

# **Klasifikasi Tumor Payudara Berbasis Ciri Tekstur pada Citra Mammografi Menggunakan Metode Naive Bayes**

**BELLA JULIA<sup>1</sup>, HENI SUMARTI<sup>2\*</sup>, HAMDAN HADI KUSUMA<sup>3</sup>**

<sup>1,2,3</sup>Jurusan Fisika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Walisongo Semarang, Indonesia

Email: [heni\\_sumarti@walisongo.ac.id](mailto:heni_sumarti@walisongo.ac.id)

*Received* 16 Agustus 2022 | *Revised* 26 September 2022 | *Accepted* 21 Desember 2022

## **ABSTRAK**

*Kanker payudara adalah jenis kanker yang terjadi pada sebagian besar wanita. Kanker payudara terjadi akibat pertumbuhan berlebih atau perkembangan sel jaringan payudara yang tidak terkendali. Mammografi merupakan metode terbaik untuk deteksi dini kanker payudara karena dapat menunjukkan lesi secara dini. Namun, analisis terhadap mammogram ini masih dilakukan secara manual oleh ahli medis, sehingga perlu perangkat tambahan. Telah banyak penelitian tentang olah citra untuk deteksi kanker secara otomatis. Pada penelitian ini digunakan metode Naive Bayes untuk klasifikasi citra tumor jinak dan tumor ganas. Tujuan dari penelitian ini untuk mengklasifikasi citra mammografi berdasarkan dua kelas yaitu tumor jinak dan tumor ganas dengan berbasis ciri tekstur menggunakan histogram dan Gray Level Co-occurrence Matrix (GLCM). Penelitian ini menunjukkan hasil akurasi sebesar 80%, sensitivitas sebesar 90%, dan spesifitas sebesar 70%. Oleh karena itu, penelitian ini bisa dijadikan perangkat tambahan untuk klasifikasi tumor payudara ganas dan jinak.*

**Kata kunci:** *Tumor Payudara, Mammografi, Klasifikasi Naive Bayes.*

## **ABSTRACT**

*Breast cancer is a type of cancer that occurs in most women. Breast cancer occurs due to overgrowth or uncontrolled development of breast tissue cells. Mammography is the best method for early detection of breast cancer because it can show lesions early. However, the analysis of this mammogram is still done manually by medical experts, so additional devices are needed. There have been many studies on image processing for automatic detection. In this study, the Naive Bayes method was used to classify images of benign tumors and malignant tumors. The purpose of this study is to classify mammographic images based on two classes, namely benign and malignant tumors based on histogram textures and Gray Level Co-occurrence Matrix (GLCM). This study showed an accuracy of 80%, sensitivity of 90%, and specificity of 70%. Therefore, this study can be used as an additional tool to classify malignant and benign tumors.*

**Keywords:** *Breast Cancer, Mammografi, Naive Bayes Classification*

## 1. PENDAHULUAN

Menurut WHO, ada sekitar 1,38 juta masalah kontemporer pada tahun 2008, dengan 458.000 wanita meninggal akibat kanker payudara setiap tahun, lebih dari setengahnya, sekitar 269.000, masih hidup di negara berkembang dengan pendapatan per kapita rendah (**WHO, 2018**). Sementara itu, di Indonesia, berdasarkan profil kesehatan tahun 2012, Kementerian Kesehatan Republik Indonesia menyebutkan bahwa kanker payudara merupakan kanker yang paling banyak terjadi pada wanita Indonesia, dengan insidensi sebesar 2,2% per 1.000 wanita. Jika hal ini tidak dikendalikan, diperkirakan 26 juta orang akan terkena kanker payudara dan 17 juta orang akan meninggal pada tahun 2030 (**Kemenkes, 2019**).

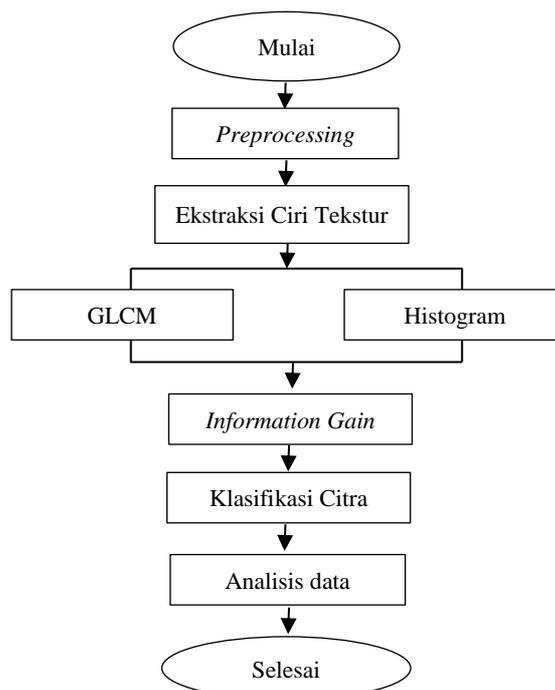
Salah satu jenis kanker yang ditemukan pada kebanyakan perempuan adalah kanker payudara. Pertumbuhan yang berlebih atau perkembangan yang tidak terkendali dalam sel jaringan payudara merupakan penyebab terjadinya kanker payudara (**Fitriani, 2014**). Munculnya tumor atau nodul pada payudara merupakan salah satu gejala yang dapat dideteksi. Cara paling efektif untuk mencegah meningkatnya angka kematian dan menekan jumlah kasus tumor payudara adalah dengan deteksi dini. Salah satu metode skrining terbaik adalah menggunakan mammografi untuk kanker payudara. Hal ini disebabkan oleh kemampuan mammografi dalam menunjukkan keberadaan lesi pada tahap awal. Ahli radiologi bertugas untuk menentukan diagnosis lesi tersebut. Hasil citra mammografi ini telah dapat dikonversi dalam bentuk digital dan meningkatkan minat pengolahan citra menggunakan sistem komputer sebagai alat bantu diagnosis (**Widodo et al., 2012**).

Metode untuk menemukan, mendiagnosis atau menyelidiki suatu penyakit dapat dilakukan dengan menggunakan pencitraan medis. Ini berhubungan dengan menciptakan gambaran tubuh manusia yang berhubungan dengan teknik dan proses diagnosis medis. Di rumah sakit, bidang ini biasa disebut dengan radiologi (**Putra, 2011**). Salah satu metode yang efektif untuk mengidentifikasi jenis tumor payudara adalah dengan pemeriksaan menggunakan mammografi. Hasil pencitraan pada mammografi disebut dengan mammogram. Jika pasien tidak merasakan adanya perubahan yang terjadi pada payudaranya, mammogram dapat digunakan sebagai skrining untuk menemukan perubahan. Selain itu, jika dokter meragukan adanya suatu perubahan dalam payudara, mammografi dapat digunakan sebagai alat diagnosis. Namun, teknik analisis terhadap mammogram ini masih dilakukan secara manual oleh dokter atau ahli radiologi (**Listia & Harjoko, 2014**). Oleh karena itu, diperlukan alat tambahan untuk diagnosa kanker payudara agar mendapatkan hasil yang lebih akurat. Telah banyak dikembangkan olah citra medis untuk klasifikasi citra kanker payudara.

Penelitian mengenai kanker payudara telah banyak dilakukan, penelitian yang dilakukan sebelumnya oleh Amrane dkk. (2018) menggunakan dua *classifier* yang berbeda yaitu Naive Bayes (NB) dan *K-Nearest Neighbors* (KNN) untuk klasifikasi kanker payudara. Evaluasi akurasi masing-masing algoritma menggunakan *cross-validation*. Hasil menunjukkan bahwa KNN menghasilkan akurasi hingga 97.51%, sedangkan Naive Bayes menghasilkan akurasi hingga 96.19%. Belum ada penelitian yang mengklasifikasi kasus kanker payudara berdasarkan ciri tekstur dengan metode Naive Bayes. Pada penelitian ini, dilakukan klasifikasi tumor payudara untuk membedakan kasus ganas dan jinak menggunakan metode Naive Bayes. Penelitian ini bertujuan untuk meningkatkan akurasi dan menunjang tenaga kesehatan dalam mediagnosis citra mammografi, khususnya pada kasus ganas dan jinak menggunakan ciri tekstur yang dapat digunakan untuk prediksi di bidang medis.

## 2. METODE PENELITIAN

Penelitian ini menggunakan citra mammogram dari database MIAS (*Mammographic Image Analysis Society*), terdiri dari 10 citra ganas dan 10 citra jinak. Citra awal hasil pemeriksaan Mammografi pada payudara yang tergolong citra RGB yang akan diubah ke *grayscale*. Kemudian diekstraksi menggunakan MATLAB, dan diolah menggunakan WEKA. Tahapan penelitian ini ditunjukkan pada Gambar 1. Terdapat 4 tahapan utama yaitu : *preprocessing*, ekstraksi ciri tekstur, menentukan nilai *gain*, dan klasifikasi menggunakan metode Naive Bayes.



Gambar 1. Flowchart Penelitian

### 2.1. Preprocessing

*Preprocessing* terdapat tiga proses yaitu proses mengambil citra dari MIAS (*Mammographic Image Analysis Society*), proses *image resize* (melalui perbaikan ukuran citra), mengubah citra input RGB menjadi *grayscale*.

### 2.2. Ekstraksi Ciri Tekstur

Proses untuk memperoleh suatu ciri yang mewakili citra itu sendiri disebut dengan ekstraksi ciri. Suatu objek dapat dikenali sesuai dengan kategorinya yang merupakan suatu ciri (**Fajrin et al., 2015**). Perulangan elemen tekstur dasar yang diklaim "texel" (*texture element – texel*) dikenal sebagai ciri tekstur. Sejumlah piksel menggunakan probabilitas penempatan secara periodik, semi periodik dan juga acak merupakan bagian dari sebuah *texel*. Biasanya tekstur buatan mempunyai pola eksklusif atau berulang secara periodik, sedangkan tekstur yang terbentuk secara alami bersifat acak (**Purwaningsih et al., 2015**). Metode ekstraksi ciri yang digunakan yaitu Histogram dan GLCM.

#### 2.2.1. Histogram

Histogram merupakan grafik yang dinyatakan dengan frekuensi kemunculan intensitas piksel atau nilai-nilai warna suatu citra (**Marlina et al., 2020**). Hasil dari histogram menyatakan skala abu-abu (*grayscale*). Adapun beberapa metode statistik histogram yang digunakan untuk ekstraksi ciri diantaranya yaitu (**Harismah et al., 2016**):

- 1) *Mean* adalah ukuran rata-rata intensitas citra.

$$\mu = \sum_{n=0}^N f_n p(f_n) \quad (1)$$

dengan  $\mu$  adalah *mean*,  $f_n$  adalah tingkat keabuan citra,  $p(f_n)$  adalah nilai histogram dari  $n$ , dan  $N$  adalah tingkat keabuan tertinggi.

- 2) Standar deviasi merupakan ukuran rata-rata kontras.

$$\sigma = \sqrt{\mu_2(z)} = \sqrt{\sigma^2} \quad (2)$$

dengan  $\sigma$  adalah standar deviasi, dan  $z$  adalah probabilitas kemunculan nilai keabuan.

- 3) Varian merupakan variasi elemen suatu citra.

$$\alpha_2 = \sum_{n=0}^N (f_n - \mu)^2 p(f_n) \quad (3)$$

dengan  $\alpha_2$  adalah skewness, dan  $n$  adalah nomor piksel dari 0 sampai  $N$ .

- 4) *Skewness* merupakan kemiringan relatif dari kurva histogram citra.

$$\alpha_3 = \frac{1}{a^3} \sum_{n=0}^N (f_n - \mu)^3 p(f_n) \quad (4)$$

dengan  $\alpha_3$  adalah *skewness*.

- 5) Kurtosis merupakan keruncingan relatif dari kurva histogram citra.

$$\alpha_4 = \frac{1}{\alpha^4} \sum_{n=0}^N (f_n - \mu)^4 p(f_n) - 3 \quad (5)$$

dengan  $\alpha_4$  adalah kurtosis.

- 6) Entropi merupakan bentuk ketidakteraturan (*randomness*) dari suatu citra.

$$H = - \sum_{n=0}^N p(f_n)^2 \log p(f_n) \quad (6)$$

dengan  $f_n$  adalah nilai intensitas keabuan,  $p(f_n)$  adalah nilai histogram, dan  $\mu$  adalah *mean*.

### 2.2.2. GLCM (*Gray Level Co-Occurrence Matrix*)

Metode untuk menghitung pola tekstur pada suatu citra menggunakan memperhitungkan interaksi ketetanggaan antar piksel disebut dengan GLCM (*Gray Level Co-Occurrence Matrix*). Metode ini menggunakan perhitungan sudut yang dibuat dengan dua buah piksel sehingga disebut dengan matriks kookurensi. Matriks ini berisi nilai intensitas kedua piksel dengan jarak tertentu dan menciptakan suatu sudut. Beberapa *descriptor* yang mampu diturunkan menurut glcm yaitu kontras, energi, entropi, varian, *cluster performance*, homogenitas, korelasi (Auliasari & Kertaningtyas, 2018). Beberapa fitur yang dipakai menurut GLCM yaitu (Olaniyi et al., 2017) (Singh & Kaur, 2012):

- 1) Energi mengembalikan jumlah elemen kuadrat dalam GLCM, dengan 1 untuk gambar konstan. Energi juga dikenal sebagai keseragaman, sudut Kedua.

$$\sum_{i,j=0}^{N-1} P(i, j)^2 \quad (7)$$

dengan N-1 adalah jumlah skala grayscale yaitu 256,  $P(i, j)$  adalah matriks GLCM yang telah ternormalisasi.

- 2) *Correlation* menentukan bagaimana dependensi linier dari *gray* tingkat yang sedang diukur. Pada setiap kesempatan, ada korelasi nilai, ini menunjukkan urutan lingkungan positif dari abu-abu tingkat. Korelasi mengembalikan ukuran seberapa berkorelasi piksel ke tetangga di seluruh gambar.

$$\sum_{i=0}^{g-1} \sum_{j=0}^{g-1} \frac{\{i \times j\} \times P\{i, j\} - \{\mu_x \times \mu_y\}}{\sigma_x \times \sigma_y} \quad (8)$$

Dengan  $\mu_x$ ,  $\mu_y$ ,  $\sigma_x$ , dan  $\sigma_y$  adalah nilai rata-rata dan standar deviasi dari matriks  $P_x$  dan  $P_y$ , berturut-turut.

- 3) Homogenitas mengukur suatu wilayah tertentu didasarkan secara seragam menurut perbedaan tingkat abu-abunya.

$$\sum_{i=0}^{g-1} \sum_{j=0}^{g-1} \left\{ \frac{P(i, j)}{1 + |i - j|} \right\} \quad (9)$$

dengan  $P(i, j)$  adalah matriks GLCM yang telah ternormalisasi.

- 4) *Contrast* adalah ukuran kontras intensitas antara a piksel dan piksel yang lainnya di seluruh gambar. Ini adalah nol untuk gambar konstan.

$$\sum_{i,j=0}^{N-1} |i - j|^2 P(i, j) \quad (10)$$

dengan  $P(i, j)$  adalah matriks GLCM yang telah ternormalisasi.

### 2.2.3. Information Gain

Perolehan informasi, merupakan suatu istilah mengukur jumlah bit informasi yang diperoleh untuk prediksi kategori ada atau tidak adanya istilah dalam dokumen. Information Gain mengukur penurunan entropi ketika fitur yang diberikan tidak ada. Ini adalah penerapan teknik yang lebih umum, pengukuran informasional entropi untuk masalah yang memutuskan seberapa penting fitur yang diberikan (**Azhagusundari & Thanamani, 2013**) (**Ladha & Deepa, 2011**):

$$gain(s_j) = E(P_j) - E(s_j) \quad (11)$$

$$E(P) = \sum_{i=1}^n P_i \log_2 P_i$$

dengan :

dengan  $P_i$  adalah rasio dari kondisi atribut  $P$  dalam dataset,  $S_j$  adalah nilai atribut dan kondisi atribut partisi  $P_i$  pada set  $P$  menggunakan atribut  $S_j$ .

#### 2.2.4. Klasifikasi

Klasifikasi adalah salah satu tugas utama dalam data *mining*. Proses klasifikasi adalah proses komputasi data yang ada dengan data baru (Saiful rahman et al., 2019). Klasifikasi menggunakan sifat numerik dari fitur gambar untuk mengatur data menjadi kategori yang berbeda. Klasifikasi menggunakan dua fase pemrosesan, yakni fase pelatihan dan fase pengujian. Dalam fase pelatihan, sifat dari ciri fitur gambar diekstraksi dan setiap kategori diklasifikasi. Pada fase pengujian, fitur ini memiliki partisi ruang yang digunakan untuk mengklasifikasikan fitur gambar (Singh & Kaur, 2012).

Klasifikasi dalam penelitian ini menggunakan metode Naive Bayes. Algoritma Naive Bayes adalah pengklasifikasi probabilistik sederhana yang menghitung kombinasi nilai dalam kumpulan data tertentu dan sekumpulan probabilitas dengan menghitung frekuensi. Algoritma yang menggunakan teorema Bayes mengasumsikan bahwa semua variabel independen ketika mempertimbangkan nilai variabel kelas. Teorema Bayes adalah rumus matematika untuk menentukan probabilitas bersyarat, yang dinamai dari ahli matematika Inggris abad ke-18 Thomas Bayes (Zamrodah, 2016).

#### 2.2.5. Analisis Data

Tahap analisis merupakan tahap akhir dari penelitian holistik, dimana dilakukan analisis terhadap hasil pengujian pada tahap klasifikasi yang telah dilakukan sebelumnya. Analisis dilakukan untuk menguji seberapa baik hasil klasifikasi diperoleh dari skenario pengujian yang dirancang. Beberapa nilai yang dibandingkan adalah dataset yang diklasifikasikan dengan benar, dataset yang salah diklasifikasikan, nilai akurasi, dan keuntungan (Oktavianto & Handri, 2020).

Keberhasilan dalam proses klasifikasi dapat ditunjukkan menggunakan nilai indeks pengukuran yang diperoleh menurut hasil klasifikasi data menggunakan machine learning WEKA pada bentuk *confusion matrix*. Hasil dari *confusion matrix* yaitu *True Positive (TP)*, *True Negatif (TN)*, *False Negative (FN)* dan *False Positive (FP)*. TP adalah data positif yang terklasifikasi positif pada sistem. TN adalah data negatif yang terklasifikasi negatif pada sistem. FN adalah data negatif yang terklasifikasi positif pada sistem dan FP merupakan data positif yang terklasifikasi negatif (Frank et al., 2017).

Olaniyi dkk. (2017) mengemukakan ide mengenai indeks pengukuran performa pada suatu metode klasifikasi yaitu dengan menggunakan perhitungan *confusion matrix*. *Confusion matrix* yang digunakan yaitu dengan nilai akurasi (*accuracy*), sensitivitas (*sensitivity*), spesifisitas dan (*specificity*) yang dirumuskan sebagai berikut :

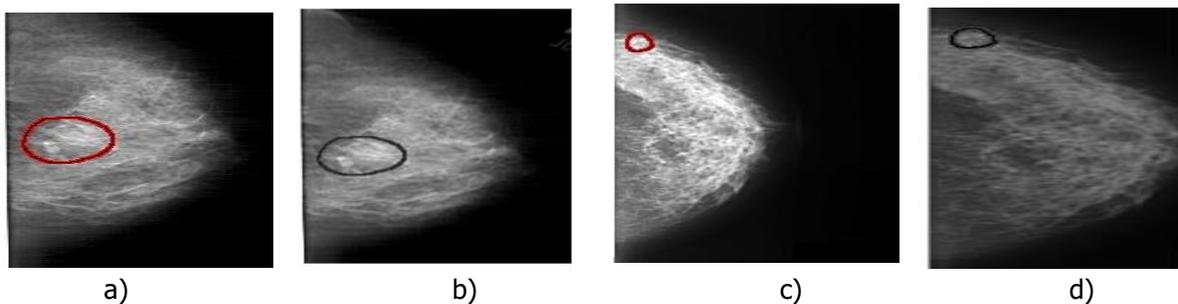
$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (12)$$

$$Sensitivitas = \frac{TP}{TP + FN} \times 100\% \quad (13)$$

$$Spesifitas = \frac{TN}{TP + FN} \times 100\% \quad (14)$$

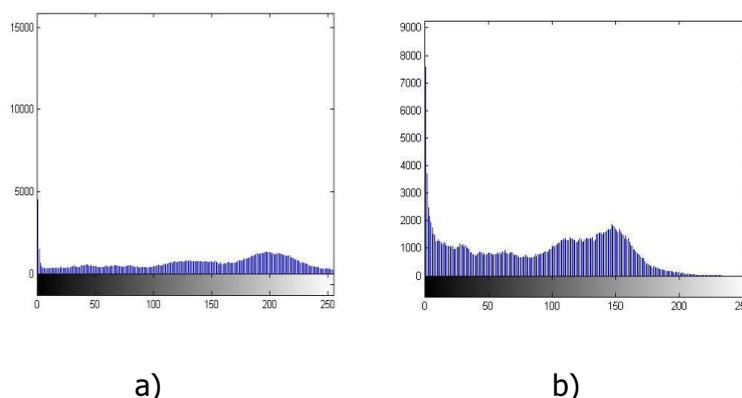
### 3. HASIL DAN PEMBAHASAN

Data penelitian yang digunakan merupakan citra mammogram dari database MIAS (*Mammographic Image Analysis Society*). Gambar 2 menunjukkan hasil citra mammografi payudara pada tumor ganas dan tumor jinak sebelum *preprocessing* dan setelah *preprocessing*. Hasil *preprocessing* yang telah dilakukan memiliki perbedaan pada keseragaman pada ukuran dan warna citra menjadi *grayscale* (abu-abu) untuk proses selanjutnya.



**Gambar 2. Citra Mammografi Kanker Payudara a) Tumor Ganas Sebelum *Preprocessing*, b) Tumor Ganas Setelah *Preprocessing*, c) Tumor Jinak Sebelum *Preprocessing*, D) Tumor Jinak Setelah *Preprocessing***

Gambar 3 menunjukkan hasil nilai histogram citra mammografi kanker payudara pada tumor ganas dan tumor jinak. Perbandingan dari kedua hasil histogram memiliki pola bentuk grafik yang berbeda, pada citra ganas grafik cenderung lebih rendah sehingga menunjukkan citra gelap dan pada citra jinak grafik cenderung tinggi sehingga menunjukkan citra terang. Sehingga hasil dari nilai histogram ini bisa dijadikan alat untuk klasifikasi.



**Gambar 3. Histogram dari Citra Kanker Payudara a) Tumor Ganas, b) Tumor Jinak**

Hasil ekstraksi ciri tekstur dengan 10 atribut, dimana 6 berasal dari histogram dan 4 dari GLCM. Hasil dari nilai histogram yang terdiri dari mean, standar deviasi, varian, *skewness*, kurtosis, dan entropi. Sementara itu, nilai dari nilai GLCM, yang terdiri dari energi, *contrast*, *correlation*, dan homogenitas. Hasil nilai rata-rata dari ciri tekstur ditunjukkan pada tabel 1. Atribut yang memiliki selisih nilai rata-rata yang besar yaitu standar deviasi, varian, entropi, dan kurtosis. Sementara itu, atribut yang memiliki selisih nilai rata-rata kecil yaitu *mean*, *skewness*, energi, *contrast*, *correlation*, dan homogenitas. Selisih nilai rerata pada citra mammografi tumor ganas dan tumor jinak menunjukkan variasi terhadap ciri tekstur. Nilai rata-rata ciri tekstur ditambahkan dengan nilai standar deviasinya untuk melihat nilai sebaran datanya. Standar deviasi merupakan atribut yang berguna untuk mengetahui tingkat kekontrasan suatu citra. Semakin rendah kontras citra, maka semakin rendah pula standar deviasinya, begitu pula

sebaliknya (**Putra, 2011**). Untuk masing-masing nilai atribut yang didapatkan, atribut yang mempunyai sebaran data seragam yaitu *skewness*, *contrast*, *correlation*, energi, dan homogenitas. Sementara itu, atribut yang mempunyai sebaran data acak yaitu *mean*, varian, entropi, standar deviasi, dan kurtosis.

**Tabel 1. Nilai Rata-rata Ciri Tekstur Histogram dan GLCM**

No	Atribut	Rata-rata Ganas	Rata-rata Jinak
1	Mean	105.236 ± 55.383	114.210 ± 34.263
2	Standar deviasi	692.889 ± 118.298	642.190 ± 204.200
3	Varian	491.115 ± 168.104	529.466 ± 154.056
4	Entropi	654.397 ± 224.436	634.074 ± 372.524
5	<i>Skewness</i>	-170.67 ± 512.03	153.61 ± 460.45
6	Kurtosis	243.558 ± 110.371	191.367 ± 824.067
7	Energi	0.3749 ± 0.0514	0.5502 ± 0.0932
8	<i>Contrast</i>	0.0993 ± 0.0204	0.0479 ± 0.019
9	<i>Correlation</i>	0.9911 ± 0.0018	0.9927 ± 0.0026
10	Homogenitas	0.9732 ± 0.061	0.9853 ± 0.0054

Nilai hasil rata-rata gain ratio ditunjukkan seperti pada tabel 2. Hasil yang didapatkan dari nilai *gain ratio* yang lebih berpengaruh yaitu *skewness*, *contrast*, energi, homogenitas, *mean*, dan entropi karena nilai yang didapatkan lebih tinggi. Sementara yang tidak berpengaruh yaitu varian, standar deviasi, *correlation* dan kurtosis karena hasil yang didapatkan nol sehingga tidak berpengaruh terhadap proses klasifikasi.

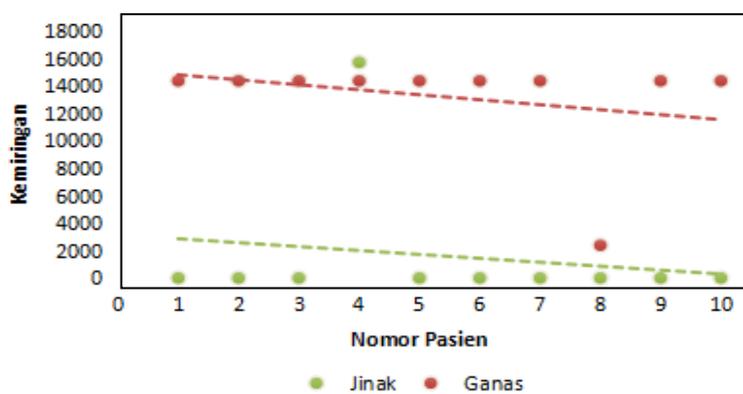
**Tabel 2. Hasil Rata-rata Gain Ratio**

No	Atribut	Gain
1	<i>Skewness</i>	0,628
2	<i>Contrast</i>	0,628
3	Energi	0,628
4	Homogenitas	0,628
5	Mean	0,531
6	Entropi	0,528
7	Varian	0
8	Standar deviasi	0
9	<i>Correlation</i>	0
10	Kurtosis	0

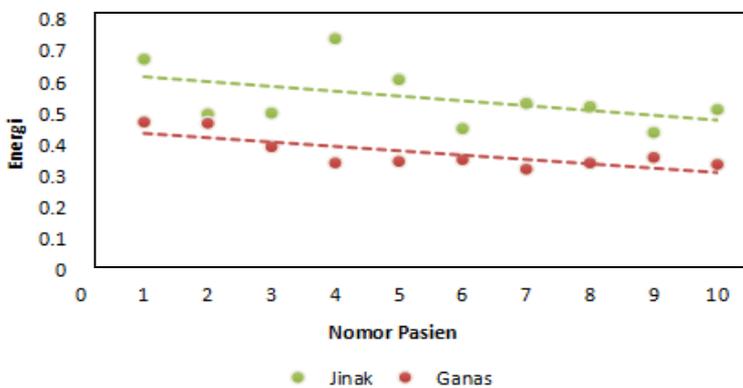
Gambar 4 menunjukkan bahwa atribut yang berpengaruh saat pengklasifikasian dengan menggunakan metode Naive Bayes diambil peringkat tiga besar. Atribut pertama yaitu Kemiringan menunjukkan bahwa kemiringan relatif dari kurva histogram pada data tumor ganas dan tumor jinak yang dihasilkan sama. Atribut kedua yaitu entropi menunjukkan nilai entropi dimana pada grafik ketidakteraturan bentuk pada citra, untuk citra ganas lebih tidak teratur dibandingkan dengan citra jinak. Atribut ketiga yaitu energi menunjukkan nilai energi bahwa keseragaman yang dihasilkan berbeda antara citra tumor ganas dan citra tumor jinak.

Hasil dari *confusion matrix* dari metode Naive Bayes pada tabel 3 yaitu nilai TP (*True Positive*) sebesar 9, TN (*True Negative*) sebesar 7, FP (*False Positive*) sebesar 3, dan FN (*False Negative*) sebesar 1. Nilai TP merupakan tumor jinak yang terdeteksi tumor jinak, TN merupakan tumor jinak terdeteksi tumor ganas, FP merupakan tumor ganas yang terdeteksi tumor ganas, FN merupakan tumor ganas terdeteksi tumor jinak. Hasil penelitian ini menunjukkan hasil yang cukup akurat untuk proses klasifikasi pada 20 data.

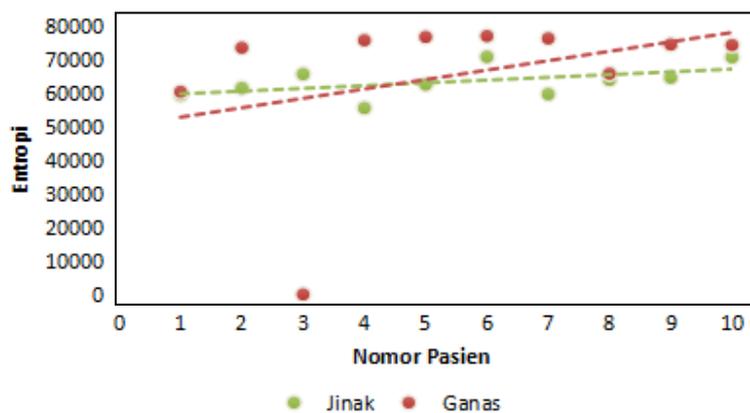
# Klasifikasi Tumor Payudara Berbasis Ciri Tekstur Pada Citra Mammografi Menggunakan Metode Naive Bayes



(a)



(b)



(c)

**Gambar 4. Nilai Atribut yang Berpengaruh pada Proses Klasifikasi a) Kemiringan, b) Entropi, c) Energi**

**Tabel 3. Hasil *Confusion Matrix***

TN	TP	FN	FP	Akurasi	Sensitivitas	Spesifitas
9	7	1	3	80%	90%	70%

Hasil klasifikasi menggunakan metode Naive Bayes menghasilkan nilai akurasi sebesar 80%, nilai sensitivitas 90%, dan nilai spesifitas 70%. Hasil ini menunjukkan bahwa akurasi pada penelitian ini lebih rendah jika dibandingkan dengan penelitian sebelumnya yang dilakukan oleh Amrane dkk. (2018) yang membandingkan metode KNN dan Naive Bayes dengan akurasi tertinggi pada Naive Bayes adalah 96.19%. Hal ini disebabkan oleh ciri yang digunakan pada penelitian ini bukan hanya dari gambar mammogram, melainkan dari biomarker berupa Glucose, Resistin, Umur, BMI, HOMA, Leptin, Insulin, Adiponectin, dan MCP-1. Berbeda dengan penelitian ini yang hanya menggunakan ciri tekstur dari citra mammogram.

#### 4. KESIMPULAN

Telah dilakukan klasifikasi kanker payudara berbasis ciri tekstur dengan data citra mammografi yang berjumlah 10 citra tumor ganas dan 10 citra tumor jinak. Penelitian ini menggunakan metode Naive Bayes dengan hasil akurasi 80%, sensitivitas 90%, dan spesifitas 70%. Hasil yang telah didapatkan pada penelitian ini mampu mengklasifikasikan pengolahan citra tumor ganas dan tumor jinak, sehingga penelitian ini bisa membantu untuk diagnosis dan menjadi perangkat tambahan dalam bidang medis.

#### DAFTAR RUJUKAN

- Amrane, M., Oukid, S., Gagaoua, I., & EnsarĪ, T. (2018). Breast cancer classification using machine learning. *IEEE Conference Publication*.
- Auliasari, K., & Kertaningtyas, M. (2018). Studi Komparasi Klasifikasi Pola Tekstur Citra Digital Menggunakan Metode K-Means Dan Naive Bayes. *Jurnal Informatika*, 18(2), 175–185.
- Azhagusundari, B., & Thanamani, A. S. (2013). Feature Selection based on Information Gain. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 2(2), 18–21.
- Fajrin, H. R., Nugroho, H. A., & Soesanti, I. (2015). Ekstraksi Ciri Berbasis Wavelet Dan Glcm Untuk Deteksi Dini Kanker Payudara Pada Citra Mammogram. *Snst*, 47–52.
- Fitriani, I. R. (2014). Peningkatan Metode Naive Bayes Classification untuk Penentuan Tingkat Keganasan Kanker Payudara Menggunakan Particle Swarm Optimiza-tion.(Graduate .... *Universitas Dian Nuswantoro. Jawa Tengah*.  
[http://eprints.dinus.ac.id/13137/1/jurnal\\_13546.pdf](http://eprints.dinus.ac.id/13137/1/jurnal_13546.pdf)
- Frank, E., Hall, M. A., & Witten, I. H. (2017). The WEKA workbench. *Data Mining*, 553–571.  
<https://doi.org/10.1016/b978-0-12-804291-5.00024-6>
- Harismah, K., Vitasari, D., Fuadi, A. M., & Setyawan, T. (2016). *Diversifikasi Minyak Sereh Dan*

- Minyak Cengkeh Sebagai Bahan Repellent Pembuatan Tinta Anti Nyamuk*. 526–531.
- Kemkes. (2019). Kanker Payudara Paling Banyak di Indonesia, Kemenkes Targetkan Pemerataan Layanan Kesehatan. In *Kementerian Kesehatan RI* (p. 1). <https://www.kemkes.go.id/article/view/19093000001/penyakit-jantung-penyebab-kematian-terbanyak-ke-2-di-indonesia.html>
- Ladha, L., & Deepa, T. (2011). Feature Selection Methods And Algorithms. *International Journal on Computer Science and Engineering*, 3(5), 1787–1797. <http://journals.indexcopernicus.com/abstract.php?icid=945099>
- Listia, R., & Harjoko, A. (2014). Klasifikasi Massa pada Citra Mammogram Berdasarkan Gray Level Cooccurrence Matrix (GLCM). *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 8(1), 59. <https://doi.org/10.22146/ijccs.3496>
- Marlina, L., Munadi, K., & Arnia, F. (2020). Klasifikasi Kanker Payudara Menggunakan Teknik Support Vector Machine (Svm) Pada Citra Termografi. *Jurnal Komputer, Informasi Teknologi, Dan Elektro*, 5(2), 11–13. <https://doi.org/10.24815/kitektro.v5i2.15571>
- Oktavianto, H., & Handri, R. P. (2020). Analisis Klasifikasi Kanker Payudara Menggunakan Algoritma Naive Bayes. *INFORMAL: Informatics Journal*, 4(3), 117. <https://doi.org/10.19184/isj.v4i3.14170>
- Olaniyi, E. O., Adekunle, A. A., Odekuoye, T., & Khashman, A. (2017). Automatic system for grading banana using GLCM texture feature extraction and neural network arbitrations. *Journal of Food Process Engineering*, 40(6). <https://doi.org/10.1111/jfpe.12575>
- Purwaningsih, N., Soesanti, I., & Nugroho, H. A. (2015). Ekstraksi Ciri Tekstur Citra Kulit Sapi Berbasis Co-Occurrence Matrix. *Seminar Nasional Teknologi Informasi Dan Multimedia*, 6–8.
- Putra, D. K. (2011). Identifikasi keberadaan kanker pada citra mammografi menggunakan metode wavelet haar. *Jurnal Ilmiah Teknik Elektro*, 1–7.
- Saiful rahman, A. F., B, A. A., & Kurniawan, S. D. (2019). Identifikasi Citra Daun Dengan Menggunakan Metode Deep Learning Convolutional Neural Network (Cnn). *Jurnal Teknik Elektro Uniba (JTE Uniba)*, 4(1), 23–28. <https://doi.org/10.36277/jteuniba.v4i1.55>
- Singh, D., & Kaur, K. (2012). Classification of Abnormalities in Brain MRI Images Using GLCM , PCA and SVM. *International Journal of Engineering and Advanced Technology (IJEAT)*, 1(6), 243–248.
- WHO. (2018). *Cancer*. <https://www.who.int/health-topics/cancer>
- Widodo, S., Informasi, J. S., Gunadarma, U., Studi, P., & Informasi, S. (2012). Representasi Kode IRMA pada. *Seminar Nasional Teknologi Informasi Komunikasi Dan Industri*

Jullia, dkk

(*SNTIKI*) 4, 198–202.

Zamrodah, Y. (2016). *Bayesian Artificial Intelligence (Second Edition)* (Vol. 15, Issue 2). CRC Press.