

Klasifikasi Citra Kanker Kulit Melanoma Menggunakan Metode *Support Vector Machine* (SVM)

Kusuma Wahaninggar, Dwi Ratna S
Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember (ITS) Surabaya
Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia
e-mail: dwratna@matematika.its.ac.id

Abstrak— Kanker Kulit melanoma merupakan salah satu jenis kanker kulit yang mematikan. Cara mendeteksi melanoma adalah dengan menggunakan metode Biopsi. Untuk melakukan metode ini membutuhkan kinerja dokter terlatih. Proses Biopsi menyakitkan dan membutuhkan waktu yang cukup lama. Karena itu pada penelitian ini telah dilakukan klasifikasi citra kanker kulit melanoma dengan menggunakan metode *Support Vector Machine* (SVM). Adapun tahapan dari proses klasifikasi citra kanker kulit melanoma pada penelitian ini meliputi proses pra-pemrosesan, segmentasi, ekstraksi fitur dengan ABC yaitu *Asymmetry*, *Border Irregularity*, *Colour Variation* dan *Gray-Level Co-Occurrence Matrix* (GLCM) serta klasifikasi citra dengan SVM. Dengan menggunakan dataset 80 citra melanoma dan 70 citra bukan melanoma. Hasil pengujian menunjukkan bahwa sistem dapat mengenali citra melanoma dengan tingkat akurasi sebesar 90%.

Kata kunci : Kanker Kulit, Ekstraksi Fitur, *Asymmetry*, *Border Irregularity*, *Colour Variation* (ABC) dan *Gray-Level Co-Occurrence Matrix* (GLCM), *Support Vector Machine* (SVM).

I. PENDAHULUAN

Kanker kulit melanoma merupakan jenis kanker kulit yang paling mematikan. Melanoma dimulai di sel-sel kulit dan bisa menyebar ke organ lain dalam tubuh melalui darah. Kemunculan tahi lalat baru atau perubahan pada tahi lalat yang sudah ada merupakan pertanda umum gejala melanoma [2].

Penyebab utama melanoma karena kulit sering terkena sinar ultra violet. Umumnya dokter menggunakan metode Biopsi untuk mendiagnosis melanoma. Metode Biopsi adalah pengambilan potongan kecil jaringan tubuh untuk sampel diagnosis. Dengan cara memotong atau menggores sepotong kecil jaringan atau dengan menggunakan jarum suntik untuk mengambil sampel, kemudian sampel diuji oleh beberapa dokter terlatih untuk didiagnosis. Proses Biopsi ini menyakitkan dan membutuhkan waktu yang cukup lama. Dari penelitian yang telah dilakukan, diagnosis dini melanoma menunjukkan lebih dari 90% dapat disembuhkan dan jika terlambat mendiagnosis melanoma menunjukkan kurang dari 50% dapat disembuhkan. Karena itu jika melanoma tidak terdiagnosis sejak tahap awal, maka akan menyebabkan

kematian pada penderita. Sehingga mendeteksi kanker kulit dari tahap awal merupakan hal yang tak terelakkan [1,2].

Pada pengolahan citra digital, ada beberapa metode untuk mendeteksi citra kanker kulit melanoma. Jika sudah didapatkan ciri citra melanoma maka proses klasifikasi dapat dilakukan. Salah satu metode yang sering mendapat perhatian dalam klasifikasi citra adalah *Support Vector Machine* (SVM). *Support Vector Machine* (SVM) berusaha menemukan *hyperplane* yang terbaik pada *input space*. Akhir-akhir ini *Support Vector Machine* (SVM) telah berhasil diaplikasikan untuk menyelesaikan masalah di dunia nyata dan secara umum memberikan solusi yang lebih baik dibandingkan metode lainnya. Penelitian yang menggunakan metode *Support Vector Machine* diantaranya Analisis Diagnosis Pasien Kanker Payudara Menggunakan Regresi Logistik dan *Support Vector Machine* (SVM) Berdasarkan Hasil Mamografi [3] dan penelitian tentang klasifikasi cacat pada selongsong peluru dengan metode *Support Vector Machine* (SVM) [5].

Berdasarkan latar belakang serta kajian pada penelitian tersebut, maka pada tugas akhir ini telah meneliti tentang klasifikasi citra kanker kulit melanoma menggunakan metode *Support Vector Machine* (SVM) berbasis citra digital.

II. KAJIAN TEORI

2.1 Pengolahan citra

Pengolahan citra adalah suatu metode yang digunakan untuk memroses citra dalam bentuk dua dimensi. Segala operasi untuk memperbaiki, menganalisa, atau melakukan transformasi suatu citra menjadi citra lain dengan menggunakan teknik tertentu disebut pemrosesan citra (*image processing*). Pemrosesan citra telah diaplikasikan dalam berbagai bentuk, dengan tingkat kesuksesan yang cukup besar. Konsep-konsep pemrosesan citra dalam tugas akhir ini meliputi pra-pemrosesan, segmentasi, ekstraksi fitur, dan klasifikasi.

Tahap pra-pemrosesan bertujuan untuk menghilangkan *noise* dan mempersiapkan citra agar bisa diproses untuk proses ekstraksi fitur. Tahapan-tahapan yang dilakukan dalam pra-pemrosesan antara lain :

2.1.1 Median Filtering

Peningkatan kualitas citra menggunakan filter median. Filter ini memilih intensitas piksel yang ditengah, setelah piksel-piksel yang tercakup dalam filter diurutkan. Filter median sering digunakan karena filter ini memberikan

kemampuan reduksi *noise* yang sangat baik dengan *blurring* yang sedikit.

2.1.2 Metode *Mapping* Nilai Intensitas

Metode penguatan nilai kekontrasan menggunakan teknik *mapping* nilai intensitas. Intensitas piksel-piksel pada citra input dimap dengan batasan atau *threshold* tertentu, misal senilai *minS* untuk batas minimal dan *maxS* untuk batas maksimal. Maka piksel-piksel pada citra input yang memiliki intensitas dibawah batas minimal, nilainya akan diset sama dengan batas minimal. Begitu juga untuk piksel-piksel pada citra input yang memiliki intensitas diatas batas maksimal, nilainya akan diset sama dengan batas maksimal. Untuk piksel-piksel pada citra input yang memiliki intensitas diantara batas minimal dan batas maksimal, maka nilainya dipetakan/ diskalakan (*mapping*) untuk memenuhi rentang nilai-nilai keabuan yang lengkap (dari batas minimal sampai dengan batas maksimal) dengan persamaan [6]:

$$S = \frac{(R - \min R)x(\max S - \min S)}{(\max R - \min R)} + \min S \quad (1)$$

Dengan,

S = citra output

maxS = batas maximal (untuk citra output)

minS = batas minimal (untuk citra output)

R = citra input

maxR = intensitas terbesar dari piksel di citra input

minR = intensitas terkecil dari piksel di citra input

2.1.3 Segmentasi Citra

Segmentasi menggunakan metode *thresholding* otsu, yaitu metode untuk segmentasi berdasarkan histogram citra. Metode otsu berasumsi bahwa citra yang akan dicari nilai *threshold*-nya memiliki 2 kelas piksel, yaitu *background* dan *foreground*. Cara kerja metode Otsu dengan memaksimalkan varians antar kelas (*between-class variance*). Varians antar kelas ini cocok digunakan untuk analisa diskriminan kelas secara statistik [7].

Algoritma dari metode *thresholding* Otsu untuk pemilihan satu *threshold*:

- Citra dibagi menjadi 2 piksel yaitu G1 dan G2
- Hitung bobot dari masing-masing kelompok piksel G1 dan G2

$$w1 = \sum_0^t p(i) \quad (2)$$

Dan

$$w2 = \sum_t^n p(i) \quad (3)$$

- Hitung rata-rata intensitas μ_1 dan μ_2 untuk piksel di daerah G1 dan G2 :

$$\mu_1 = [\sum_0^t p(i)x(i)]/w1 \quad (4)$$

Dan

$$\mu_2 = [\sum_t^n p(i)x(i)]/w2 \quad (5)$$

- Hitung varians antar kelas (*between-class variance*) pada persamaan :

$$Vb = w1.w2((\mu_1 - \mu_2)^2) \quad (6)$$

- Ulangi langkah a sampai d, sehingga didapatkan nilai *Vb* terbesar. Maka itulah nilai k^* yang digunakan.

Setelah mendapatkan nilai ambang batas atau *threshold* maka kita gunakan untuk segmentasi citra. Metode Otsu hanya untuk menemukan *threshold* yang optimum. Teknik segmentasi menggunakan persamaan :

$$g(x,y) = \begin{cases} 1, & \text{jika } f(x,y) > k^* \\ 0, & \text{jika } f(x,y) \leq k^* \end{cases} \quad (7)$$

Hasil segmentasi adalah citra biner yang memiliki nilai intensitas hanya 0 dan 1. Nilai intensitas 0 menyatakan warna hitam dan (dalam Tugas Akhir ini dianggap *background*) sedangkan nilai intensitas 1 menyatakan warna putih (dianggap *foreground* atau objek).

2.1.4 Metode *Flood-Filling*

Metode *flood-fill* disebut juga dengan *seed fill*. Algoritma *flood-fill* memiliki 3 parameter masukan, yaitu node awal, warna target, dan warna pengganti. Algoritma ini bekerja dengan mencari semua node yang terkoneksi dengan node awal sebagai sebuah jalur berwarna target, lalu mengubahnya dengan warna pengganti.

2.2 Ekstraksi fitur

Tujuan dari ekstraksi fitur adalah untuk mendapatkan karakteristik pembeda yang mewakili sifat utama dengan memisahkannya dari fitur yang tidak diperlukan pada proses klasifikasi. Ekstraksi fitur yang digunakan untuk penelitian ini adalah ekstraksi fitur ABC dan GLCM.

2.2.1 *Asymmetry*

Sebuah tahi lalat umumnya berbentuk simetris dan sirkular sedangkan melanoma biasanya tumbuh secara irregular dan asimetris. Pengukuran nilai asimetri bergantung pada suatu properti geometris yang disebut momen inersial. Momen inersial digunakan untuk mengkalkulasi dua fitur, yaitu *asymmetry index AI* dan *lengthening index A* [8].

1. *Asymmetry Index* (AI)

Pada pencarian nilai *asymmetry index*, titik *origin* (0,0) dari koordinat kartesius ditetapkan pada pusat massa *G* dari area luka *L*, dengan *L* didefinisikan oleh fungsi biner $z(i,j)$, dengan $z(i,j) = 1$ jika $(i,j) \in L$, selain itu 0. Momen inersia kuadratik $I(\varphi)$ dari suatu citra luka *L* yang berhubungan dengan sembarang sumbu yang melalui *G* menunjukkan sudut φ terhadap sumbu horizontal dari koordinat kartesius Δ , dan diberikan dalam persamaan :

$$I(\varphi) = \sum_{(i,j) \in L} D_\varphi^2(i,j) = \sum_{(i,j) \in L} [-i \sin \varphi + j \cos \varphi]^2 \quad (8)$$

$D_\varphi(i,j)$ adalah jarak antara piksel yang ditinjau (i,j) dan proyeksinya pada Δ , sepanjang normal garis Δ . Sumbu utama mayor didapat dari momen inersia terkecil dari *L* dan memiliki arah tegak φ_0 . Sumbu utama mayor tersebut bisa didapat dengan menghitung turunan dari persamaan tersebut dan mengeset turunannya dengan 0 seperti dalam persamaan berikut :

$$\frac{\partial I(\varphi)}{\partial \varphi} = 0 \Rightarrow \varphi_0 = \frac{1}{2} \tan^{-1} \left[\frac{2m_{11}^c}{m_{20}^c - m_{02}^c} \right] \quad (9)$$

m_{11}^c merepresentasikan momen standar, m_{20}^c merepresentasikan momen kuadratik yang berhubungan dengan sumbu horizontal koordinat kartesius G_x dan m_{02}^c merepresentasikan momen kuadratik yang berhubungan dengan sumbu vertical koordinat kartesius G_y . Sumbu utama minor dari *L* memiliki arah $\varphi_0 + \pi/2$ menghasilkan arah yang melintang terhadap *L* yang berarti memiliki momen inersia terbesar. Sumbu melintang dan tegak yang telah didapatkan bisa dilakukan untuk mengkalkulasi nilai

asymmetry index. Dilakukan dengan melipat area luka L pada kedua sumbu tadi dan mengukur area *non overlap* seperti dalam persamaan

$$AI = \frac{1}{2} \sum_{k=1}^2 \frac{\Delta A_k}{A_L} \quad (10)$$

Subscript k mengidentifikasi sumbu utama (mayor vs. minor), ΔA_k merupakan area *non-overlap* saat luka dilipat dan $A_L (A_L = m_{00}^c)$ adalah area luka.

2. Lengthening Index

Pengukuran ini digunakan untuk mendeskripsikan pemanjangan sebuah luka, misalnya derajat *anisotropy* luka. Pemanjangan luka berhubungan dengan nilai eigen λ', λ'' dari matriks tensor inersia. *Lengthening Index* ini didefinisikan oleh perbandingan antara momen inersia λ' (terhadap sumbu mayor) dan momen inersia λ'' (terhadap sumbu minor).

$$\hat{A} = \frac{\lambda'}{\lambda''} \quad (11)$$

$$\lambda' = \frac{m_{20}^c + m_{02}^c - \sqrt{(m_{20}^c - m_{02}^c)^2 + 4(m_{11}^c)^2}}{2} \quad (12)$$

$$\lambda'' = \frac{m_{20}^c + m_{02}^c + \sqrt{(m_{20}^c - m_{02}^c)^2 + 4(m_{11}^c)^2}}{2} \quad (13)$$

2.2.2 Border Irregularity

Luka tidak berbahaya umumnya mempunyai pembatas yang jelas. Sebuah melanoma sering menunjukkan pembatas yang tidak jelas atau kabur yang mensinyalkan pertumbuhan dan penyebaran kanker.

Iregularitas pembatas luka/ tepi area luka (*Border Irregularity Lesion*) ini sudah mempresentasikan faktor diagnostik yang sangat signifikan ketika menduga bahaya sebuah luka. Dalam Tugas Akhir ini, digunakan lima fitur untuk mengkuantifikasi *Border Irregularity*, antara lain: indeks kepadatan (*Compactness Index*) CI , dimensi fraktal (*Fractal Dimension*) fd , kecuraman tepi (*Edge Abruptness*) Cr , dan transisi pigmentasi (*Pigmentation Transition*) m_e dan v_e [8].

1. Compactness Index

Indeks kepadatan (*Compactness Index/ CI*) adalah pengukuran bentuk pembatas paling populer yang mengestimasi kebutuhan obyek 2D. Namun, pengukuran ini sangat sensitive terhadap noise di sepanjang pembatas/ tepi area luka yang nilainya dikuatkan oleh kuadrat dari keliling luka. Berikut ini persamaan Indeks kepadatan (*Compactness Index/ CI*)

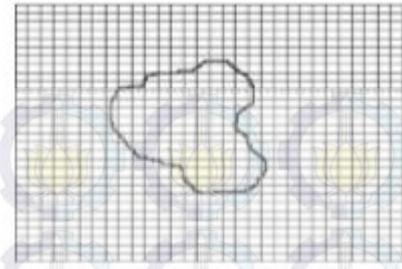
$$CI = \frac{P_L^2}{4\pi A_L} \quad (14)$$

P_L = keliling luka

2. Fractal Dimension

Dimensi fractal dapat dihitung dengan metode perhitungan kotak (*box-counting*). Metode ini membagi citra menjadi kotak-kotak dengan berbagai variasi ukuran (r).

Salah satu contoh pengukuran nilai r adalah $2k$, dengan $k = 0,1,2, \dots$ dst, dan $2k$ lebih kecil dari ukuran citra. Contoh untuk citra luka pada kulit, digunakan grid kotak yang membagi citra dengan masing-masing kotak berukuran $r \times r$ (Gambar 1). $N(r)$ adalah sebuah piksel yang berisi potongan pembatas/ tepi luka. Ukuran piksel berbeda (r berbeda) digunakan dan fd diperoleh dari gradient atau kemiringan garis regresi $\log(r)$ terhadap $\log(N(r))$.



Gambar 1. Kalkulasi fractal dimension (fd) menggunakan metode box-counting

Secara keseluruhan, hubungan antara $N(r)$ dan fd dapat diungkapkan sebagai berikut:

$$N(r) \propto r^{-fd} \quad (15)$$

3. Edge Abruptness

Luka dengan pembatas atau tepi yang irregular (*Edge Abruptness*) memiliki nilai variance jarak radial yang besar (contoh d_2 , jarak antara centroid G_L dan pembatas atau tepi C). Untuk mengestimasi iregularitas pembatas dilakukan analisa nilai variance dari distribusi jarak radial

$$C_r = \frac{\frac{1}{P_L} \sum_{p \in C} (d_2(p, G_L) - m_d)^2}{m_d^2} \quad (16)$$

m_d adalah mean jarak d_2 anantara titik-titik pembatas dan centroid G_L (titik pusat luka).

4. Pigmentation Transition

Fitur penting ini menjelaskan transisi pigmentasi antara area luka dan kulit sekitarnya. Tepi area luka yang mengalami perubahan pigmentasi secara tajam mengindikasikan bahwa luka adalah melanoma, sedangkan perubahan secara perlahan (*gradasi*) mengindikasikan luka tidak berbahaya. Karena itu, dipertimbangkan komponen luminansi $lum(i, j)$ (Persamaan 17) dari warna citra asli sebagai tiga komponen warna yang berbobot sama. Kemudian diestimasi *gradient magnitude* dari komponen intensitas lum sepanjang pembatas/ tepi luka C di kulit. Diperoleh set-set nilai *gradient magnitude* K , $e(k)$ ($1 \leq k \leq K$, dimana K adalah jumlah sampel tepi luka) yang mendeskripsikan secara local transisi antara luka dan kulit di setiap piksel tepi luka pada citra. Untuk menjelaskan transisi secara global, digunakan mean m_e dan variance v_e dari nilai *gradient magnitude* yang mendeskripsikan level kecuraman dan variasinya secara global (Persamaan 18).

$$lum(i, j) = \frac{1}{3} [r(i, j) + g(i, j) + b(i, j)] \quad (17)$$

$$m_e = \frac{1}{K} \sum_{k=1}^K e(k), v_e = \sum_{k=1}^K e^2(k) - m_e^2 \quad (18)$$

2.3.3 Color Variation

Salah satu tanda awal dari melanoma adalah munculnya variasi warna dalam warna. Karena melanoma tumbuh dalam sel penumbuh pigmen, umumnya berwarna coklat, coklat gelap, atau hitam, tergantung kondisi pigmen melanin di berbeda pada kulit. Untuk membatasi diagnosis lebih lanjut, variasi warna pada sebuah luka dijelaskan dengan homogenitas warna (*Color Homogeneity*) C_h dan korelasi antara geometri dan fotometri (*Correlation between Photometry and Geometry*) C_{pg} [8].

1. Color Homogeneity

Histogram dari luminansi warna pada luka dibagi dalam tiga interval yang panjangnya sama. Interval yang berelasi ke nilai luminansi ketiga terkecil mendefinisikan area gelap. *Level intermediate* berelasi ke yang lain dan tidak terlibat dalam kuantifikasi warna. Kemudian homogenitas warna dijelaskan sebagai jumlah transisi zona lebih terang/zona lebih gelap ke zona lebih gelap/zona lebih terang ketika memindai luka secara horizontal dan vertical.

2. Correlation between Photometry and Geometry

Atribut ini mengevaluasi distribusi warna pada luka. Termasuk penjelasan evaluasi level warna dari centroid G_L menuju pembatas/ tepi luka (persamaan 18). Nilai lebih besar menunjukkan luka tak-berbahaya karena luka tak-berbahaya memiliki aspek target, sedangkan nilai yang kecil menandakan bahaya melanoma.

$$C_{pg} = \frac{1}{A_L} \cdot \sum_{P \in L} \frac{(lum(p) - m_1) \cdot (d_2(p, G_L) - m_d)}{v_1 \cdot v_d} \quad (19)$$

m_d dan v_d adalah mean dan variance, d_2 adalah jarak antara titik pembatas dengan centroid, m_1 dan v_1 berhubungan dengan luminansi, P adalah semua elemen piksel dan L adalah area luka.

2.3.4 GLCM

GLCM (*Gray-Level Co-occurrence Matrix*) merupakan ekstraksi fitur yang sering digunakan untuk analisis tekstur. Piksel yang bertetangga (i, j) dan memiliki jarak d diantara keduanya, dapat terletak di delapan arah yang berlainan baik horizontal, vertical, maupun diagonal. Berikut adalah persamaan ekstraksi fitur GLCM [9]:

1. Entropi

$$Entropy = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} p(i, j) \log(p(i, j)) \quad (20)$$

dengan $P_{(i,j)}$ = Piksel baris ke $-i$ dan kolom ke $-j$

2. Energi

$$Energy = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} p(i, j)^2 \quad (21)$$

dengan $P_{(i,j)}$ = Piksel baris ke $-i$ dan kolom ke $-j$.

3. Kontras

$$Contrast = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |i - j|^2 p(i, j) \quad (22)$$

dengan $P_{(i,j)}$ = Piksel baris ke $-i$ dan kolom ke $-j$.

4. Homogenitas

$$Homogenitas = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{p(i, j)}{1 + |i - j|} \quad (23)$$

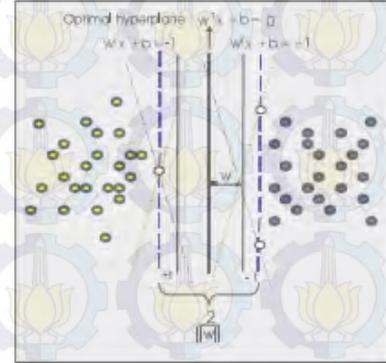
dengan $P_{(i,j)}$ = Piksel baris ke $-i$ dan kolom ke $-j$

2.3 Klasifikasi dengan SVM

Support Vector Machine (SVM) merupakan sebuah metode klasifikasi linier dengan *feature set* yang telah ditentukan sebelumnya. Metode SVM akan mencari sebuah *hyperplane* linier dengan margin terbesar untuk memisahkan kelas yang ada. Margin terbesar, yang kemudian disebut sebagai *Maximum Marginal Hyperplane* (MMH) tersebut akan memberikan jarak terbesar antar kelas. Jarak antara *hyperplane* dengan sebuah sisi dari margin tersebut sama dengan jarak antara *hyperplane* dengan margin pada sisi lainnya.

Dalam kasus real, biasanya data yang ada bersifat *non linearly separable* data. Metode ini mentransformasikan data ke dalam dimensi ruang fitur (*feature space*) sehingga dapat diipisahkan secara linear pada *feature space*. *Feature space* dalam prakteknya biasanya berdimensi lebih tinggi dari

vektor input (*input space*). Hal ini mengakibatkan komputasi pada *feature space* mungkin akan sangat besar, karena ada kemungkinan *feature space* dapat memiliki jumlah fitur yang tidak terhingga. Selain itu sulit untuk mengetahui fungsi transformasi yang tepat. Untuk mengatasi masalah ini, pada SVM digunakan "*kernel trick*". Untuk menyelesaikan masalah pada *non linearly separable* data harus melakukan modifikasi pada beberapa fungsi yang ada.



Gambar 2. SVM Berusaha Mencari Hyperplane Terbaik [10].

Ada banyak jenis kernel yang biasa digunakan dalam proses SVM. Pada tugas akhir ini metode yang telah diteliti untuk memetakan dari *input space* ke *feature space* ada tiga yaitu kernel polynomial, kernel linear, dan kernel gaussian/rbf [10]. Dari ketiga kernel tersebut dicari kernel mana yang memberikan hasil lebih akurat.

Fungsi kernel polynomial memiliki bentuk persamaan :

$$K(x_i, x_j) = ((x_i x_j^T) + 1)^n \quad (24)$$

Fungsi kernel linear memiliki bentuk persamaan :

$$K(x_i, x_j) = x_i x_j^T \quad (25)$$

Fungsi kernel rbf memiliki bentuk persamaan :

$$K(x_i, x_j) = \exp \frac{\|x_i - x_j\|^2}{2\sigma^2} \quad (26)$$

$i, j = 1, 2, \dots, l$ dengan :

K = kernel

x_i = data inputan (*training*)

x_j = data inputan (*training/testing*) yang telah di transpos

σ = parameter bebas

l = banyaknya data

Setelah dilakukan transformasi dari *input space* ke *feature space* dilakukan, langkah selanjutnya adalah pencarian *hyperplane* klasifier. Untuk itu dilakukan langkah-langkah berikut [10]:

1. Data yang tersedia dinotasikan sebagai $x \in R$ sedangkan label masing-masing dinotasikan $y_i \in \{-1, +1\}$ untuk $i=1, 2, \dots, l$ yang mana l adalah banyaknya data. Diasumsikan kedua *class* -1 dan $+1$ dapat terpisah secara sempurna oleh *hyperplane* berdimensi d , yang didefinisikan:

$$\bar{w} \cdot x + b = 0 \quad (27)$$

Sebuah *pattern* x_i yang termasuk *class* -1 (sampel negatif) dapat dirumuskan sebagai *pattern* yang memenuhi persamaan :

$$\bar{w} \cdot x + b = -1 \quad (28)$$

sedangkan *pattern* yang termasuk *class* $+1$ (sampel positif):

$$\bar{w} \cdot x + b = +1 \quad (29)$$

2. Margin terbesar dapat ditemukan dengan memaksimalkan nilai jarak antara *hyperplane* dan

titik terdekatnya, yaitu $1/\|w\|$. Hal ini dapat dirumuskan sebagai *Quadratic Programming* (QP) *problem*, yaitu mencari titik minimal dengan memperhatikan *constraint* :

$$\min_w \tau(w) = \frac{1}{2} \|w\|^2 \quad (30)$$

dengan batasan :

$$y_i(x_i \cdot w + b) - 1 \geq 0 \quad (31)$$

3. Problem diatas dapat dipecahkan dengan berbagai teknik komputasi, diantaranya *Lagrange Multipliers* sebagaimana ditunjukkan pada Persaman :

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i (y_i((x_i \cdot w + b) - 1)) \quad (32)$$

α_i adalah *Lagrange Multipliers*, yang bernilai nol atau positif ($\alpha_i \geq 0$).

4. Nilai optimal dari persamaan diatas dapat dihitung dengan meminimalkan L terhadap w dan b , dan memaksimalkan L terhadap α_i . Dengan memperhatikan sifat bahwa pada titik optimal gradient $L=0$, persamaan dapat dimodifikasi sebagai maksimalisasi problem yang hanya mengandung α_i saja, sebagaimana Persamaan :

$$= \max \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (33)$$

dengan batasan :

$$\alpha_i \geq 0 \quad (i = 1, 2, \dots, l) \quad (34)$$

$$\sum_{i=1}^l \alpha_i y_i = 0 \quad (35)$$

Dari hasil dari perhitungan ini diperoleh α_i yang kebanyakan bernilai positif. Data yang berkorelasi dengan α_i yang positif inilah yang disebut sebagai *support vector*.

III. METODOLOGI PENELITIAN

3.1 Gambaran Sistem Secara Umum

Gambaran sistem klasifikasi citra kanker kulit menggunakan metode SVM mempunyai beberapa tahapan, yakni input data, pra-pemrosesan, segmentasi, ekstraksi fitur dengan ABC dan GLCM, klasifikasi citra dengan SVM.

a. Input Data

Data citra masukan dari proses ini adalah citra dermatoskopis yaitu citra melanoma dan citra bukan melanoma. Citra-citra tersebut akan diproses untuk menentukan apakah citra tersebut melanoma atau hanya citra luka yang tidak berbahaya.

b. Pra-pemrosesan

Proses pra-pemrosesan merupakan tahap pertama yang dilakukan sebelum memperoleh fitur dari suatu citra. Dalam pra-pemrosesan terdapat dua subproses yaitu median filtering dan mapping nilai intensitas citra.

c. Segmentasi

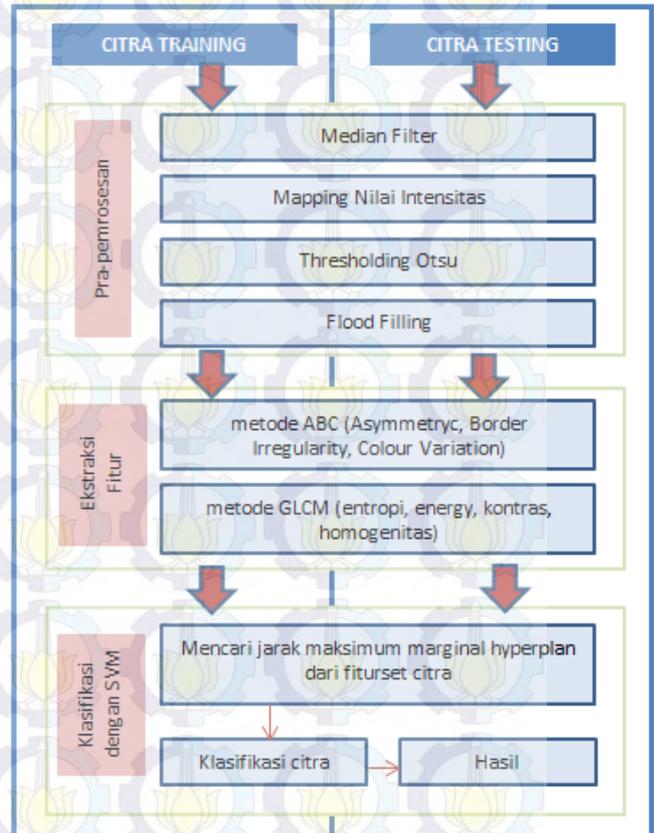
Setelah mendapatkan hasil keluaran dari proses pra-pemrosesan, maka citra output dari proses pra-pemrosesan tersebut diproses untuk mendapatkan daerah luka melalui proses segmentasi. Proses segmentasi dengan thresholding kemudian dengan flood-filling.

d. Ekstraksi Fitur

Ekstraksi fitur merupakan suatu pengambilan ciri dari suatu bentuk yang nilainya akan dianalisis untuk proses selanjutnya. Hal ini bertujuan untuk memperoleh ciri citra melanoma yang akan digunakan untuk proses berikutnya. Pada proses ekstraksi fitur ini akan dihitung masing-masing nilai ABC yaitu Asymmetry, Border Irregularity, Color Variation, dan GLCM yaitu Energi, Entropi, Kontras, Homogenitas.

e. Klasifikasi dengan SVM

Proses ini merupakan proses pengklasifikasian antara citra melanoma dan citra bukan melanoma. Ada 2 proses yang dilakukan pada proses ini, yaitu proses *training* dan *testing*.



Gambar 3. Diagram Alir Keseluruhan Proses.

IV. PENGUJIAN DAN PEMBAHASAN

4.1 Lingkungan Pengujian Sistem

Perangkat yang digunakan dalam pengujian sistem terdiri dari beberapa perangkat keras dan perangkat lunak. Perangkat keras yang digunakan yaitu komputer dengan Prosesor Intel(R) Core(TM)i5-2410M CPU @ 2.30GHz, Memory 4 GB. Sedangkan perangkat lunak yang digunakan adalah Matlab.

4.2 Pengujian Tahap Pra-pemrosesan

Pengujian ini dilakukan terhadap proses-proses pada tahap pengolahan citra. Pengujian bertujuan untuk mengetahui bahwa proses-proses pada tahap pengolahan citra sudah benar. Pada tahap ini terdapat 4 proses yaitu : median *filtering*, *mapping* nilai intensitas, segmentasi citra dengan *thresholding* otsu, dan *flood filling*.

1. Pengujian proses median *filtering*

Pengujian proses ini bertujuan untuk mengetahui bahwa sistem telah berhasil menghapus *noise* pada citra.

2. Pengujian proses *mapping* nilai intensitas

Pengujian proses ini bertujuan untuk mengetahui bahwa sistem telah berhasil menguatkan nilai kekontrasan citra.

3. Pengujian proses segmentasi citra dengan *thresholding* otsu

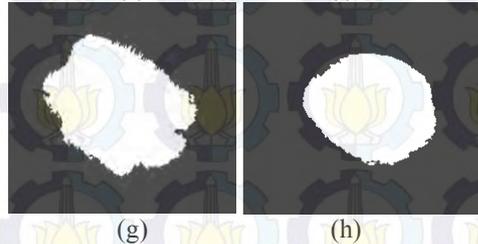
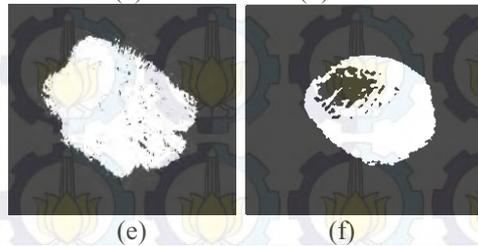
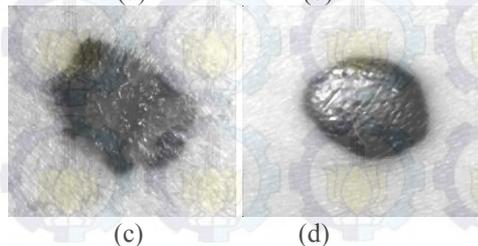
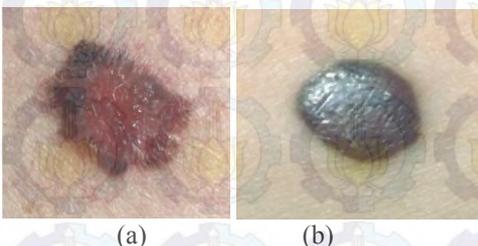
Pengujian proses ini bertujuan untuk mengetahui bahwa system telah berhasil memisahkan antara *foreground* dan *background*.

4. Pengujian proses *flood filling*

Pengujian proses ini bertujuan untuk menghilangkan piksel-piksel yang terisolasi.



Gambar 4. Citra input (a) citra melanoma, (b) citra bukan melanoma



Gambar 5. Hasil tahap pra-pemrosesan (a) Median filtering citra melanoma, (b) median filtering citra bukan melanoma, (c) mapping nilai intensitas citra melanoma, (d) mapping nilai intensitas citra bukan melanoma, (e) segmentasi citra melanoma, (f) segmentasi citra bukan melanoma, (g) flood filling citra melanoma, (h) flood filling citra bukan melanoma

4.3 Pengujian Tahap Ekstraksi Fitur

Pengujian proses ini menggunakan data citra yang telah diproses pada pra-pemrosesan. Tahap ekstraksi fitur ini menggunakan metode ABC yaitu *Asymmetry*, *Border Irregularity*, *Colour Variation* dan GLCM yaitu entropi,

energy, kontras, homogenitas. Data pengujian dari proses ekstraksi fitur dengan citra input pada Gambar 4 dapat dilihat pada Tabel 1 berikut:

Tabel 1. Hasil Proses Ekstraksi Fitur dengan ABC dan GLCM

Ekstraksi Fitur	Citra Melanoma	Citra Bukan Melanoma	
ABC	<i>Asymmetry</i>	0.3016	0.3968
	<i>Border Irregularity</i>	0.1515	0.1764
	<i>Colour Variation</i>	0.1334	0.2147
GLCM	Entropi	3.6171	3.4156
	Energi	0.0005	0.0008
	Kontras	78.6318	110.1517
	Homogenitas	0.3348	0.3416

4.3 Pengujian Tahap Klasifikasi dengan SVM

4.3.1 Proses Training

Pengujian sistem menggunakan 80 data melanoma dan 70 data bukan melanoma. Data dibagi menjadi data pelatihan dan data uji. Data pelatihan dibagi ke dalam 3 kelompok data, yakni :

1. Data pertama yang terdiri dari 64 citra melanoma dan 56 citra bukan melanoma
2. Data kedua yang terdiri dari 56 citra melanoma dan 49 citra bukan melanoma
3. Data ketiga yang terdiri dari 48 citra melanoma dan 42 citra bukan melanoma

Sedangkan data uji merupakan sisa dari jumlah keseluruhan data yang tidak digunakan untuk data training.

Pengujian proses ini bertujuan untuk mengetahui data yang diperlukan dalam proses berikutnya seperti alpha, bias, jumlah support vector, kernel dan sebagainya. Pengujian proses ini menggunakan fitur yang telah diperoleh dari proses ekstraksi fitur dengan ABC dan GLCM.

4.3.2 Proses Testing

Pengujian sistem menggunakan data uji yang merupakan sisa dari jumlah keseluruhan data yang tidak dipakai untuk training. Pengujian proses ini untuk menentukan apakah suatu citra yang diuji merupakan citra melanoma atau citra bukan melanoma. Hasil dari pengujian system dapat dilihat pada Tabel 2.

Tabel 2. Pengaruh Banyak Data *Training* Terhadap Pengujian dan Waktu Komputasi menggunakan Kernel Polynomial.

Data Latih	Akurasi	Waktu Komputasi (detik)	
		<i>Training</i>	<i>Testing</i>
1	90%	71.67 detik	0.513 detik
2	88.88%	62.02 detik	0.520 detik
3	73.33%	52.73 detik	0.516 detik

Berdasarkan Tabel 2. dengan parameter kernel polynomial dapat dilihat bahwa banyaknya data pelatihan dapat mempengaruhi hasil pengujian atau akurasi. Data 1 yaitu data pelatihan 64 citra melanoma dan 56 citra bukan melanoma dengan waktu komputasi 71.76 detik menghasilkan akurasi paling tinggi, yaitu 90%

Dengan menggunakan data pelatihan data 1 yaitu 64 citra melanoma dan 56 citra bukan melanoma pengujian selanjutnya adalah dengan menggunakan parameter yang berbeda yaitu parameter linear, rbf, dan polynomial (dapat dilihat pada Tabel 2.). Hasil dari pengujian ini dapat dilihat pada Tabel 3.

Tabel 3. Pengaruh Parameter Polynomial, Linear dan Rbf Terhadap Pengujian dan Waktu Komputasi dengan menggunakan Data Latih 1

Kernel	Akurasi	Waktu Komputasi (detik)	
		<i>Training</i>	<i>Testing</i>
Polynomial	90%	71.67 detik	0.513 detik
Linear	80%	69.69 detik	0.512 detik
Rbf	73.33%	67.55 detik	0.528 detik

Berdasarkan Tabel 3. parameter polynomial menghasilkan tingkat akurasi paling tinggi yaitu 90%.

V. KESIMPULAN

Berdasarkan eksperimen dan pembahasan terhadap hasil pengujian yang telah dilakukan terhadap sistem klasifikasi citra kanker kulit menggunakan metode *Support Vector Machine*, maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Tugas Akhir ini telah berhasil melakukan klasifikasi citra melanoma dan bukan melanoma dengan menggunakan metode *Support Vector Machine* sebagai metode klasifikasi dengan urutan proses pra-pemrosesan, segmentasi citra, ekstraksi fitur dengan ABC dan GLCM, dan proses klasifikasi SVM.
2. Tugas akhir ini telah berhasil membuat prototipe perangkat lunak untuk klasifikasi citra melanoma dan bukan melanoma.

3. Metode *Support Vector Machine* pada tugas akhir ini dapat mengenali citra kanker kulit melanoma dengan tingkat akurasi terbaik sebesar 90%. Pengujian dengan hasil terbaik ini dilakukan pada data 64 citra melanoma dan 56 citra bukan melanoma dengan waktu komputasi untuk pelatihan terbaik adalah 71.67 detik.
4. Pengujian pada beberapa kernel dihasilkan kesimpulan bahwa kernel yang memberikan hasil paling baik adalah kernel polynomial dengan akurasi 90% pada data latih 1.

DAFTAR PUSTAKA

- [1] Immagulate, I., Vijaya, M.S. 2015. "Categorization of Non-Melanoma Skin Lesion Diseases Using Support Vector Machine and Its Variants". International Journal of Medical Imaging 3(2): 34-40.
- [2] Jaleel, J.A., Salim, S., Aswin, R.B. 2013. "Computer Aided Detection of Skin Cancer". International Conference on Circuits, Power and Computing Technologies.
- [3] Novianti, F.A., Purnami, S.W. 2012. "Analisis Diagnosis Pasien Kanker Payudara Menggunakan Regresi Logistik dan Support Vector Machine (SVM) Berdasarkan Hasil Mamografi". JURNAL SAINS DAN SENI ITS Vol. 1, No. 1, ISSN: 2301-928X
- [4] Listia, R., Harjoko, A. 2014. "Klasifikasi Massa pada Citra Mammogram Berdasarkan Gray Level Cooccurrence Matrix (GLCM)". IJCCS, Vol.8, No.1, pp. 59-68
- [5] Herdha, Tyara (2015). "Klasifikasi Kecatatan Peluru dengan Menggunakan Metode *Support Vector Machine* (SVM)" Tugas Akhir Jurusan Matematika ITS.
- [6] M. Rinaldi (2004). "Pengolahan Citra Digital dengan Pendekatan Algoritmik" Teknik Informatika ITB. Bandung.
- [7] Gonzales, R.C. dan Woods, Richard E. (2010). "Digital Image Processing". New Jersey. Prentice Hall.
- [8] Barhoumi W, Zagrouba E. 2002. "A PRELIMINARY APPROACH FOR THE AUTOMATED RECOGNITION OF MALIGNANT MELANOMA"
- [9] Immagulate I, Vijaya M. S. 2015. "Categorization of Non-Melanoma Skin Lesion Diseases Using Support Vector Machine and Its Variants". International Journal of Medical Imaging.
- [10] Santosa, Budi. (2007). "Data Mining, Teknik Pemanfaatan Data untuk Keperluan Bisnis". Yogyakarta. Graha Ilmu.