

Breast Cancer Detection

Habib Alfarobi

Departemen Teknik Elektro
Fakultas Teknik, Universitas Indonesia
Depok, Indonesia
habib.alfarobi@ui.ac.id

Anissa Maharani

Departemen Teknik Elektro
Fakultas Teknik, Universitas Indonesia
Depok, Indonesia
anissa.maharani81@ui.ac.id

Nabila Nurul Khalisha

Departemen Teknik Elektro
Fakultas Teknik, Universitas Indonesia
Depok, Indonesia
nabila.nurul82@ui.ac.id

Josh Frederich

Departemen Teknik Elektro
Fakultas Teknik, Universitas Indonesia
Depok, Indonesia
josh.frederich@ui.ac.id

Vidi Miranda Pakpahan

Departemen Teknik Elektro
Fakultas Teknik, Universitas Indonesia
Depok, Indonesia
vidi.miranda@ui.ac.id

Hilda Auliana

Departemen Teknik Elektro
Fakultas Teknik, Universitas Indonesia
Depok, Indonesia
hilda.auliana@ui.ac.id

Abstract— Banyak wanita di dunia juga di Indonesia meninggal karena kanker payudara dan deteksi dini akan meningkatkan peluang kelangsungan hidup pasien. Mamografi adalah pencitraan medis khusus yang menggunakan sistem sinar X dosis rendah untuk melihat bagian dalam payudara. Pada umumnya ahli radiografi akan menganalisis gambar mamogram untuk mendeteksi dan mengklasifikasikan gambar tersebut, namun metode manual ini sering kali menyebabkan kesalahan diagnosis. Oleh karena itu, pembahasan ini bertujuan untuk merancang dan menganalisa algoritma untuk mendeteksi kanker payudara dari dataset gambar mamografi yang disimulasikan dengan software Matlab. Pada algoritma ini, gambar mammogram ditingkatkan dengan menerapkan filter Wiener untuk menghilangkan noise dan menerapkan *Contrast Limited Adaptive Histogram Equitation* (CLAHE) untuk meningkatkan kualitas gambar mammogram. Kemudian ROI (*region of interest*) dari gambar ditentukan menggunakan teknik segmentasi *Gaussian Mixture Model* (GMM) berbasis K-means, kemudian gambar diekstraksi kembali melalui analisis tekstur menggunakan metode GLCM. Setelah itu, gambar akan diproses kembali dengan sistem klasifikasi *Probabilistic Neural Network* (PNN) yang dapat mendiagnosis jenis tumor menjadi tumor ganas (*malignant*), tumor jinak (*benign*), atau tidak terdapat tumor apapun (*normal*). Berdasarkan hasil evaluasi menggunakan *confusion matrix* dan model ROC diperoleh tingkat akurasi 100% yang menandakan sistem ini dapat bekerja dengan baik untuk pendeteksian kanker payudara,

Keywords—kanker payudara, filter Wiener, CLAHE, segmentasi GMM, PNN

I. PENDAHULUAN

A. Latar Belakang

Kanker terjadi ketika terdapat sel abnormal yang tumbuh dengan tidak terkendali, biasanya menyerang bagian tubuh yang berdekatan, dan dapat menyebar ke

organ lain. Kanker dapat terjadi di hampir semua organ atau jaringan tubuh, salah satunya pada jaringan yang menyusun payudara [1]. Kanker payudara merupakan tumor ganas yang berada pada jaringan payudara. Kanker payudara disebabkan oleh kegagalan dalam koordinasi fungsi gen. Tumor ini dapat timbul dari epitel duktus ataupun lobulus payudara dan merupakan salah satu jenis kanker yang paling umum ditemui di Indonesia [2]. Berdasarkan data dari The Global Cancer Observatory (Globocan) secara global terdapat 2.261.419 kasus baru dari kanker payudara pada tahun 2020, dan dengan jumlah kematian sebanyak 684.996 kasus [3]. Sementara di Indonesia berdasarkan data dari Badan Registrasi Kanker Perhimpunan Dokter Spesialis Patologi Indonesia (IAPPI) dan Yayasan Kanker Indonesia (YKI)), diperkirakan angka kejadian kanker payudara pada tahun 2010 adalah 12/100.000 wanita, dengan mortalitas yang cukup tinggi yaitu 27/100.000 atau 18 % dari total kasus. Lebih dari 80% kasus kanker payudara di Indonesia ditemukan pada stadium lanjut, dimana upaya pengobatan sulit dilakukan [2]. Oleh karena itu diagnosis dini kanker payudara sangat penting untuk meningkatkan peluang kelangsungan hidup pasien.

Mamografi adalah pencitraan medis khusus yang menggunakan sistem sinar X dosis rendah untuk melihat bagian dalam payudara. Pemeriksaan mamografi, yang disebut mammogram, merupakan tes yang digunakan untuk deteksi dini dan diagnosis kanker payudara [4]. Pada umumnya ahli radiologi memeriksa gambar hasil mamografi untuk mendeteksi dan mengklasifikasikan citra menjadi ganas (*malignant*) dan jinak (*benign*). Namun, metode ini membutuhkan ahli radiologi yang sangat terampil yang umumnya tidak tersedia, terutama bagi orang-orang di daerah terpencil dan miskin. Selain itu, terdapat risiko tinggi terjadinya human error dalam pemeriksaan manual, oleh karena itu dibutuhkan sistem Computer Aided Detection (CAD) yang dapat membantu ahli radiologi dalam mendiagnosis dan membantu menurunkan tingkat kesalahan diagnosis.

B. Studi Literatur

Probabilistic Neural Network (PNN) adalah kelompok *artificial neural network* yang terbentuk dengan pendekatan milik Parzen untuk merancang kelompok fungsi estimator densitas probabilitas yang dapat melakukan pendekatan untuk teorema Bayes tanpa adanya gejala [5]. PNN ini lebih berorientasi, tapi tidak eksklusif, kepada masalah pengklasifikasian. Hal tersebut dikarenakan awalnya PNN dibuat untuk memberikan sarana yang dapat melatih permasalahan dunia nyata lebih cepat dibandingkan dengan propagasi balik [6]. Maka, PNN ini tidak menggunakan propagasi balik pada pelatihan komputasi tapi menggunakan representasi untuk setiap pola data untuk melihat kesamaan dari pola *input* dengan pola data. Beberapa kategori permasalahan yang dicoba diselesaikan dengan menggunakan PNN adalah [5]:

- Klasifikasi pola data diam yang berlabel
- Klasifikasi pola data yang memiliki fungsi densitas probabilitas yang berbeda waktu
- Berurusan dengan bentuk gelombang sebagai pola data pada pemrosesan sinyal
- Algoritma yang bekerja dengan kumpulan data yang tidak berlabel

Kelebihan dari PNN adalah waktu pelatihan yang lebih cepat dibandingkan dengan propagasi balik, tidak ada masalah *local minima*, dan memiliki jaminan akan pengklasifikasi optimal saat ukuran dari *training set* meningkat [7]. Meskipun waktu pelatihannya cepat, PNN memakan waktu lama untuk pengenalan tergolong lama akibat dari jumlah pola sampel yang banyak yang harus dikategorikan ke setiap neuron [8].

Mekanisme dari PNN berawal dari *probability density function* (PDF) yang dikemukakan oleh Pazen untuk variabel acak X_1, X_2, \dots, X_n dengan PDF $f(x)$ yang tidak diketahui adalah

$$f_n(x) = \frac{1}{nh(n)} \sum_{j=1}^n K\left(\frac{x - X_j}{h(n)}\right) \quad (I.1)$$

dengan $h(n)$ adalah fungsi yang memenuhi syarat:

$$\lim_{n \rightarrow \infty} h(n) = 0 \quad (I.2)$$

Fungsi K pada persamaan I.1 harus memenuhi kondisi-kondisi sebagai berikut:

$$1. \sup_{-\infty < y < \infty} K(y) < \infty \quad (I.3)$$

$$2. \int_{-\infty}^{\infty} |K(y)| dy < \infty \quad (I.4)$$

$$3. \lim_{y \rightarrow \infty} |yK(y)| dy = 0 \quad (I.5)$$

$$4. \int_{-\infty}^{\infty} K(y) dy = 1 \quad (I.6)$$

lalu

$$\lim_{n \rightarrow \infty} E[f_n(x) - f(x)]^2 = 0 \quad (I.7)$$

Persamaan terakhir menunjukkan bahwa kelompok PDF milik Parzen dapat mengestimasi PDF dengan variabel X yang tidak diketahui sebagai $n \rightarrow \infty$. Specht menggunakan konsep ini pada jurnalnya di tahun 1990 untuk memformulasikan pendekatan untuk pengklasifikasian pola data dengan kelas

yang tidak diketahui dari pengaturan pola awal yang diketahui. Lebih jelasnya, semisal sebuah rangkaian variabel independen yang terdistribusikan rata adalah $\{X_i, Y_i\}$, $i = 1, 2, \dots, n$, ini dikenal sebagai rangkaian training, $Y_i \in \{1, 2, \dots, M\}$ adalah kelas yang diasosiasikan dengan pola $X_i \in A \subset \mathbb{R}^p$. Variabel acak X_i memiliki PDF $f_m(x)$ yang dijuluki “class conditional density” (Rutkowski (2004)). Definisikan fungsi diskriminan dari kelas j sebagai:

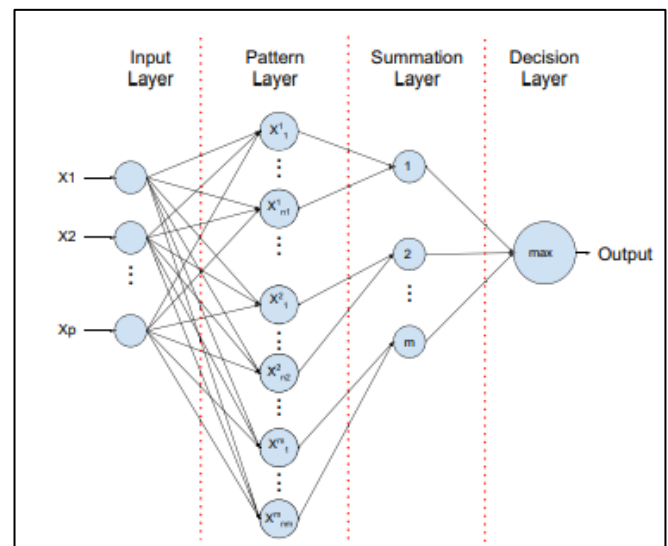
$$d_j(x) = p_j f_j(x) \quad (I.8)$$

Dimana p_j adalah probabilitas awal dari kelas j dengan rumus:

$$p_j = \frac{n_j}{n} \quad (I.9)$$

Dan n adalah jumlah pola dalam rangkaian pelatihan dan n_j adalah jumlah pola yang ada pada kelas j . $f_j(x)$ adalah PDF dari variabel x . Kelas dari pola x yang tidak diketahui kelasnya akan dideterminasikan sebagai m jika:

$$d_m(x) > d_j(x) \quad \forall j \in \{1, 2, \dots, M\}, j \neq m \quad (I.10)$$



Gambar 1. Struktur PNN

Akibat dari nilai $f_j(x)$ biasanya tidak diketahui, estimator sering digunakan untuk menggantikan posisi sehingga dengan menggunakan persamaan 1.1, $f_j(x)$ dapat diestimasi sebagai:

$$\hat{f}_{j,n}(x) = \frac{1}{h_{nj}^p} \sum_{i=1}^{n_j} K(x, X_i^{(j)}) \quad (I.11)$$

Dengan $X^{(j)}$ adalah pola ke i dari rangkaian pelatihan pada kelas j . p adalah dimensi dari input vector dan pola. Gambar 1. merupakan gambaran dari seluruh persamaan sejauh ini. Jaringan yang dikemukakan Specht pada tahun 1990 memiliki empat lapisan:

- Lapisan input
- Lapisan pola
- Lapisan penjumlahan
- Lapisan keputusan

Lapisan pola hanyalah lapisan yang berisi persimpangan untuk mendapatkan input vector dan mendistribusikannya ke lapisan berikutnya. Sebagian besar kalkulasi berada pada lapisan pola, setiap pola memiliki pelatihan yang berbeda-beda untuk setiap unit berupa proses yang terdiri dari dua tahap sebelum diteruskan kepada input berikutnya. Pada tahap pertama, pola unit membandingkan polanya dengan vector input. Ini merupakan jarak yang diukur untuk menjelaskan kemungkinan fungsi-fungsinya. Definisikan $D(x, X_i^{(j)})$ sebagai jarak antara input vector x dengan pola ke- i pada rangkaian pelatihan kelas j :

- Dot product, didefinisikan sebagai

$$D(x, X_i^{(j)}) = \sum_{k=1}^p x_k \cdot X_{i,k}^{(j)} \quad (I.12)$$

Dengan x_k adalah elemen ke- k dari input vector dan $X_{i,k}^{(j)}$ adalah elemen ke- k dari pola ke- i dari rangkaian training pada kelas j . Kedua vector dinormalisasikan sebagai unit panjang sebelum dilakukan dot product.

- Jarak Euclidean didefinisikan sebagai

$$D(x, X_i^{(j)}) = \sqrt{\sum_{i=1}^p (x_i - x_i^{(j)})^2} \quad (I.13)$$

- Jarak Manhattan:

$$D(x, X_i^{(j)}) = \sum_{i=1}^p |x_i - x_i^{(j)}| \quad (I.14)$$

Tahap kedua adalah mengaplikasikan fungsi nonlinear (kernel) kedalam jarak antara input dengan pola pelatihan. Definisikan kernel sebagai $K(x, u) = K(D(x, u))$. Fungsi kernel harus memenuhi empat kondisi yang dimasukkan untuk fungsi berpotensi:

1. Kernel harus memiliki nilai maksimum pada $x = u$
2. $\lim_{D(x,u) \rightarrow \infty} K(x, u) = 0$ (I.15)
3. $K(x, u)$ adalah kontinu pada $-\infty < x < \infty$
4. Jika $K(x_1, u) = K(x_2, u)$, berarti $D(x_1, u) = D(x_2, u)$, maka vektor x_1 dan x_2 memiliki tingkat kesamaan dengan u .

Implikasi dari kondisi pertama dan kedua adalah kernel harus dipilih dengan memikirkan jarak. Implikasi dari kondisi keempat adalah vector harus memiliki jarak yang sama dengan pola data u akan memiliki nilai yang sama setelah diaplikasikan oleh kernel.

Unit pola akan melewati seluruh kalkulasi dan memasuki lapisan penjumlahan. Setiap kelas $j \in \{1, 2, \dots, M\}$ memiliki unit tersendiri di lapisan ini. Unit penjumlahan dari kelas j menghitung rata-rata dari nilai yang ada pada unit pola yang diasosiasikan dengan kelas j , seperti pada gambar 1. Fitur lain yang dapat dimasukkan pada lapisan penjumlahan adalah mendeterminasi tingkat kepentingan dari keputusan yang salah dari kelas j . Definisikan l_j sebagai koefisien *loss* yang diasosiasikan dengan kepurusan yang diberikan kepada input x yang

berada pada kelas j padahal kelasnya sangat berbeda. Koefisien ini dapat dimasukkan kedalam fungsi diskriminan sebagai berikut:

$$\hat{d}_j(x) = p_j l_j \hat{f}_j(x) \quad (I.15)$$

Nilai dari l_j tidak bisa diambil dari data dan bergantung pada aplikasi dan kepentingan dari keputusan *false positive* untuk kelas j . Pada pengaplikasian yang tidak terdapat perbedaan antara kelas, l_j dapat dibuat menjadi 1 untuk $j \in \{1, 2, \dots, M\}$. Fungsi lapisan keputusan adalah untuk mengambil nilai $\hat{d}_j(x)$ terbesar dan mendeklarasikan j sebagai kelas dari input vector x [5].

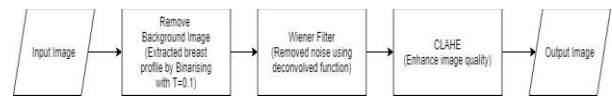
Tujuan

Secara umum, tujuan utama dari pembahasan ini adalah merancang dan menganalisa algoritma untuk mendeteksi kanker payudara dari dataset gambar mamografi yang disimulasikan dengan Matlab software. Pada akhirnya, diharapkan sistem berbasis Computer Aided Detection (CAD) ini mampu mendeteksi kanker payudara secara otomatis sehingga dapat membantu ahli radiologi dan dokter mengambil keputusan dengan cepat dan mengurangi tingkat kesalahan diagnosis.

II. METODE

A. Proses Pembuatan

1. Preprocessing



Gambar 2. Alur Diagram Preprocessing Gambar

1.1 Thresholding

Tahap ini merupakan tahap awal yang bertujuan untuk mengekstrak profil payudara yang relevan serta meminimalisir area citra mamograf yang akan diolah kembali pada tahap segmentasi untuk mendeteksi letak kemungkinan kanker. Prosesnya dimulai dengan mengubah setiap pixel pada gambar menjadi bilangan biner [0, 1] pada batas nilai ambang 0.1 [4], dimana setiap pixel yang bernilai diatas batas tersebut akan diubah menjadi 1 (putih) dan pixel lain akan diubah menjadi 0 (hitam). Saat setiap nilai telah diubah menjadi bilangan biner, latar hitam pada gambar akan terabaikan saat dikali dengan nilai pixel awal sehingga akan terlihat batas profil payudara yang lebih jelas.

1.2 Enhancement

Tahap selanjutnya merupakan rekonstruksi citra mamograf dengan menerapkan metode wiener filter dan CLAHE.

1.2.1 Wiener Filter

Tahap ini bertujuan untuk memperbaiki citra asli dengan mengurangi sensitivitas noise pada filter

inversi menggunakan metode dekonvolusi wiener filter. Dekonvolusi merupakan sebuah operasi inverse (kebalikan) dari proses konvolusi. Dalam hal ini, metode dekonvolusi digunakan untuk mengembalikan sebuah citra yang telah dimanipulasi dengan sebuah filter inversi pada poin dimana citra tersebut berkonvolusi. Fungsi dari metode ini sendiri dinyatakan dengan $\text{deconvwnr}(i)$ dimana parameter yang digunakan antara lain PSF (*point-spread function*) yang dinyatakan dalam fungsi $\text{fspecial}('avarange', [3, 3])$, dimana $[3, 3]$ merupakan parameter letak citra tersebut berkonvolusi, dan NSR (*noise-to-signal ratio*) yang merupakan besaran power ratio dari *additive noise* ($\text{NSR}=0,2$) pada domain transformasi fourier.[9]

1.2.2 Contrast Limited Adaptive Histogram Equalization (CLAHE)

Hasil citra yang telah difilter kemudian akan ditingkatkan kualitasnya menggunakan CLAHE dengan meningkatkan nilai kontras pada setiap bagain kecil pada citra (*tiles*), namun tetap dibatasi untuk menghindari penguatan *noise* yang mungkin terdapat pada gambar. Hal ini dapat membantu proses deteksi pada tahap awal kanker agar lebih akurat. Metode ini dinyatakan dalam fungsi $\text{adapthisteq}(i)$ dengan parameter sebagai berikut: 'NumTiles' (*number of tiles*) yang bernilai $[8, 8]$, dan 'ClipLimit' (*contrast enhancement limit*) yang berada pada level 0.01.[9]

2. Segmentasi

Pada tahap ini, output dari tahap preprocessing akan digunakan sebagai input dalam menentukan ROI (*region of interest*) menggunakan teknik GMM (*Gaussian Mixture Model*) segmentasi berbasis K-means. Untuk menentukannya, bagian tumor akan diidentifikasi berdasarkan letak intensitas tertinggi, intensitas ini dinilai berdasarkan esitmasi distribusi probabilitas setiap pixel pada masing-masing klaster seperti mean dan varian acak. [10] Lalu nilai parameter yang digunakan dalam segmentasi GMM ini antara lain jumlah region ($k=2$), jumlah komponen GMM ($g=2$), dan jumlah iterasi maksimum 10. Kemudian diterapkan pula HMRF-EM (Hidden Markov Random Field Model) dan Algoritma Ekspektasi-Maksimalisasi.

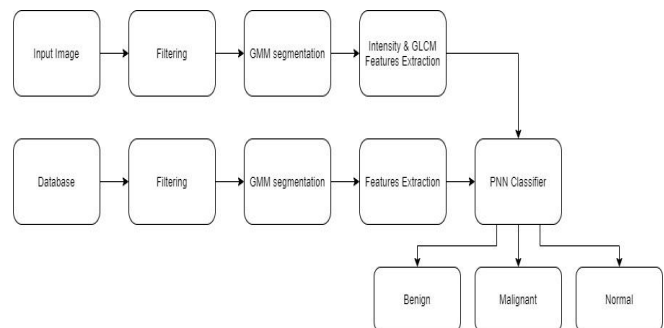
Setelah itu, hasil segmentasi akan dialokasikan untuk ekstraksi fitur. Proses ekstraksi tersebut menerapkan analisis tekstur untuk

mengidentifikasi informasi fitur tekstur tentang distribusi spasial variasi nada dan mendeskripsikan pola variasi pada tingkat *grayscale* di suatu bagian menggunakan metode GLCM.[10]

3. Klasifikasi

Metode klasifikasi yang kita gunakan merupakan probabilistic neural network (PNN). Metode ini banyak digunakan karena kemampuannya untuk mengidentifikasi dan mengklasifikasi pola dengan akurasi yang tinggi. Pada kasus ini, input yang telah kita olah sebelumnya akan kita gunakan sebagai variable bebas untuk memprediksi tipe kanker tersebut berdasarkan database 60 kasus penderita kanker, dimana 20 diantaranya normal.

B. Diagram Alir



Gambar 3. Diagram Alir Klasifikasi Kanker Payudara

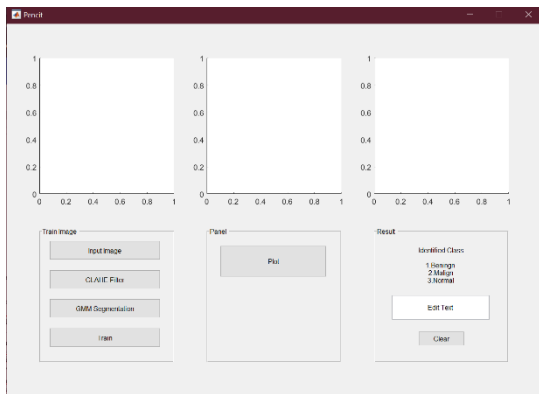
Diagram tersebut memperlihatkan algoritma deteksi jenis kanker payudara secara menyeluruh yang nantinya akan ditampilkan dalam bentuk GUI. Proses ini dimulai dengan memasukkan input berupa citra mammograf. Citra tersebut kemudian difilter untuk ditingkatkan kualitasnya sehingga mempermudah proses segmentasi untuk menentukan ROI. Setelah itu, citra akan diekstraksi kembali melalui analisis tekstur menggunakan metode GLCM. Setelah itu, output citra akan diproses kembali pada sistem klasifikasi PNN melalui serangkaian training menggunakan database yang telah diolah dengan proses yang sama seperti input citra yang akan diidentifikasi sehingga pada hasil akhir nanti akan terlihat apakah jenis tumor tersebut ganas atau jinak atau bahkan tidak terdapat tumor apapun (normal).

III. SIMULASI DAN PEMBAHASAN

A. Hasil Simulasi

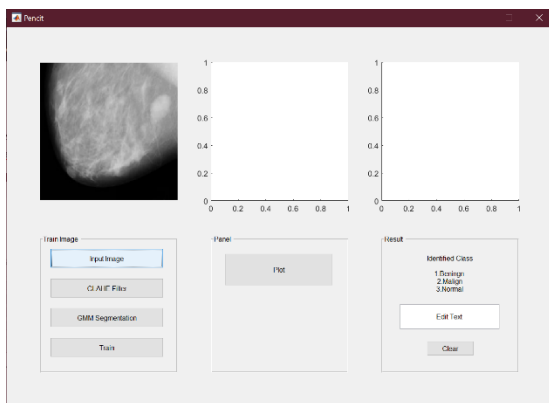
Simulasi dilakukan menggunakan aplikasi MATLAB. Simulasi ini digunakan untuk mengidentifikasi apakah seseorang memiliki sel kanker payudara atau tidak. Berikut langkah-langkah simulasi dengan MATLAB:

1. Menjalankan program



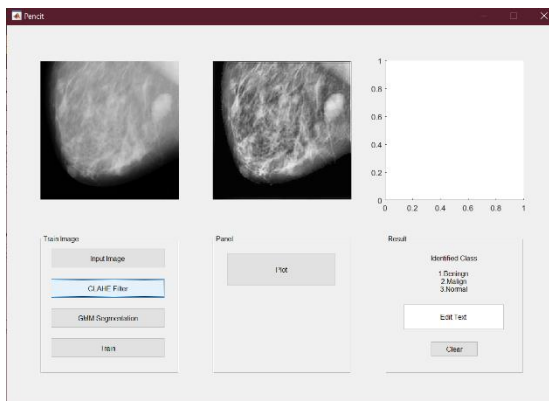
Gambar 4. Window yang Muncul Ketika Pertama di-run

- Memasukkan hasil scan payudara yang akan diidentifikasi



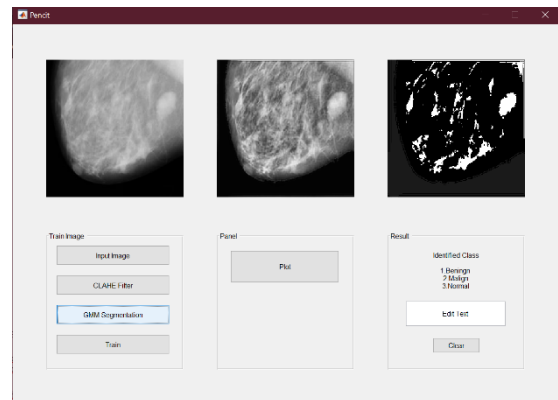
Gambar 5. Hasil Scan di-Input

- Memperjelas gambar dengan CLAHE filter



Gambar 6. Kontras Gambar diubah dengan CLAHE Filter

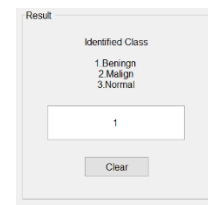
- Menentukan Region of Interest



Gambar 7. Teknik GMM Segmentation digunakan untuk Menentukan ROI

- Mengidentifikasi hasil scan

Gambar scan payudara tersebut menunjukkan hasil 1 yang merupakan *benign*, yaitu terdapat tumor jinak yang tidak berpotensi kanker.



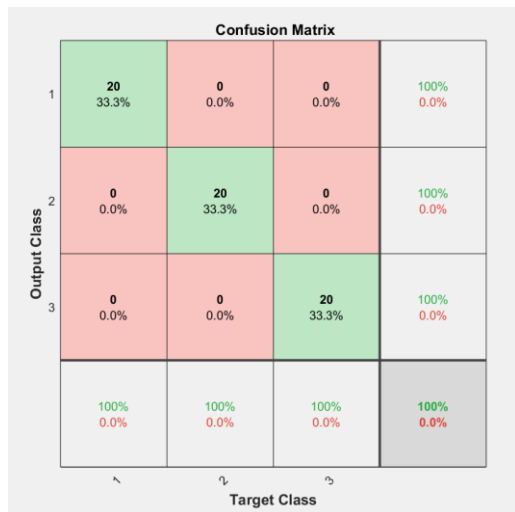
Gambar 8. Identifikasi Hasil Scan

B. Analisis dan Pembahasan

Klasifikasi yang telah dibuat perlu membutuhkan validasi sehingga tidak menciptakan kesalahan dalam proses diagnosa ketika diaplikasikan. Untuk mengukur akurasi dari klasifikasi kanker payudara yang sudah kami buat, Kami menggunakan metode confusion matrix serta Region of Curve (ROC) untuk melakukan validasi klasifikasi tersebut.

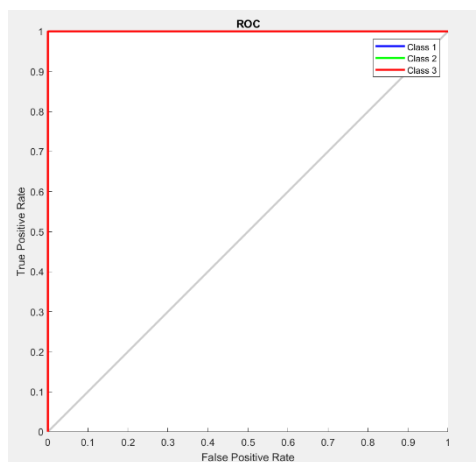
Confusion matrix adalah pengukuran performa untuk masalah klasifikasi machine learning dimana keluaran dapat berupa dua kelas atau lebih [binus]. Confusion Matrix adalah tabel dengan 4 kombinasi berbeda dari nilai prediksi dan nilai aktual. Ada empat istilah yang merupakan representasi hasil proses klasifikasi pada confusion matrix yaitu True Positif, True Negatif, False Positif, dan False Negatif.

Sedangkan ROC adalah metode representasi grafis dari hubungan antara sensitivitas dan 1-spesifisitas [ugm]. Dalam penelitian medis kurva ROC banyak digunakan untuk menggambarkan keakuratan diagnostik dan menentukan nilai cut-off yang optimal. Keakuratan diagnosis berasal dari area di bawah kurva ROC dan optimal cut-off digunakan untuk mengidentifikasi kondisi positif dan negatif dalam diagnosis. Terlihat pada gambar dibawah merupakan hasil confusion matrix dan ROC dari klasifikasi yang sudah dibuat.



Gambar 9. Confusion Matrix Model

Hasil dari evaluasi confusion matrix terlihat bahwa model yang telah dilatih dapat melakukan diferensiasi terhadap klasifikasi yang diinginkan. Pada matrix tersebut terlihat bahwa setiap kategori tepat diprediksikan dengan kategori 1, 2, dan 3 bernilai 100% akurasi. Hal ini membuktikan bahwa Masukan gambar yang diberikan oleh model dapat diprediksi secara tepat oleh model sehingga tidak terdapatnya false negative maupun false positive.



Gambar 10. ROC Model

Hasil dari evaluasi ROC dapat terlihat bahwa model memiliki batasan cut-off optimal di titik (0,1). Hal ini menyatakan bahwa model yang telah dibuat dapat mengatur Batasan pada titik yang optimal tersebut sehingga dapat secara tepat mendiagnosa gambar mammography yang diberikan. ROC pada titik optimal tersebut terlihat bahwa True Positive Rate (TPR) yang ada berada pada nilai 1 sedangkan False Positive Rate (FPR) berada di 0 membuktikan bahwa model mampu mengklasifikasikan masukan gambar yang diberikan tanpa kesalahan.

Berdasarkan 2 evaluasi tersebut, terlihat bahwa model yang telah dibuat dapat mengklasifikasikan masukan gambar yang diberikan. Akan tetapi, terlihat bahwa hasil evaluasi tersebut sangat ideal dengan mencapai akurasi 100% pada confusion matrix dan nilai titik optimal (0,1) pada ROC. Hal tersebut dapat saja diakibatkan oleh masukan gambar yang kurang ketika proses testing sehingga diperlukannya proses testing dengan masukan gambar yang lebih banyak untuk

melakukan evaluasi lebih lanjut. Namun dengan berdasarkan percobaan yang telah dilakukan, kita dapat mengetahui bahwa klasifikasi kanker payudara yang telah dibuat mampu melakukan diferensiasi gambar yang telah diberikan sehingga dapat mampu mendiagnosa gambar mammography berdasarkan klasifikasi-klasifikasi yang diberikan.

IV. KESIMPULAN

Kanker terjadi ketika terdapat sel abnormal yang tumbuh dengan tidak terkendali, biasanya menyerang bagian tubuh yang berdekatan, dan dapat menyebar ke organ lain. salah satu kanker yang paling berbahaya adalah kanker payudara. kanker payudara menempati urutan kedua sebagai penyebab utama kematian wanita diseluruh dunia, sekitar 1 dari 8 wanita menderita kanker payudara dalam hidupnya dan hanya 5% sampai 10% kanker payudara terjadi pada wanita dengan hubungan genetik yang jelas. Mamografi adalah tes pemindaian paling dasar untuk kanker payudara. Mamografi merupakan pencitraan medis khusus yang menggunakan sistem sinar X dosis rendah untuk melihat bagian dalam payudara[4].

Pada percobaan ini dibuat simulasi untuk mendeteksi kanker payudara menggunakan *software* MATLAB dengan menggunakan metode *probabilistic neural network* (PNN) untuk mengklasifikasikan kanker payudara. Metode PNN dicapai dengan melewati input gambar, threshold dengan binerisasi, merekonstruksi sekaligus meningkatkan kualitas gambar dengan wiener filter, meningkatkan kontras dengan CLAHE, dan menentukan *region of interest* dengan GMM. Hasil dari simulasi pengklasifikasian kanker payudara adalah kita dapat mengetahui apakah input merupakan tumor ganas, jinak, atau normal. Pada *confusion matrix* yang dihasilkan dapat diketahui bahwa hasil simulasi ini memiliki keakuratan 100%, sistem ini perlu diperbanyak datasetnya untuk memvalidasi lebih lanjut tentang keakuratan sistem ini.

REFERENSI

- [1] WHO, "Cancer," 2021. [Online]. Available: https://www.who.int/health-topics/cancer#tab=tab_2. [Accessed 12 January 2021].
- [2] Kementerian Kesehatan Republik Indonesia, "Panduan Penataklasaan Kanker payudara," 2010. [Online]. Available: <http://kanker.kemkes.go.id/guidelines/PPKPayudara.pdf>. [Accessed 12 January 2021].
- [3] The Global Cancer Observatory, "Breast," December 2020. [Online]. Available: <https://gco.iarc.fr/today/data/factsheets/cancers/20-Breast-fact-sheet.pdf>. [Accessed 12 January 2021].
- [4] A. Makandar and B. Halalli, "Pre-processing of Mammography Image for Early," *International Journal of Computer Applications*, vol. 144, no. 3, p. 0975 – 8887, 2016.
- [5] B. Mohebbi, A. Tahmassebi, A. Meyer-Baese and A. Gandomi, "Probabilistic neural networks: a brief overview of theory, implementation, and application", *Handbook of Probabilistic Models*, pp. 347-359, 2020.
- [6] F. Ancona, A. Colla, S. Rovetta and R. Zunino, "Implementing probabilistic neural networks", *Neural Computing and Applications*, pp. 152-159, 1998 [Online]. Available:

https://www.researchgate.net/publication/220372521_Implementing_Probabilistic_Neural_Networks
[Accessed: 12- Jan- 2021]

- [7] S. Satapathy, S. Dehuri, A. Jagadev and S. Mishra, EEG Brain Signal Classification for Epileptic Seizure Disorder Detection. Elsevier Inc., 2019.
- [8] N. Aibe, M. Yasunaga, I. Yoshihara and J. H. Kim, "A probabilistic neural network hardware system using a learning-parameter parallel architecture," Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No.02CH37290), Honolulu, HI, USA, 2002, pp. 2270-2275 vol.3, doi: 10.1109/IJCNN.2002.1007495.
- [9] T. Abbas and M. Nafea Jafaar, "Mammogram image with Wiener filter and CLAHE", Journal of AL-Qadisiyah for computer science and mathematics, vol.9, no.2, p. 97 – 102, 2017.
- [10] S. Shamy and J. Dheeba, "A Research on Detection and Classification of Breast Cancer using k- means GMM & CNN Algorithms", International Journal of Engineering and Advanced Technology (IJEAT), vol. 8, August 2019.