

**PENGENALAN POLA JARINGAN NORMAL DAN
JARINGAN BERMIKROKALSIFIKASI PADA CITRA
MAMMOGRAFI DIGITAL MENGGUNAKAN
SUPPORT VECTOR MACHINES (SVM)**

Indah Susilawati

Abstract: Features can be extracted from images based on pixel values. A classifier uses these image features to classify the image by the mean of pattern recognition.

In this research, six features are extracted from mammogram image, i.e. normal mammogram and mammogram with microcalcification, based on pixel values. Because of the presence of microcalcification in mammograms can be an indicator of breast cancer, a pattern recognition of these two type of mammograms is important. In this research, Support Vector Machines (SVM) is used to recognize the pattern and classify the images based on six-feature input. The proposed method is developed and evaluated using 98 samples. All samples are preprocessed using linear scalling function and tophat transformation.

The result show that the SVM best performance in recognition the microcalcification is 100%, using linear kernel and regularization parameter $C = 100$.

Keywords: normal, microcalcification, features, linear scalling function, tophat, SVM

Citra dapat dianggap sebagai suatu sumber informasi dengan simbol-simbol berupa nilai-nilai piksel yang mungkin pada citra tersebut. Misalnya, untuk citra aras keabuan maka ada 256 simbol yang mungkin yaitu nilai piksel 0 hingga 255. Dengan menggunakan asumsi ini maka terdapat banyak ciri yang dapat diekstrak dari sebuah citra, termasuk citra mamografi, berdasarkan nilai piksel citra.

Ciri-ciri yang diekstrak dari suatu objek dapat digunakan untuk mengenali jenis objek yang bersangkutan. Dalam penelitian ini akan dilakukan pengenalan pola jaringan normal dan jaringan bermikrokalsifikasi pada citra mamografi menggunakan *Support Vector Machines* (SVM) dengan masukan ciri-ciri citra yang diekstrak berdasarkan nilai piksel citra. Keberadaan mikrokalsifikasi pada jaringan payudara menjadi tanda awal adanya kanker. Semakin dini mikrokalsifikasi dapat dideteksi maka semakin cepat pula penanganan medis dapat dilakukan. Pada penelitian ini, jaringan normal dan jaringan bermikrokalsifikasi akan dikategorikan sebagai dua kelas yang berbeda

dan SVM akan digunakan untuk mengenali kedua kelas tersebut.

Penelitian ini bertujuan untuk melakukan pengenalan pola jaringan normal dan jaringan bermikrokalsifikasi pada citra mamografi menggunakan *Support Vector Machines* (SVM) dengan input ciri-ciri citra mamografi yang diekstrak berdasarkan nilai pikselnya. Dalam penelitian ini akan dilakukan prapengolahan citra menggunakan fungsi penskalaan linear (*linear scalling function*) dan transformasi *tophat*. Pada proses selanjutnya akan dilakukan pengenalan pola menggunakan SVM dengan masukan ciri-ciri yang telah diekstrak. Hasil pengenalan pola dengan kedua jenis prapengolahan tersebut selanjutnya akan dibandingkan untuk mengetahui tingkat pengenalan tertinggi yang dapat dicapai.

Penelitian ini dilakukan dengan bantuan Matlab 7.1, prapengolahan citra menggunakan fungsi penskalaan linear (*linear scalling function*) dan transformasi *tophat*, serta ciri-ciri citra yang diekstrak adalah rerata, varians, standar deviasi, entropi, beda nilai aras maksimum dan minimum (selanjutnya disebut *bedaan I*), dan beda nilai aras maksimum dengan nilai rerata (selanjutnya

Indah Susilawati (susilawati.indah@yahoo.co.id) adalah dosen di Program Studi Teknik Elektro, Fakultas Teknik dan Ilmu Komputer, Universitas Mercu Buana, Yogyakarta

disebut bedaan2). Data citra mamografi diperoleh dari basis data MIAS (*Mammographic Image Analysis Society*).

LANDASAN TEORI
Perentangan Kontras

Perentangan kontras (*contrast stretching*) termasuk dalam kategori operasi titik dalam proses peningkatan citra (*image enhancement*). Operasi titik merupakan operasi tanpa memori (*zero memory operations*) yang memetakan suatu aras keabuan tertentu yaitu $u \in [0, 255]$ ke suatu aras keabuan lain.

Citra mamografi seringkali ditemukan sebagai citra dengan kontras rendah. Untuk meningkatkan kontras citra dapat digunakan transformasi menggunakan fungsi penskalaan linear (*linear scaling function*) untuk tiap nilai piksel,

$$P_o = (P_i - c) \times \frac{b - c}{d - c} + a \quad (1)$$

dengan

P_i adalah nilai piksel input,
 P_o adalah nilai piksel output,
 a adalah nilai piksel minimum yang diinginkan,
 b adalah nilai piksel maksimum yang diinginkan,
 c adalah nilai piksel terendah pada citra input,
 d adalah nilai piksel tertinggi pada citra input.

Transformasi Tophat

Transformasi *tophat* merupakan jenis pemrosesan morfologi yang dikembangkan oleh Meyer. Transformasi diperoleh dengan mencari selisih citra hasil *opening* secara morfologi terhadap citra aslinya. Transformasi ini juga dapat digunakan untuk meningkatkan kontras pada citra. Bergantung pada apakah objek merupakan objek yang terang dengan latar belakang gelap atau objek yang gelap dengan latar belakang terang, transformasi *tophat* didefinisikan sebagai berikut.

Untuk objek terang,

$$\begin{aligned} TopHat(A, B) &= A - (A \circ B) \\ &= A - \max_B(\min(A)) \end{aligned} \quad (2a)$$

Untuk objek gelap,

$$\begin{aligned} TopHat(A, B) &= (A \bullet B) - A \\ &= \min_B(\max(A)) - A \end{aligned} \quad (2b)$$

dengan A adalah objek yang akan ditransformasikan, B adalah elemen penstruktur

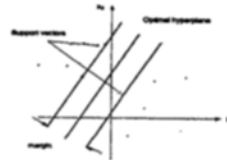
(*structuring element*), notasi \circ menyatakan operasi pembukaan (*opening*) untuk objek terang dengan latar belakang gelap dan notasi \bullet menyatakan operasi pembukaan (*opening*) untuk objek gelap dengan latar belakang terang.

Dalam penelitian ini, transformasi *tophat* yang dilakukan menggunakan elemen penstruktur berbentuk *disk*. Pemilihan elemen penstruktur berbentuk *disk* didasarkan pada kenyataan bahwa mikrokalsifikasi secara visual terlihat sebagai titik-titik kecil yang lebih terang dari daerah sekitarnya. Transformasi bertujuan untuk meningkatkan kontras citra sehingga ciri-ciri yang membedakan antara citra normal dan citra bermikrokalsifikasi akan lebih menonjol.

Support Vector Machines (SVM)

Prinsip dasar dari pengklasifikasi *Support Vector Machines* atau SVM diperlihatkan pada Gambar 1. Pada gambar tersebut diperlihatkan kasus paling sederhana dimana vektor-vektor data (titik \times dan \bullet) dapat dipisahkan oleh sebuah *hyperplane*. Terdapat banyak *hyperplane* yang mungkin pada kasus yang demikian. SVM akan mencari *hyperplane* yang menghasilkan margin pemisah terbesar. Skema ini berhubungan dengan minimisasi resiko struktural (*structural risk minimization*).

Pendekatan yang digunakan dalam penelitian ini adalah sebagai berikut. Jaringan normal pada citra mamografi akan dikelompokkan dalam satu kelas, yaitu kelas -1 . Sedangkan jaringan bermikrokalsifikasi akan dikelompokkan dalam kelas yang lain, yaitu $+1$. Sehingga ada dua kelas klasifikasi citra termasuk ke dalam kelas -1 atau kelas $+1$.



Gambar 1. Klasifikasi SVM dengan hyperplane linear yang memaksimalkan margin pemisah antara dua kelas

Dengan ℓ data pelatihan yang dapat dinyatakan dengan $(x_i, y_i) \in R^N \rightarrow \pm 1$ dan $i = 1, 2, \dots, \ell$, SVM akan mengestimasi suatu fungsi $f: R^N \rightarrow \pm 1$ sedemikian sehingga f dapat mengklasifikasikan data uji (x, y) . Data akan diklasifikasikan ke dalam kelas -1 jika $f(x) < 0$

dan akan diklasifikasikan ke dalam kelas +1 jika $f(x) \geq 0$.

SVM akan memilih *hyperplanes* sebagai fungsi yang memisahkan kedua kelas. *Hyperplane* yang optimal adalah *hyperplane* yang memaksimalkan margin pemisah dari data-data yang terdekat. *Hyperplane* dengan margin yang maksimal ditentukan sebagai suatu bentuk fungsi keputusan,

$$f(x) = \text{sgn} \sum_{i=1}^l y_i \alpha_i K(x, x_i) + b \quad (3)$$

$$\begin{cases} \text{maksimalkan} & \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j K(x_i, x_j) y_i y_j \\ \text{dengan} & \sum_{i=1}^l \alpha_i y_i = 0 \quad 0 \leq \alpha_i \leq C \end{cases} \quad (4)$$

dengan $K(x, x_i)$ adalah fungsi kernel, sedangkan α_i dan b ditentukan dengan menyelesaikan persyaratan yang dinyatakan pada formula (4) dan C adalah parameter regularisasi yang dipilih oleh peneliti yang akan menentukan nilai-tawar (*tradeoff*) antara kesalahan empiris dan kerumitan SVM.

JALAN PENELITIAN

Bahan Penelitian

Citra mamografi yang digunakan dalam penelitian ini diperoleh dari basis data MIAS. Citra merupakan citra aras keabuan 8 bit berukuran 1024 x 1024 piksel dengan ukuran piksel 200 mikron. Basis data MIAS memberikan informasi mengenai jenis abnormalitas dan posisinya pada citra. Pada penelitian ini digunakan 10 citra mamografi normal dan 10 citra mamografi bermikrokalsifikasi.

Metode Penelitian

Langkah-langkah yang dilakukan pada penelitian ini diperlihatkan pada Gambar 2.



Gambar 2. Diagram blok langkah penelitian

Proses prapengolahan terdiri atas dua operasi yaitu *cropping* dan transformasi untuk meningkatkan kontras citra. *Cropping* dilakukan untuk memperoleh sampel-sampel jaringan normal dan jaringan bermikrokalsifikasi dari citra mamografi. Dalam penelitian ini digunakan sampel citra berukuran 40 x 40 piksel, sampel

jaringan normal yang digunakan sebanyak 64 sampel dan sampel jaringan bermikrokalsifikasi sebanyak 34 sampel. Jumlah sampel jaringan normal yang digunakan lebih banyak daripada jumlah sampel jaringan bermikrokalsifikasi dengan alasan bahwa jaringan normal jauh lebih bervariasi daripada jaringan bermikrokalsifikasi.

Pada penelitian ini digunakan dua jenis transformasi untuk meningkatkan kontras citra. Yang pertama adalah transformasi menggunakan fungsi penskalaan linear (*linear scaling function*) untuk tiap nilai piksel seperti dinyatakan pada persamaan (1). Yang kedua adalah transformasi *tophat*. Oleh karena mikrokalsifikasi adalah objek terang dengan latar belakang gelap, maka dalam penelitian ini digunakan persamaan (2a) untuk proses peningkatan kontras citra mamografi.

Ekstraksi ciri-ciri citra dilakukan berdasarkan nilai-nilai piksel pada citra sampel. Ekstraksi ciri dilakukan dua kali, yaitu pada sampel dengan transformasi menggunakan fungsi penskalaan linear (*linear scaling function*) dan pada sampel dengan transformasi *tophat*. Ada enam macam ciri yang diekstrak, yaitu rerata, varians, standar deviasi, entropi, beda nilai aras maksimum dan minimum (selanjutnya disebut *bedaan1*), serta beda nilai aras maksimum dengan nilai rerata (selanjutnya disebut *bedaan2*). Dalam pemrograman untuk memperoleh ciri-ciri tersebut digunakan rumusan-rumusan sebagai berikut.

$$\text{rerata} = \frac{1}{N} \sum u(m, n) \quad (5)$$

$$\text{varians} = \frac{1}{N} \sum [u(m, n) - \text{rerata}]^2 \quad (6)$$

$$\text{std} = \sqrt{\frac{1}{N} \sum [u(m, n) - \text{rerata}]^2} \quad (7)$$

$$\text{bedaan1} = \max[u(m, n)] - \min[u(m, n)] \quad (8)$$

$$\text{bedaan2} = \max[u(m, n)] - \text{rerata} \quad (9)$$

$$\text{Entropi} = E[-\log_2 p_u] = - \sum_{x=0}^{L-1} p_u(x) \log_2 p_u(x) \quad (10)$$

dengan $u(m, n)$ adalah nilai piksel citra pada baris ke- m kolom ke- n dan N adalah jumlah piksel dalam sampel citra. Hasil ekstraksi ciri menggunakan dua jenis prapengolahan ini akan

digunakan sebagai masukan pengklasifikasi SVM.

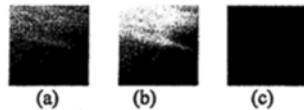
Support Vector Machines (SVM) merupakan metode pembelajaran yang terawasi. Pada penelitian ini digunakan vektor input yang terdiri atas enam ciri citra yang telah diekstrak dari sampel-sampel jaringan normal dan jaringan bermikrokalsifikasi pada tahap sebelumnya. Jaringan normal akan dimasukkan ke dalam satu kelas (yaitu -1) dan jaringan bermikrokalsifikasi dimasukkan ke dalam kelas yang lain (yaitu +1). Pengenalan pola dilakukan dua kali, yaitu dengan masukan ciri-ciri citra dengan prapengolahan menggunakan fungsi penskalaan linear dan transformasi *tophat*. Hasil pengenalan pola menggunakan kedua jenis prapengolahan ini selanjutnya akan dianalisis dan dibandingkan.

HASIL PENELITIAN DAN PEMBAHASAN Hasil Penelitian

Hasil prapengolahan citra menggunakan fungsi penskalaan linear dan transformasi *tophat* diperlihatkan pada Gambar 3, yaitu untuk sampel jaringan normal *norm_11.tif*. Sedangkan pada Gambar 4 diperlihatkan hasil prapengolahan untuk sampel jaringan bermikrokalsifikasi *calc_18.tif*.

Berikut adalah contoh hasil ekstraksi ciri dengan prapengolahan transformasi *tophat* untuk sampel *norm_11.tif* dan *calc_18.tif* (berturut-turut adalah ciri rerata, varians, standar deviasi, entropi, beda1, beda2).

stats = 1.8825 6.9562 2.6375 2.5988 13.0000 11.1175
stats = 3.8706 69.7926 8.3542 3.2360 72.0000 68.1294



Gambar 3. (a) Sampel jaringan normal, *norm_11.tif* (b) Sampel setelah prapengolahan dengan fungsi penskalaan linear (c) Sampel setelah prapengolahan dengan transformasi *tophat*



Gambar 4. (a) Sampel jaringan bermikrokalsifikasi; *calc_18.tif* (b) Sampel setelah prapengolahan dengan fungsi penskalaan linear (c) Sampel setelah prapengolahan dengan transformasi *tophat*

Tabel 1 memperlihatkan jangkauan nilai masing-masing ciri untuk semua sampel yang digunakan dalam penelitian dengan prapengolahan menggunakan fungsi penskalaan linear. Sedangkan Tabel 2 memperlihatkan jangkauan nilai masing-masing ciri untuk semua sampel yang digunakan dalam penelitian dengan prapengolahan transformasi *tophat*. Ciri-ciri yang telah diekstrak dari sampel akan digunakan sebagai masukan SVM.

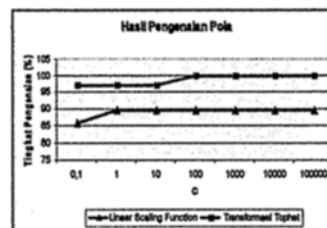
Tabel 1 Jangkauan nilai ciri (dengan transformasi menggunakan fungsi penskalaan linear)

Ciri	Jangkauan Nilai	
	Jaringan Normal	Jaringan Bermikrokalsifikasi
Rerata	0,2 – 182,9	21,2 – 85,9
Varians	40,6 – 7829	114,4 – 1938,9
Std	6,4 – 88,5	10,7 – 44
Entropi	0 – 7,1	4,6 – 6,2
Bedaan1	52 – 255	78 – 204
Bedaan2	19,2 – 254,8	52,9 – 146,2

Tabel 2 Jangkauan nilai ciri (dengan transformasi *tophat*)

Ciri	Jangkauan Nilai	
	Jaringan Normal	Jaringan Bermikrokalsifikasi
Rerata	0,0050 – 2,8169	1,8425 – 3,8706
Varians	0,0040 – 10,3858	4,3215 – 69,7926
Std	0,1999 – 3,2227	2,0788 – 8,3542
Entropi	0,0076 – 7,0703	4,5707 – 6,1755
Bedaan1	3 – 20	16 – 72
Bedaan2	2,6981 – 17,6806	13,7125 – 68,1294

Hasil pengenalan pola jaringan normal dan jaringan bermikrokalsifikasi dengan SVM diperlihatkan dalam bentuk grafik pada Gambar 5, masing-masing menggunakan data ciri dari sampel dengan prapengolahan menggunakan fungsi penskalaan linear dan transformasi *tophat*. Pengenalan pola dilakukan menggunakan kernel linear dengan nilai C yang bervariasi (0,1 hingga 10^5).



Gambar 5. Hasil pengenalan pola jaringan normal dan jaringan bermikrokalsifikasi menggunakan SVM

Dengan membandingkan hasil penelitian yang diperlihatkan pada Gambar 3 dan 4, dapat diketahui bahwa prapengolahan menggunakan fungsi penskalaan linear dan transformasi *tophat* secara visual dapat mempertegas perbedaan antara jaringan normal dan jaringan bermikrokalsifikasi. Transformasi *tophat* secara visual lebih baik untuk membedakan kedua jenis jaringan. Hal ini juga dapat diketahui dengan memperhatikan rentang atau jangkauan masing-masing ciri yang diekstrak untuk kedua jenis jaringan yang diperlihatkan pada Tabel 1 dan Tabel 2. Meskipun terdapat tumpang-tindih (*overlapping*) jangkauan pada semua ciri yang diekstrak, namun kombinasi enam ciri sebagai masukan SVM ternyata mampu menghasilkan tingkat pengenalan tertinggi yaitu 100%. Tingkat pengenalan tertinggi ini dicapai dengan prapengolahan menggunakan transformasi *tophat* dengan nilai C sebesar 100. Pada Gambar 5 juga diperlihatkan hasil pengenalan pola dengan prapengolahan menggunakan fungsi penskalaan linear. Dengan prapengolahan jenis ini, tingkat pengenalan pola tertinggi yang dapat dicapai adalah 89,5%.

KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan dapat ditarik beberapa kesimpulan sebagai berikut.

Pertama, transformasi *tophat* secara visual dapat lebih mempertegas perbedaan antara jaringan normal dan jaringan bermikrokalsifikasi dibandingkan transformasi menggunakan fungsi penskalaan linear.

Kedua, *Support Vector Machines* (SVM) dapat digunakan untuk pengenalan pola jaringan normal dan jaringan bermikrokalsifikasi pada citra mamografi.

Ketiga, tingkat pengenalan tertinggi pada penelitian ini adalah 100%. Hasil ini dicapai saat SVM menggunakan kernel linear dan nilai C sebesar 100, serta prapengolahan citra dilakukan dengan menggunakan transformasi *tophat*.

Saran

Hasil pengenalan pola pada penelitian ini selanjutnya dapat digunakan untuk mendeteksi keberadaan jaringan bermikrokalsifikasi pada citra mamografi digital dengan bantuan komputer. Sistem deteksi yang otomatis berbantuan komputer dapat digunakan sebagai alat bantu bagi dokter atau ahli radiologi dalam melakukan analisis hasil citra rekam medis, khususnya rekam medis mamografi.

DAFTAR RUJUKAN

- Abubaker, A.A., Qahwadi, R.S., Aqel, M.J., Al-Osta, H., Saleh, M.H., 2006, *Efficient Pre-Processing of USF and MIA S Mammograms Images*, Journal of Computer Science 3 (2): 67 – 75, ISSN 1549 – 3636.
- Campanini, R., Dongiovanni, D., Iampieri, E., Lanconelli, N., Massoti, M., Palermo, G., Riccardi, A., Roffilli, M., *A Novel Featureless Approach to Mass Detection in Digital Mammograms Based on Support Vector Machines*, Department of Physics and Department of Computer Science, University of Bologna Italy.
- Candade, N.V., 2004, *Application of Support Vector Machines and Neural Networks in Digital Mammography: A Comparative Study*, Thesis, Biomedical Engineering Department of Chemical Engineering College of Engineering, University of South Florida.
- El-Naqa, I., Yang, Y., Wernick, M.N., Galatsanos, N.P., Nishikawa, R., 2002, *Support Vector Machine Learning for Detection of Microcalcifications in Mammograms*, Dept. Of Electrical and Computer Engineering, Illinois Institute of Technology.
- Iqbal, K., Salam, R.A., Osman, A., Talib, A.Z., 2007, *Underwater Image Enhancement Using an Integrated Colour Model*, IAENG International Journal of Computer Science.
- Jain, A.K., 1989, *Fundamental of Digital Image Processing*, Prentice Hall International, Inc. Singapore.
- Liu, Sheng., Babbs, Charles F., dan Delp, Edward J., *Normal Mammogram Analysis and Recognition*, Purdue University, Indiana.
- Melloul, M., Joskowicz, L., 2002, *Segmentation of Microcalcification in X-ray Mammograms Using Entropy Thresholding*, CARS 2002.
- Munir, R., 2004, *Pengolahan Citra Digital dengan Pendekatan Algoritmik*, Informatika Bandung.
- Sheshadri, H.S. dan Kandaswamy, A., 2006, *Computer Aided Decision System for Early Detection of Breast Cancer*, Indian J Med Res 124, pp 149 – 154.
- Srivastava, V., 2005, *Perfomance of Microcalcification Detection Algorithms*, Thesis, Department of Elecktrical and

- Computer Engineering, North Carolina State University, Raleigh.
- Tembey, M., 2003, Computer - Aided Diagnosis for Mammographic Microcalcification Clusters, Thesis, Department of Computer Science and Engineering, College of Engineering, University of South Florida.
- Yaffe, Martin J., 1995, *Mammography, Biomedical Engineering Handbook*, IEEE Press, USA.
- Anonim, Image Processing Toolbox, <http://www.mathworks.com/access/helpdesk/help/toolbox/images/>
- Anonim, MIAS Database, <http://peipa.essex.ac.uk/ipa/pix/mias>
- Anonim, Breast Cancer <http://www.imaginis.com/breasthealth/statistics.asp>,