010 *Wavelet Gabor*. http://www.itelkom.ac.id/library/index.php?view=article&catid=15:pemrosesan-sinyal&id=663:wavelet-gabor&option=com_content&Itemid=15. Didownload pada tanggal 23 Februari 2011
- [5] Effendy, Nazrul. Imanto, Rifqi. P.T, Ayodya. *Deteksi Pornografi Pada Citra Digital Menggunakan Pengolahan Citra dan Jaringan Syaraf Tiruan*. <http://achmad.blog.undip.ac.id/files/2010/02/deteksi-pornografi.pdf>. Didownload pada tanggal 23 Februari 2010.
- [6] Eni Puspitawati, Diah. 2008. *Pengenalan Iris Mata dengan Menggunakan Metode Wavelet dan Jaringan Saraf Tiruan Radial Basis Function*. Fakultas Informatika. Institut Teknologi Telkom Bandung.
- [7] Inwijayati. Prihandoko. Bertalya. *Identifikasi Ciri Kain Menggunakan Fitur Tekstur dan Gray Level Difference Method*. <http://openstorage.gunadarma.ac.id/~mwiryana/KOMMIT/per-artikel/01-02-007-Identifikasi%5BIrwijayati%5D.pdf>. Didownload pada tanggal 6 Februari 2011.
- [8] Isnanto, R. Rizal. 2009. *Identifikasi Iris Mata Menggunakan Tapis Gabor Wavelet dan Jaringan Syaraf Tiruan Learning Vector Quantization (LVQ)*. http://eprints.undip.ac.id/5374/1/Identifikasi_Iris.pdf. Didownload pada tanggal 24 Maret 2010.
- [9] Kepenekci, Burcu. 2001. *Face Recognition Using Gabor Wavelet Transform*. www.eee.metu.edu.tr/~alatan/PAPER/MSburcu.pdf. Didownload pada tanggal 26 Maret 2010.
- [10] Leung, K. Ming. 2009. *Learning Vector Quantization*. <http://cis.poly.edu/~mleung/CS6673/s09/LVQ.pdf>. Didownload pada tanggal 29 Maret 2010.

- [11] Munir, Renaldi. 2004. *Pengolahan Citra Digital dengan Pendekatan Algoritmik*. Bandung : Informatika.
- [12] Pramadihanto, Dadet. Sukaridhoto, Sritrusta. Fatoni A, Muhammad. Cahyadi HP, Rudy. *Sistem Pendeteksi Gambar Porno*. <http://lecturer.eepis-its.edu/~dphoto/publikasi/porn-detect2007/Paper%20Sistem%20Deteksi%20Gambar%20Porno.pdf>. Didownload pada tanggal 20 Januari 2011.
- [13] Ranadhi, Djalu. Indarto, Wawan. Hidayat, Taufiq. 2006. *Implementasi Learning Vector Quantization (Lvq) Untuk Pengenal Pola Sidik Jari Pada System Informasi Narapidana LP Wirogunan*. <http://journal.uii.ac.id/index.php/media-informatika/article/viewFile/121/82>. Didownload pada tanggal 23 Februari 2010.
- [14] Rizal, Achmad. *Modul 6 : Analisis Tekstur & Ekstraksi Ciri.pdf*. Departemen Teknik Elektro. Institut Teknologi Telkom Bandung.
- [15] Sukmana, Indriani. 2008. *Analisis dan Implementasi Direct Linear Discriminant Analysis (DLDA) dan Jaringan Syaraf Tiruan Learning Vector Quantization (LVQ) pada Pengenalan Iris Mata Manusia* Fakultas Informatika. Institut Teknologi Telkom Bandung.
- [16] Suyanto. 2008. *Soft Computing Membangun Mesin Ber-IQ Tinggi*. Informatika : Bandung.
- [17] Syawaluddin, Mochammad Taufik. 2010. *Pengenalan Plat Nomor Otomatis Menggunakan Principal Component Analysis (PCA) dan Learning Vector Quantization (LVQ)*. Fakultas Informatika. Institut Teknologi Telkom Bandung.
- [18] Tim Undang-Undang Pornografi. 2008. *Undang – Undang Pornografi*. <http://www.bpkp.go.id/unit/hukum/uu/2008/44-08.pdf>. Didownload pada tanggal 25 Desember 2010.
- [19] Wijaya, Marvin Ch. dan Agus Prijono. *Pengolahan Citra Digital Menggunakan Matlab*. 2007. Informatika: Bandung.

Telkom
University