

Penerapan Segmentasi Citra untuk Deteksi Kanker Paru-paru pada Citra *CT Scan*

Anzaldi Sulaiman Oemar 13517098
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jalan Ganesha 10 Bandung
anzaldy@gmail.com

Abstract—Kanker yang paling umum dialami pada manusia baik seorang pria maupun wanita adalah kanker paru-paru. Untuk meningkatkan peluang tingkat keselamatan seseorang yang terinfeksi kanker paru-paru adalah mendeteksinya sedini mungkin. Ada tantangan sendiri dalam proses deteksi awal keberadaan kanker paru-paru yaitu sel pada paru-paru saling tumpang tindih sehingga tidak terlihat mengenai awal mula kemunculan sel kanker di paru-paru seseorang. Kemajuan teknologi pemindaian citra CT (*Computed Tomography*) memungkinkan adanya proses deteksi awal keberadaan sel kanker pada paru-paru lebih efektif dan efisien. Salah satu konsep yang cukup populer digunakan adalah dengan memanfaatkan pemrosesan citra (*image processing*). Nantinya, citra CT akan dipisahkan berdasarkan bagian tertentu untuk mendeteksi keberadaan sel kanker melalui proses segmentasi.

Kata kunci—kanker paru-paru, citra *CT Scan*, pemrosesan citra medis, segmentasi citra, klasifikasi

I. PENDAHULUAN

Kanker merupakan suatu sel yang berawal dari sel liar yang tidak terkontrol. Paru-paru menjadi salah satu organ di tubuh manusia yang memungkinkan terjadi kemunculan sel kanker, biasanya pada saluran udara. Sel kanker yang menyebar ke paru-paru bisa berawal dari darah atau cairan limfa di jaringan sekitar paru-paru. Sel kanker yang liar ini tidak akan tumbuh dan berkembang di paru-paru yang sehat, tetapi jika sudah berhasil bereproduksi lebih jauh maka akan menyebabkan tumor. Mencari tahu keberadaan atau tingkat infeksi dari kanker paru-paru sangatlah penting, terutama mengingat dua alasan. Pertama, bisa menentukan metode terapi yang sesuai. Kedua, bisa memperhatikan seberapa banyak sel kanker liar yang sudah tersebar di sekitar organ. Kedua alasan tersebut membantu penentuan metode penanganan terbaik dan ancaman terburuk yang mungkin dihadapi pasien. Mendeteksi keberadaan sel kanker di paru-paru mampu membantu menurunkan tingkat kematian akibat penyakit ini. Sudah ada teknologi pencitraan *CT Scan* yang membantu dunia medis berupa pencitraan suatu organ di dalam tubuh. Menurut beberapa penelitian sebelumnya, citra CT sendiri dapat diandalkan untuk deteksi awal keberadaan kanker paru-paru, tetapi banyaknya kemungkinan citra CT yang tergambar memungkinkan adanya salah penilaian antara masing-masing dokter. Peneliti dari bidang teknologi informasi telah membantu mengkembangkan sistem yang

mampu membaca citra CT secara konsisten dengan representasi yang sama sehingga bisa melanjutkan proses deteksi sel kanker paru-paru lebih konsisten juga, mengurangi kemungkinan salah penilaian. Pengujian lebih lanjut tentu diperlukan untuk menentukan model terbaik dalam membangun sistem deteksi awal kanker paru-paru.

II. LANDASAN TEORI

A. Pemrosesan Citra Digital

Suatu teknik pemrosesan citra merupakan cara penggunaan komputer untuk memanipulasi citra digital. Teknik ini memiliki banyak kelebihan dalam aspek fleksibilitas, adaptif, struktur penyimpanan data, dan komunikasi. Diiringi dengan perkembangan teknik representasi, penyimpanan, atau ukuran citra, maka citra dapat dijaga tetap efisien. Teknik pemrosesan citra memiliki susunan aturan untuk dieksekusi terhadap citra secara asinkron. Citra dua dimensi dan tiga dimensi memungkinkan bisa diproses dalam berbagai dimensi.

Teknik pemrosesan citra mulai dikembangkan sejak era 1960an. Pada saat itu, teknik tersebut banyak digunakan pada berbagai bidang yang sedang berada di awal perkembangannya seperti kesenian, eksplorasi langit, kedokteran, dan televisi. Di era 1970an dengan semakin maju perkembangan sistem komputer, beban dari pemrosesan citra bisa ditekan dan semakin cepat. Sampai era 2000an, pemrosesan citra bukan lagi suatu proses yang kompleks dalam sistem komputer, perkembangannya menjadi lebih sederhana dan lebih cepat.

Pemrosesan citra digital terus berkembang dan sudah banyak diterapkan dalam beberapa tahun terakhir ini pada berbagai bidang seperti penginderaan jauh, kedokteran, fotografi, produksi film dan video, serta keamanan. Inovasi teknologi terbaru mulai bermunculan di bidang pemrosesan citra, terutama dalam hal menerapkan metode segmentasi citra (*image segmentation*).

B. Pengolahan Citra Medis

Pencitraan medis adalah proses menghasilkan gambar struktur bagian dalam dari tubuh manusia untuk kebutuhan studi ilmiah atau penanganan medis beserta pemanatauan terhadap fungsi jaringan interior. Proses ini mampu membantu manajemen dan identifikasi gangguan. Proses ini

juga membuat kumpulan data dari struktur dan fungsional organ yang normal untuk lebih mempermudah pengenalan apabila terjadi anomali.

Ada banyak teknologi yang digunakan untuk mencatat informasi mengenai lokasi dan fungsional dalam tubuh. Berbagai teknik tersebut juga memiliki batasan dibandingkan teknik yang menghasilkan pencitraan. Setiap tahun, jutaan gambar medis berhasil dicatat untuk tujuan diagnosis berbagai penyakit. Sekitar setengah dari catatan tersebut dimanfaatkan melalui metode radiasi ionisasi dan non ionisasi. Citra medis menghasilkan gambar struktur dalam tubuh tanpa memerlukan prosedur yang membebani tubuh. Citra tersebut dihasilkan menggunakan prosesor yang cepat dan konversi sinyal secara logika maupun aritmatika. Sinyal ini kemudian diubah menjadi citra digital. Masing-masing sinyal ini mampu merepresentasikan berbagai jenis jaringan yang berbeda dalam tubuh.

Citra digital memainkan peran penting sehari-hari. Pengolahan citra medis mengacu pada penanganan citra dengan menggunakan komputer. Cara dan operasi yang termasuk dalam proses ini antara lain peningkatan citra, penyimpanan, presentasi, dan komunikasi. Suatu citra adalah suatu fungsi yang mampu membedakan pengukuran secara signifikan dengan melihat karakteristiknya, seperti warna yang dihasilkan atau pantulan sinar.

Citra digital memiliki beberapa kelebihan seperti cepat dan ringan dalam beban komputasi, mudah disimpan dan dikomunikasikan, dapat dinilai secara instan, mudah diperbanyak tanpa mengurangi kualitas, mudah dimanipulasi, dan prosedurnya mudah untuk diulang dalam kasus lain. Sementara itu kekurangannya adalah kurangnya pengawasan hak cipta, sulitnya penyederhanaan dengan tetap menjaga kualitas, kebutuhan penyimpanan yang cukup besar, dan kebutuhan sistem yang semakin kompleks apabila manipulasi yang ingin dilakukan juga cukup kompleks.

C. Segmentasi Citra

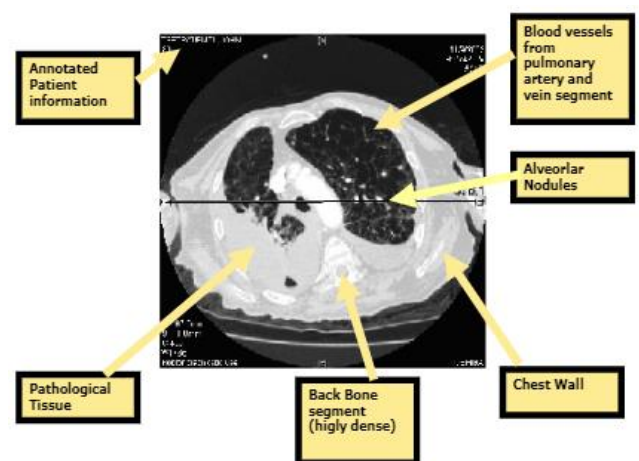
Segmentasi citra mengarah pada proses membagi sebuah citra menjadi beberapa bagian berbeda dengan mengelompokkan satuan pixel bertetangga berdasarkan kriteria kemiripan yang sudah ditentukan di awal. Kriteria kemiripan ini bisa ditentukan menggunakan aspek tertentu atau fitur dari pixel yang ada pada citra. Dengan kata lain, segmentasi adalah teknik klasifikasi titik pixel yang memungkinkan pembentukan bagian-bagian yang mirip pada citra.

Tujuan akhir dari segmentasi citra adalah menemukan bagian citra yang koheren atau objek spesifik. Citra disegmentasi berdasarkan property yang dipilih seperti kecerahan, warna, tekstur, dan sebagainya. Segmentasi membagi citra menjadi sejumlah region yang terhubung, tiap region bersifat homogen berdasarkan properti yang dipilih. Segmentasi citra merupakan tahapan sebelum melakukan image/object recognition, image understanding, dll.

Metode segmentasi sudah menjadi suatu proses yang penting dalam pemrosesan citra medis dan sangat berguna pada banyak aplikasi penerapan. Contoh penerapan pada teknik pemrosesan citra medis antara lain deteksi garis

koroner pada citra angiogram, memantau luka-luka dari kasus sklerosis, simulasi operasi, perencanaan operasi, mengukur volume tumor dan keterkaitannya dengan proses terapi, pemetaan fungsional jaringan, klasifikasi otomatis sel darah, mempelajari perkembangan otak, mendeteksi mikro kalsifikasi pada alat mammogram, ekstraksi citra jantung dari alat angiogram, dan deteksi tumor.

Dalam pencitraan medis, segmentasi adalah salah satu proses penting untuk ekstraksi fitur, pengukuran citra, dan penampilan citra. Pada beberapa penerapan aplikasinya, pencitraan medis bisa bermanfaat untuk klasifikasi citra menjadi bagian anatomis seperti tulang, otot, sel darah, pembuluh, atau bisa juga klasifikasi patologi anatomi lainnya seperti kanker, deformitas jaringan, dan luka sklerosis. Di beberapa studi sebelumnya, tujuan dari pencitraan medis dengan segmentasi citra adalah memisahkan bagian dengan pewarnaan zat putih, zat abu, cairan serebrospinal pada otak, atau bisa juga pada bagian yang lebih spesifik seperti kanker payudara dari citra MRI.



Gambar 1. Pemrosesan citra medis beserta klasifikasinya

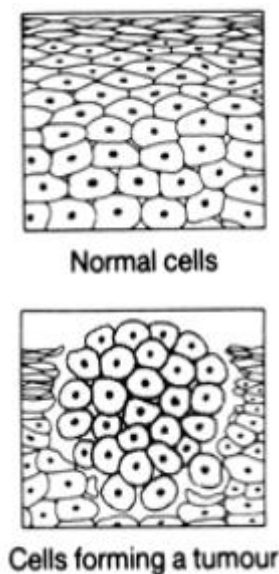
D. Kanker Paru

Organ dan jaringan pada tubuh manusia terbentuk oleh blok atau unit kecil yang disebut sel. Sel di setiap bagian tubuh mungkin terlihat berbeda dan bekerja dengan cara yang berbeda juga, tapi sebagian besar dari sel tersebut melakukan reproduksi dengan cara yang serupa. Secara konsisten, sel akan tumbuh menua dan mati, sel-sel baru akan terbentuk untuk menggantikan mereka yang mati. Pada kondisi normal, pembelahan dan pertumbuhan sel biasanya teratur dan bisa dikontrol tetapi jika proses ini menjadi tidak terkontrol karena suatu alasan, sel-sel akan terus membelah dan berkembang menjadi sebuah benjolan yang biasa disebut tumor.

Tumor bisa mungkin bersifat jinak atau ganas. Kanker merupakan sebutan untuk tumor yang bersifat ganas. Tumor yang ganas mengandung sel kanker yang memiliki kemampuan untuk menyebar melewati batas lokasi asli dari sel. Jika dibiarkan tidak ditangani, sel kanker ini mungkin menyerbu, membahayakan, bahkan menghancurkan jaringan di sekitarnya. Suatu waktu, sel memisahkan diri dari sel kanker awal dan menyebar menuju organ lainnya di dalam

tubuh dengan mengalir melalui aliran darah atau sistem limfatik. Ketika sel kanker yang ganas tersebut mencapai lokasi baru di dalam tubuh, mereka mungkin saja membelah diri dan membentuk tumor baru, sering disebut sebagai metastasis atau tumor sekunder. Hal penting yang perlu disadari bahwa kanker bukanlah satu penyakit khusus yang bisa ditangani atau diobati dengan cara yang sama. Ada sebanyak 200 lebih jenis kanker, masing-masing sudah memiliki klasifikasi namanya sendiri termasuk cara penanganannya sendiri.

Kanker paru-paru adalah tumor ganas yang terletak di paru-paru serta merupakan jenis penyakit kanker paling mematikan secara global baik untuk kategori laki-laki maupun perempuan, tingkat prevalensi kanker paru-paru mencapai 11.4 persen. Data dari Global cancer statistics (Globocan) 2020 mencatat bahwa kematian karena kanker paru di Indonesia meningkat menjadi 30.843 orang dengan kasus baru mencapai 34.783 kasus. Jumlah kematian tahunan hampir sama dengan jumlah kasus baru. Gejala kanker paru-paru tidak terlihat sampai penyakit ini sudah cukup parah menyerang tubuh seseorang. Faktor utama yang bisa meningkatkan risiko kanker paru-paru adalah polusi udara seperti paparan asap rokok, tinggal atau bekerja di area pabrik dengan polusi tinggi.



Gambar 2. Perbandingan sel normal dan sel liar

E. Penelitian Terdahulu

Beberapa peneliti sudah mengajukan dan melakukan implementasi deteksi kanker paru-paru menggunakan beberapa pendekatan yang berbeda berdasarkan pemrosesan citra (*image processing*) dan pembelajaran mesin.

Aggarwal, Furquan, dan Kalra [2] mengajukan suatu model yang mampu melakukan klasifikasi berdasarkan struktur anatomi paru-paru yang normal dan kemunculan nodul pada paru-paru. Mereka menggunakan ekstraksi fitur geometri, statistik, dan level keabuan. *Linear Discriminant Analysis* (LDA) digunakan sebagai algoritma untuk klasifikasi dan menghitung *threshold* optimal ketika segmentasi. Sistem ini menghasilkan akurasi senilai 84%,

sensitivitas senilai 97%, dan tingkat spesifik senilai 53%. Walaupun sistem ini berhasil melakukan deteksi awal kanker melalui nodul paru-paru, nilai akurasi yang diperoleh tetap terasa kurang. Tidak ada teknik *machine learning* khusus yang digunakan untuk klasifikasi, hanya segmentasi citra secara sederhana yang digunakan pada sistem ini.

Sangamithraa dan Govindaraju [3] menggunakan algoritma *supervised* yaitu k-means baik untuk klusterisasi atau segmentasi. Sistem ini mengelompokkan setiap pixel pada dataset berdasarkan karakteristik tertentu. Untuk melakukan klasifikasi, model pada sistem ini mengimplementasikan algoritma *backpropagation*. Fitur yang dibutuhkan seperti entropi, korelasi, homogenitas, PSNR, SSIM dilakukan ekstraksi menggunakan metode GLCM (*Gray-Level Co-occurrence Matrix*). Sistem ini menghasilkan akurasi bernilai 91%. *Median filtering* digunakan di sistem ini dalam melakukan *image preprocessing*, hal ini bermanfaat supaya bisa menghilangkan *noise* dan menghasilkan model yang meningkat nilai akurasinya.

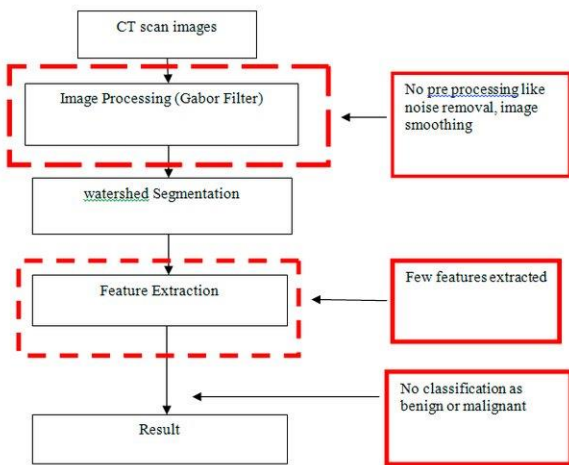
Ignatious dan Joseph [4] mengembangkan suatu sistem menggunakan *watershed segmentation*. Pada tahap *pre processing*, mereka menggunakan *Gabor filtering* untuk meningkatkan kualitas gambar. Hasil dan akurasi dari sistem ini dibandingkan dengan model serupa yang pernah dibangun yaitu model *neural fuzzy* dan model *region growing*. Akurasi dengan model ini diperoleh sebesar 90% yang artinya lebih tinggi daripada model apabila proses segmentasi menggunakan *neural fuzzy* dan *region growing*. Kelebihan dari model ini adalah penggunaan *marker* yang bisa diatur pada algoritma *watershed segmentation*, hal ini mampu memperbaiki masalah segmentasi yang melewati batas. Kekurangan yang diperoleh adalah sistem ini tidak mengklasifikasikan sel kanker sebagai sel kanker yang ganas atau jinak, nilai akurasi yang diperoleh juga masih dirasa kurang tinggi bagi kalangan peneliti. Beberapa perbaikan dan kontribusi dari peneliti lain untuk menangani model ini dapat dilakukan apabila ingin meningkatkan akurasinya.

Gonzalez dan Ponomaryvo [5] mengajukan sistem yang mampu mengklasifikasikan kanker paru-paru sebagai sel kanker yang jinak dan ganas. Sistem ini menggunakan informasi apriori dan HousefieldUnit (HU) untuk menghitung *Region of Interest* (ROI). Fitur dari bangun datar seperti luas, eksentrisitas, kelengkungan, dimensi fractal, serta fitur pada tekstur seperti rata-rata, variansi, energi, entropi, distorsi, kontras, dan kelembutan akan diambil untuk pelatihan dan klasifikasi melalui *support vector maching* (SVM). Selanjutnya, SVM akan mengidentifikasi sifat dari nodul yang jinak atau ganas. Tentu saja hal tersebut menjadi kelebihan dari model yang dibangun pada sistem ini, tetapi kekurangannya adalah adanya batasan perlunya informasi mengenai ROI. Klasifikasi antara sel kanker jinak atau ganas menggunakan SVM akan berguna sehingga mungkin dicoba untuk membangun model yang ideal.

Berdasarkan beberapa penelitian terkait yang sudah dilakukan, apabila mempertimbangkan akurasi dan kelebihan yang diperoleh, maka sistem yang diajukan oleh Ignatious dan Joseph merupakan yang paling baik. *Gabor filter*

digunakan untuk meningkatkan kualitas gambar, menggunakan *marker* dalam *watershed segmentation*, hingga berhasil melakukan deteksi awal kanker paru-paru berdasarkan struktur nodul. Model ini juga berhasil mengambil fitur seperti luas, keliling, dan eksentrisitas pada nodul sel kanker. Dapat dilihat juga bahwa model terbaik ini mencapai nilai akurasi sebesar 90% tetapi masih memiliki beberapa batasan sebagai berikut :

- Hanya sedikit fitur yang berhasil diekstrak pada nodul sel kanker.
- Tidak ada proses *pre processing* seperti penghapusan *noise*, atau pelembutan citra yang kemungkinan membantu deteksi sel kanker lebih akurat.
- Belum ada klasifikasi sel kanker termasuk kategori jinak atau ganas.



Gambar 3. Model terbaik dengan kekurangannya (Ignatious dan Joseph, 2015)

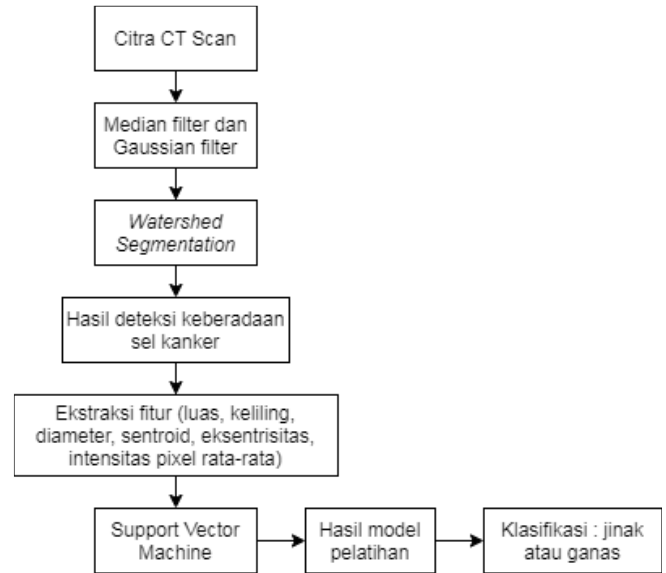
III. IMPLEMENTASI DAN PEMBAHASAN

A. Rancangan Model

Perubahan yang akan dilakukan dari model terbaik yang pernah diteliti dengan rancangan model yang baru akan digambarkan setelah ini. Dibandingkan menggunakan Gabor filter, tahap *pre processing* pada model baru ini akan mengimplementasikan median filter dan Gaussian filter. Setelah *pre processing*, citra terkait akan dijalankan segmentasi menggunakan *watershed*. Beberapa proses yang sudah dilalui ini akan memberikan citra dengan penanda pada nodul sel kanker paru-paru.

Sebagai tambahan dari fitur awal yaitu luas, keliling, dan eksentrisitas, fitur lain seperti sentroid, diameter, dan intensitas rata-rata juga diekstrak untuk mendeteksi keberadaan nodul sel kanker. Model terbaik akan ditutup dengan ekstraksi fitur setelah berhasil mendeteksi nodul sel kanker paru-paru yaitu dengan perhitungan akurasi, tetapi pada rancangan model ini akan ditambahkan tahap klasifikasi untuk memperoleh kategori sel kanker yang jinak atau ganas. Proses klasifikasi ini menggunakan metode SVM. Fitur-fitur yang berhasil diperoleh pada proses sebelumnya akan digunakan sebagai data latih. Terakhir,

klasifikasi kategori sel kanker baik jinak maupun ganas akan bisa dihasilkan berhasillan model pelatihan sebelumnya.



Gambar 4. Rancangan model baru perbaikan

1) *Image pre processing*

Pertama, pada tahap ini yaitu *pre processing* akan menggunakan median filter pada citra CT Scan yang sudah diubah menjadi citra *grayscale*. Pada citra awal, tentu saja mungkin mengalami beberapa *noise* yang menyebabkan kesalahan ketika mendeteksi nodul sel kanker karena warnanya yang mirip. Maka dari itu, diperlukan penghapusan *noise* untuk bisa mendeteksi keberadaan nodul sel kanker menjadi lebih akurat. Median filter mampu mengatasi hal ini pada citra CT Scan. Setelah median filter, ada penerapan Gaussian filter untuk melembutkan citra dan menghapus *noise* dengan sifat yang berbeda.

2) Segmentasi

Proses ini adalah pencarian objek dan batasan sekelilingnya yang akan membantu perolehan *region of interest* pada citra. Proses ini akan membagi citra menjadi bagian wilayah tertentu dengan tujuan memperoleh informasi berdasarkan ilmu anatominya. Dalam proses deteksi kanker paru-paru, proses ini akan membatasi nodul sel kanker pada citra CT. Sesuai rancangan model yang akan dicoba, implementasi untuk segmentasi adalah menggunakan *watershed segmentation*. Manfaat dari proses segmentasi ini adalah menghasilkan fitur lain yaitu adalah mendeteksi nodul sel kanker yang bersentuhan dengan objek lain pada citra. Kelebihannya adalah mampu mendeteksi bagian nodul sel kanker yang lebih tepat setelah melalui tahap segmentasi.

3) Ekstraksi Fitur

Tahap ekstraksi fitur akan mengambil fitur-fitur seperti luas, keliling, sentroid, eksentrisitas, dan rata-rata

intensitas keabuan. Fitur-fitur ini yang akan dimanfaatkan untuk mengembangkan model pelatihan, bisa juga membantu tahap klasifikasi lanjutan.

4) Klasifikasi

Proses ini menghasilkan klasifikasi suatu sel kanker akan termasuk sebagai kategori sel kanker jinak atau sel kanker ganas. SVM menyatakan suatu fungsi yang bisa mengklasifikasikan data tepat menjadi dua kelas. Fungsi dari SVM yang digunakan adalah $D(x) = w^T x_i + b$, dengan x_i adalah masukkan dari data latih, w^T adalah vektor berdimensi M, dan b adalah nilai bias. Nilai i merupakan iterasi $i = 1$ sampai M.

$$D(x) = w^T x_i + b \geq 1 \rightarrow y_i = 1$$

$$D(x) = w^T x_i + b \leq -1 \rightarrow y_i = -1$$

Setelah melalui keempat proses besar di rancangan tersebut, diperoleh bahwa kelebihan dari rancangan model tersebut adalah sebagai berikut :

- Meningkatkan akurasi dalam mendeteksi nodul sel kanker dibandingkan model terbaik oleh Ignatious dan Joseph.
- Mampu mengklasifikasikan sel kanker sebagai kelas jinak atau ganas.
- Menghapus *salt-pepper noise* dan *speckle noise* yang mungkin memberikan salah deteksi pada citra.

Walaupun banyak kelebihan, rancangan model ini tidak terlepas dari beberapa kekurangan, yaitu :

- Terdapat peningkatan nilai akurasi, tetapi masih kurang tinggi, diharapkan mendekati 100%.
- Baru sampai bisa mendeteksi klasifikasi sel kanker jinak atau ganas. Perlu dikembangkan lebih jauh untuk kebutuhan medis sampai mendeteksi stadium I, II, III, IV, dan V.

B. Implementasi

Untuk bagian implementasi, citra CT Scan pasien secara nyata diperoleh dari arsip *Lung Image Database Consortium* (LIDC). Arsip ini adalah basis data yang berisi tentang citra CT Scan dari kanker paru-paru untuk pembelajaran, pelatihan, pengujian, dan evaluasi dari diagnosis dengan bantuan sistem komputer dalam mendeteksi kanker paru-paru. Basis data ini diinisiasi oleh *National Cancer Institute*. Secara mentah, basis data ini berisi 1018 kasus yang didaftarkan oleh tujuh pusat peneliti kesehatan dan delapan perusahaan pencitraan medis. Gambar-gambar pada dataset tersebut memiliki format DICOM dengan ukuran piksel sebesar 512x512. Format DICOM ini akan diubah menjadi citra abu bersifat JPEG menggunakan aplikasi MicroDicom.

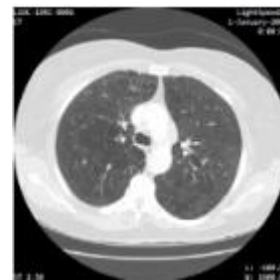
Rancangan model akan diimplementasikan menggunakan software MATLAB versi R2021a. MATLAB adalah salah satu kaskas dan bahasa pemrograman untuk analisis dan penelitian. Semua proses baik ekstraksi fitur atau pendeteksi diimplementasikan langsung di MATLAB, dengan klasifikasi diimplementasikan menggunakan *machine learning toolbox* yang tersedia pada kaskas MATLAB. Kaskas

pembelajaran yang disebutkan terakhir membantu pelatihan model untuk klasifikasi menjadi lebih mudah dan lebih cepat. Cross validation lima tingkat digunakan untuk mencegah terjadinya overfit ketika proses pelatihan. Dalam proses pelatihan, digunakan sebanyak 18 sampel citra dari LIDC sebagai bahan klasifikasi, serta hasil pelatihan akan dilakukan validasi menggunakan 5 citra yang menghasilkan 15 citra nodul sel kanker. Beberapa tantangan yang dihadapi ketika melakukan implementasi rancangan model, antara lain :

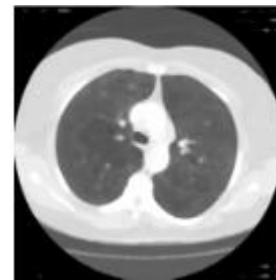
- Basis data LIDC memiliki ukuran yang sangat besar yaitu sekitar 125GB, maka dari itu dipilih hanya beberapa sampel.
- Deskripsi lanjutan mengenai citra kanker di basis data LIDC dijelaskan pada berkas XML yang terlampir, tetapi relatif lebih sulit dipahami.

C. Hasil dan Evaluasi

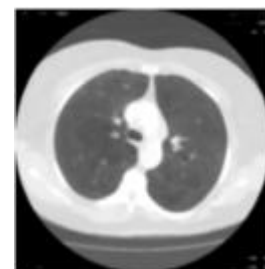
Gambar-gambar berikut menunjukkan masing-masing citra original, *grayscale*, hasil median filter, hasil Gaussian filter, citra biner, hasil segmentasi *watershed*, dan penanda di kanker.



Gambar 5. Citra CT asli dari basis data



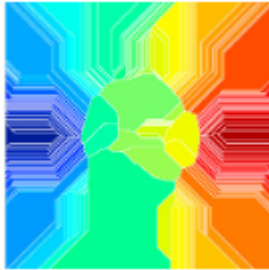
Gambar 6. Citra hasil median filter



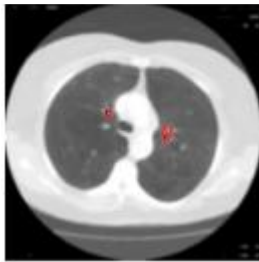
Gambar 7. Citra hasil Gaussian filter



Gambar 8. Citra biner



Gambar 9. Citra hasil segmentasi *watershed*



Gambar 10. Sel kanker yang berhasil dideteksi

Hasil dari rancangan model kali ini berhasil memperoleh tingkat akurasi senilai 93.3%, sedikit peningkatan dibandingkan model terbaik pada penelitian-penelitian sebelumnya yang hanya mencapai 88.4% ketika diuji menggunakan sampel gambar yang sama dari basis data tersebut. Tingkat presisi yang mampu diperoleh juga meningkat dari 87.5% menjadi 92.85%, sementara nilai recall tidak berubah.

Dari nodul sel kanker yang berhasil dideteksi, fitur seperti luas, keliling, sentroid, diameter, eksentrisitas, dan intensitas pixel rata-rata diekstrak. Hasil ekstraksi fitur tersebut akan digunakan untuk pelatihan SVM dan model hasil pelatihan terbentuk. Waktu yang dihabiskan untuk model klasifikasi dengan memanfaatkan sampel basis data yang sudah dijelaskan pada bagian sebelumnya adalah 7.62 detik. Model klasifikasi ini memperoleh nilai evaluasi berupa waktu prediksi untuk model yang sudah dilatih adalah 310 observasi per detik.

Faktor	Rancangan model	Model terbaik
Nodul sel kanker terdeteksi	28	24
True Positive	26	21
True Negative	2	2
False Positive	2	3
False Negative	0	0
Akurasi	28/30 = 93.3%	23/26 = 88.4%
Presisi	26/28 = 92.85%	21/24 = 87.5%
Recall	26/26 = 100%	21/21 = 100%

Tabel 1. Analisis perbandingan model terbaik dan model rancangan

Diperlukan validasi dari model yang sudah dikembangkan. Seperti yang sudah dijelaskan sebelumnya, validasi mengambil 5 citra CT dan diperoleh 15 nodul sel kanker. Tahap validasi memiliki nilai perbandingan untuk memeriksa ketepatan dari prediksi berdasarkan model.

IMAGE_ID	ID_NODUL	Kelas	Prediksi	Koreksi
0	0	Ganas	Ganas	TRUE
	1	Ganas	Ganas	TRUE
	2	Ganas	Ganas	TRUE
1	0	Jinak	Jinak	TRUE
	1	Jinak	Jinak	TRUE
	2	Jinak	Jinak	TRUE
2	0	Jinak	Jinak	TRUE
	1	Ganas	Ganas	TRUE
	2	Ganas	Ganas	TRUE
3	0	Jinak	Ganas	FALSE
	1	Jinak	Ganas	FALSE
	2	Jinak	Jinak	TRUE
	3	Jinak	Jinak	TRUE
4	0	Ganas	Ganas	TRUE
	1	Ganas	Ganas	TRUE

Tabel 2. Hasil klasifikasi berdasarkan model rancangan

Implementasi dari rancangan model dengan pengujian klasifikasi yang bisa membedakan sel kanker ganas atau jinak berdasarkan data validasi menghasilkan nilai akurasi setinggi $13/15 = 86.6\%$. Dengan demikian, fungsionalitas klasifikasi sel kanker yang ganas dan jinak berhasil juga diimplementasikan, karena fungsionalitas ini belum diimplementasikan pada model terbaik yang dirancang oleh Ignatious dan Joseph (2015).

IV. KESIMPULAN DAN SARAN PENGEMBANGAN

Sistem terbaik yang ditemukan pada penelitian sebelumnya masih belum mencapai tingkat akurasi yang memuaskan (kurang tinggi). Pada penelitian yang ditemukan, masih jarang yang mampu mengklasifikasikan sifat atau kategori sel kanker termasuk jinak atau ganas dan pada sistem yang direncanakan ini terdapat fungsionalitas untuk mengklasifikasikan kategori sel kanker tersebut.

Rancangan model pada penelitian ini diharapkan mampu meningkatkan kualitas deteksi keberadaan nodul sel kanker serta mengklasifikasikannya terhadap kategori tertentu yaitu sel kanker jinak atau ganas. Usaha untuk meningkatkan kualitas deteksi adalah tambahan proses *filtering* dan *watershed segmentation*, sementara untuk metode klasifikasi digunakan SVM karena algoritma SVM cukup ideal apabila ingin membagi dua kelas yang cukup signifikan. Model perbaikan yang sudah dirancang pada penelitian ini memperoleh tingkat akurasi sebesar 93.3%, lebih besar dari model terbaik pada literatur sebelumnya yaitu hanya 88.4%. Pada fungsionalitas klasifikasi, mampu diperoleh tingkat akurasi sebesar 86.6%.

Saran pengembangan yang mungkin bisa dilakukan adalah tambahan pendeteksi serta klasifikasi untuk kanker paru-paru stadium I, II, III, IV, dan V. Perbaikan pada tahap *pre processing* juga bisa dilakukan dengan harapan meningkatkan akurasi dalam mendeteksi keberadaan nodul sel kanker paru-paru.

V. UCAPAN TERIMA KASIH

Penulis mengucapkan banyak terima kasih kepada seluruh pihak yang telah membantu penulis baik secara langsung, dan tidak langsung selama pembuatan makalah ini. Penulis juga mengucapkan terima kasih kepada dosen mata kuliah IF4073 Interpretasi dan Pengolahan Citra, Dr. Ir. Rinaldi Munir, yang telah memberikan bimbingan serta wawasan yang sangat bermanfaat dalam penulisan makalah ini.

REFERENCES

- [1] Munir, Rinaldi. 2020. Slide Perkuliahan IF4073 Interpretasi dan Pengolahan Citra : Segmentasi Citra.
- [2] Aggarwal, T., Furqan, A., & Kalra, K. (2015) "Feature extraction and LDA based classification of lung nodules in chest CT scan images." 2015 International Conference On Advances In Computing, Communications And Informatics (ICACCI).
- [3] Sangamithraa, P., & Govindaraju, S. (2016) "Lung tumour detection and classification using EK-Mean clustering." 2016 International Conference On Wireless Communications, Signal Processing And Networking (Wispnet).
- [4] Ignatious, S., & Joseph, R. (2015) "Computer aided lung cancer detection system." 2015 Global Conference On Communication Technologies (GCCT).
- [5] Rendon-Gonzalez, E., & Ponomaryov, V. (2016) "Automatic Lung nodule segmentation and classification in CT images based on SVM." 2016 9Th International Kharkiv Symposium On Physics And Engineering Of Microwaves, Millimeter And Submillimeter Waves (MSMW).

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 25 Mei 2021



Anzaldi Sulaiman Oemar

13517098