

# **EKSTRAKSI FITUR CITRA MAMMOGRAM BERBASIS TEKSTUR MENGUNAKAN METODE *GRAY LEVEL COOCCURRENCE MATRIX* (GLCM)**

**Fifin Dewi Ratnasari<sup>\*</sup>, Adlinia Hidayati**

*Program Studi Fisika, Fakultas Matematika Ilmu Pengetahuan Alam, Universitas Negeri Semarang*

*\*Corresponding author: fifin\_fisika@mail.unnes.ac.id*

## **ABSTRAK**

*Kanker adalah penyakit yang disebabkan oleh pembelahan sel yang tidak normal di dalam tubuh. Kanker dapat menyerang semua jaringan tubuh. Kanker payudara secara umum terbagi menjadi 2 jenis, yaitu jinak (benign) dan ganas (malignant). Menurut World Health Organization (WHO) Global Cancer Observatory tahun 2018-2020 mengalami peningkatan kasus kanker baru di Indonesia. Pengolahan data citra medis berbasis komputer memudahkan dokter untuk mengamati kelainan secara cepat dan akurat. Gambar yang ditinjau secara manual kurang efektif untuk diagnosis. Oleh karena itu penelitian bertujuan sebagai sarana pengembangan penelitian pada sistem ekstraksi fitur citra mammogram berbasis tekstur sebagai pendukung diagnosis kanker payudara dengan metode Gray Level Cooccurrence Matrix (GLCM) pada software Matlab R2017a. Untuk objek penelitian yang digunakan yaitu citra mammogram yang berasal dari Kaggle RSNA Mammogram Dataset dengan jumlah 1600 citra normal dan 1600 citra abnormal yang disimpan dalam bentuk \*.png. Berdasarkan dari ujicoba yang dilakukan didapatkan total 8 fitur mammogram, pada penelitian sebelumnya hanya menggunakan 4 fitur. 8 fitur tersebut terdiri dari contrast, correlation, homogeneity, energy, mean, skewness, kurtosis, dan variance.*

Kata kunci : ekstraksi fitur, GLCM, kanker payudara, mammogram.

## **PENDAHULUAN**

### **A. Kanker Payudara**

Berdasarkan statistik Organisasi Kesehatan Dunia, kuantitas pengidap kanker di dunia meningkat sekitar 7 juta orang setiap tahun, dan dua pertiganya berada di negara berkembang (Handayani *et al.*, 2014). Lebih dari 21% wanita pengidap kanker adalah pasien kanker payudara, dan lebih dari 60% kematian di negara berkembang disebabkan oleh kanker payudara (Novar Setiawan & Suwija Putra, 2018).

Kanker adalah penyakit yang disebabkan oleh pembelahan sel yang tidak normal di dalam tubuh. Kanker dapat menyerang semua jaringan tubuh (Setiawan, 2021). Kanker payudara adalah penyakit yang tidak menular atau menyebar (WHO, 2021). Menurut *World Health Organization (WHO) Global Cancer Observatory* tahun 2018, kasus kanker terbanyak di Indonesia adalah kanker payudara, yaitu

sebanyak 58.256 dari total 348.809 kasus kanker (Hari Widowati, 2019). Pada tahun 2020, dari total 396.914 kasus kanker baru di Indonesia, jumlah kasus kanker payudara baru mencapai 68.858. Sementara itu, korban meninggal mencapai lebih dari 22.000 jiwa kasus (Rokom, 2022). Kanker payudara secara umum terbagi menjadi 2 jenis, yaitu jinak (*benign*) dan ganas (*malignant*) (Cahyanti *et al.*, 2020). Pengolahan data citra medis berbasis komputer memudahkan dokter untuk mengamati kelainan secara cepat dan akurat. Gambar yang ditinjau secara manual kurang efektif untuk diagnosis (Syaputri, 2019). Oleh karena itu penelitian bertujuan sebagai sarana pengembangan penelitian pada *sistem ekstraksi fitur citra mammogram berbasis tekstur sebagai pendukung diagnosis kanker payudara* dengan metode Gray Level Cooccurrence Matrix (GLCM) pada software Matlab R2017a.

## **B. Dasar Teori**

Tabung sinar-X mamografi dirancang khusus dan didedikasikan untuk pencitraan jaringan payudara. Unit mamografi dirancang untuk fleksibilitas dalam penentuan posisi pasien dan dapat dimodifikasi untuk melakukan pemeriksaan payudara secara khusus (Robert Fosbinder, 2012). Dari beberapa prosedur deteksi dan diagnosis, mamografi merupakan alat yang paling efektif untuk deteksi dini kanker payudara, sebelum gejala fisik kanker muncul. Mammografi adalah jenis pencitraan sinar-X yang memberikan gambar *visual* payudara yang mendetail dan sistem sinar-X yang dirancang khusus untuk pencitraan payudara (Hermawan *et al.*, 2019).

Ada pula metode ekstraksi fitur yang menggunakan perhitungan tekstur yang disebut *Gray Level Co-Occurrence Matrix* (GLCM) (Novar Setiawan & Suwija Putra, 2018). Ekstraksi fitur merupakan langkah penting dalam pendukung proses klasifikasi pola, yang bertujuan untuk mengekstraksi informasi relevan yang mencirikan setiap kategori. Fitur ini digunakan untuk mengidentifikasi *unit input* dan menargetkan *unit output*. Klasifikasi ini lebih mudah dilakukan antara kelas yang berbeda dengan melihat fitur yang ada (Varuna Shree & Kumar, 2018). Konsep

dasarnya melibatkan pemeriksaan gambar untuk melacak tingkat keabuan antara setiap dua piksel yang dipisahkan oleh jarak tertentu (Ginantra, 2016).

Ciri fitur GLCM yang digunakan diantaranya adalah sebagai berikut:

1) *Contrast*

*Contrast* merupakan hasil perhitungan yang berkaitan dengan jumlah keberagaman intensitas keabuan dalam citra. Apabila suatu piksel dengan piksel tetangganya mempunyai nilai intensitas (tingkat keabuan) yang berdekatan, maka kontras tekstur sangat rendah atau kontras akan bernilai 0. Dan apabila suatu piksel dengan piksel tetangganya mempunyai nilai intensitas (tingkat keabuan) yang berjauhan, maka kontras tekstur tinggi atau kontras akan bernilai 1 (Maulida *et al.*, 2022).

2) *Correlation*

*Correlation* merupakan representasi dari keterkaitan linear atau menunjukkan ukuran ketergantungan linear pada derajat citra grayscale atau derajat keabuan pada citra dari piksel-piksel yang saling bertetangga dalam suatu citra abu-abu sehingga dapat memberikan petunjuk adanya struktur linear dalam citra (Maulida *et al.*, 2022).

3) *Energy*

*Energy* merupakan hasil perhitungan yang berkaitan dengan jumlah keberagaman atau banyaknya perbedaan intensitas (tingkat keabuan) dalam citra. Energi memiliki nilai yang tinggi ketika citra memiliki homogenitas yang baik atau nilai piksel yang hampir serupa. Jika semakin tinggi nilai energi, maka semakin tinggi pula kemiripan suatu citra, dan sebaliknya (Maulida *et al.*, 2022).

4) *Homogeneity*

*Homogeneity* bertujuan untuk mengukur tingkat homogenitas atau kesamaan variasi dalam intensitas keabuan (tingkat keabuan) citra pada matriks *co-occurrence*. *Homogeneity* bernilai tinggi jika pasangan piksel mempunyai nilai keabuan yang seragam. Sama halnya dengan energi, jika semakin tinggi nilai *homogeneity* maka semakin tinggi pula kemiripan suatu citra dan begitu pula sebaliknya (Maulida *et al.*, 2022).

Dari beberapa tabel hasil ekstraksi fitur citra *mammogram* dapat dilihat bahwa nilai-nilai ciri tekstur menggunakan GLCM pada masing-masing citra *mammogram* kanker payudara berbeda-beda, sehingga hasil dari ekstraksi fitur citra *mammogram* ini sangat membantu bagi radiolog maupun dokter ahli untuk membedakan citra satu dengan citra yang lainnya. Ekstraksi fitur tidak bisa dilakukan jika nilai-nilai ekstraksi dari citra sama, karena tujuan dari ekstraksi fitur itu sendiri adalah untuk menghasilkan nilai fitur citra yang membedakan antara citra satu dengan citra yang lainnya (Maulida *et al.*, 2022).

Ciri orde pertama yang digunakan yaitu:

1) *Mean*

*Mean* merupakan ukuran dari nilai rata-rata yang ada pada citra.

2) *Variance*

*Variance* merupakan variasi elemen-elemen matriks kookurensi. Pada citra yang memiliki transisi derajat keabuan kecil akan memiliki variansi yang kecil.

3) *Skewness*

*Skewness* merupakan fitur yang menunjukkan tingkat kemiringan relatif kurva histogram suatu citra.

4) *Kurtosis*

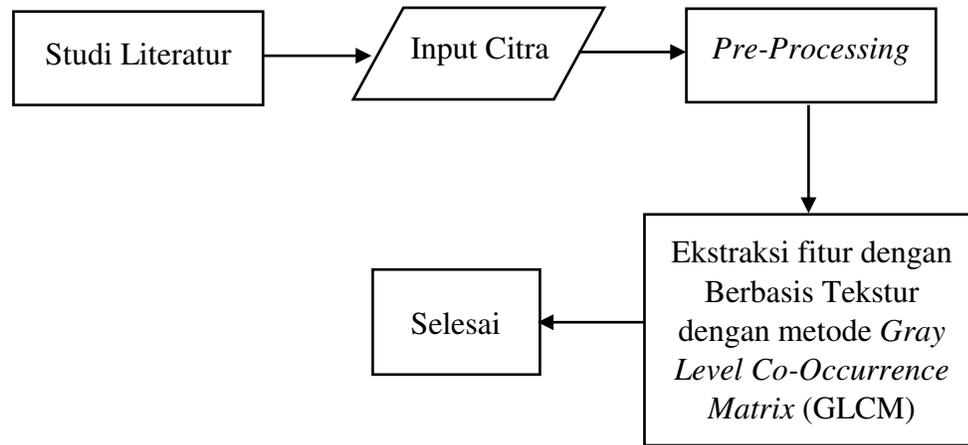
*Kurtosis* merupakan tingkat keruncingan relatif kurva histogram yang ada pada suatu citra (Retno Paras Rasmi, 2020).

## **METODE**

Metode yang digunakan dalam penelitian ini menggunakan pendekatan kualitatif dengan metode *Gray Level Co-Occurrence Matrix* (GLCM). Penelitian ini akan dilaksanakan di Laboratorium Fisika Medik D9, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Semarang. Alat yang digunakan untuk penelitian adalah Laptop merk Lenovo Thinkpad X240 dan *software* Matlab R2017a sedangkan bahan penelitian menggunakan objek yaitu citra *mammogram* payudara yang normal dan abnormal dengan disimpan dalam bentuk *\*.png*. Data yang

dipergunakan dalam penelitian ini merupakan dengan data sekunder dari citra *mammogram* payudara yang terindikasi mengandung lesi kanker payudara diambil dari *Kaggle RSNA Mammogram Dataset*.

### Alur Penelitian



Gambar 1. Alur penelitian.

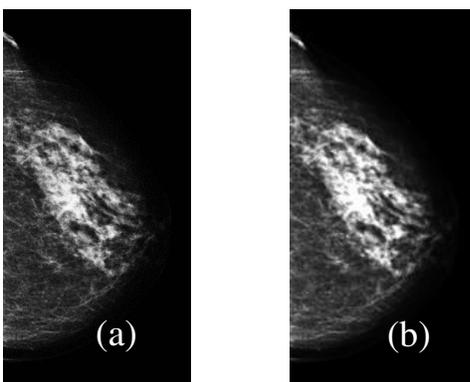
Gambar 1 menjelaskan mengenai skema alur penelitian yang akan dilakukan dimana penelitian dimulai dengan studi literatur Selanjutnya input citra *mammogram* payudara yang dari *Kaggle RSNA Mammogram Dataset*. Kemudian, hasil *input* data citra dianalisis menggunakan program Matlab R2017a untuk proses *pre-proccesing* dilanjut dengan ekstraksi fitur berbasis tekstur dengan program Matlab R2017a dengan metode *Gray Level Co-Occurrence Matrix* (GLCM). Hasil ekstraksi fitur yang didapatkan akan dianalisa dan dibahas untuk mendapatkan kesimpulan penelitian.

## HASIL DAN PEMBAHASAN

### A. Hasil Penelitian

Penelitian ini melakukan proses *preprocessing* sebelum melakukan ekstraksi fitur. Tahap *preprocessing* terdiri dari beberapa tahapan yang mempengaruhi langkah pengolahan selanjutnya dan hasil yang diperoleh. Terdapat dua proses pada tahap

*preprocessing*, yang pertama adalah proses normalisasi citra yaitu proses *resize* pada proses normalisasi citra yang digunakan untuk mengubah ukuran citra. Tahap kedua adalah proses peningkatan citra, yaitu proses perbaikan kualitas citra dengan menggunakan teknik penajaman. (*sharpening*) (Junita, 2017). Lakukan *preprocessing* untuk meningkatkan kualitas gambar (Sholihin & Rohman, 2018). Hasil dari proses ini adalah jelasnya perbedaan kontras antara daerah yang memiliki lesi kanker payudara dengan daerah payudara yang tidak memiliki lesi kanker dapat dilihat pada Gambar 2.



Gambar 2. Perbedaan kontras setelah dilakukan *pre-processing*: (a) sebelum (b) sesudah

Setelah melakukan berbagai tahapan dalam proses *image preprocessing*. Pada bagian ini, hasil ekstraksi fitur diperoleh dengan menggunakan hasil ekstraksi input yang dilakukan oleh perangkat lunak Matlab. Ekstraksi fitur dilakukan dengan menggunakan 8 fitur tekstur dan data tambahan dari *dataset mammogram Kaggle RSNA*. Ekstraksi fitur dilakukan untuk menentukan kelas normal dan *abnormal*. Dapat dilihat pada Tabel 1 merupakan tabel rata-rata ekstraksi fitur dari kedua jenis *mammogram*.

Tabel 1. Rata-Rata Ekstraksi Fitur dari Kedua Jenis *Mammogram*.

Ekstraksi Ciri	Kelas 1 (Normal)	Kelas 2 ( <i>Abnormal</i> )
<i>Contrast</i>	1.802622993	2.282232233
<i>Correlation</i>	0.736429582	0.767615394
<i>Homogeneity</i>	0.967810304	0.959245853
<i>Energy</i>	0.658146744	0.599795747
<i>Mean</i>	33.80922879	32.97928415

<i>Skewness</i>	2.437992612	2.439967274
<i>Kurtosis</i>	9.114105459	8.631666234
<i>Varians</i>	3196.35115	3817.255096

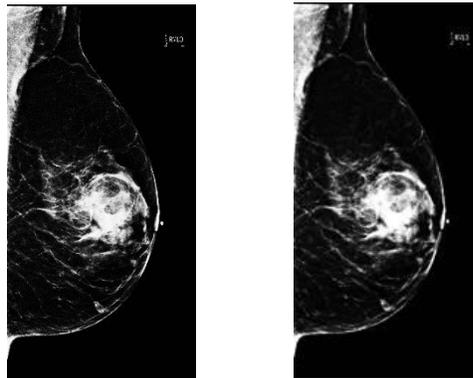
## B. Pembahasan Penelitian

### a) *Pre-Processing*

Implementasi proses *preprocessing* menggunakan program Matlab sebagai berikut.

#### 1) *Resizing* citra

Proses ini bertujuan untuk mengatur ukuran citra dengan format gambar *\*.png* secara serentak bersamaan yang semula memiliki  $256 \times 512$  kemudian diubah menjadi ukuran  $128 \times 256$  *pixel* dapat dilihat pada Gambar 3.

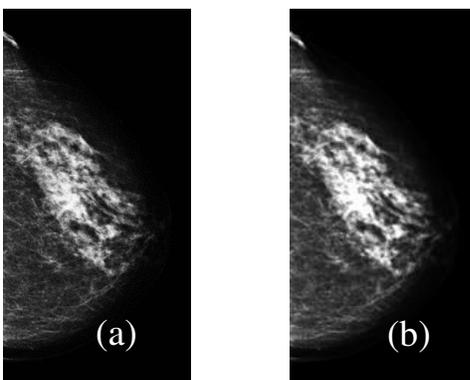


Gambar 3. (a) Citra *mammogram* sebelum *resizing*  $256 \times 512$  *pixel* (b) Citra *mammogram* sesudah *resizing*  $128 \times 256$  *pixel*.

*Resizing* gambar adalah teknik *preprocessing* yang umum digunakan dalam banyak penelitian karena dapat mengurangi kompleksitas komputasi, mempercepat waktu pelatihan, dan mengoptimalkan penggunaan memori (Sunardi *et al.*, 2022).

#### 2) *Sharpening*

Proses ini bertujuan untuk peningkatan kualitas citra dengan menggunakan teknik penajaman. Sehingga terdapat perbedaan kontras citra antara daerah yang memiliki lesi kanker payudara dengan daerah payudara yang tidak memiliki lesi kanker (Firma & Oktamuliani, 2022). Sebelum melakukan *sharpening*, gambar citra diubah terlebih dahulu menjadi *grayscale* agar informasi warna yang kurang relevan dihilangkan (Ainani Shabrina Febrianti *et al.*, 2020).



Gambar 4. Perbedaan kontras setelah dilakukan proses perbaikan kualitas citra: (a) sebelum (b) sesudah.

#### **b) Proses Ekstraksi Fitur Berbasis Tekstur**

Ekstraksi fitur adalah proses mendapatkan fitur unik dari suatu gambar yang dapat membedakannya dari objek lain. Ekstraksi fitur adalah langkah pertama dalam mengklasifikasikan dan menginterpretasikan citra. Fitur yang diekstraksi dari objek yang diduga bersifat kanker (Fuad *et al.*, 2016). Pada penelitian ini untuk ekstraksi fitur yang digunakan yaitu contrast, correlation, homogeneity, energy, mean, skewness, kurtosis, dan variance yang dilakukan dengan menggunakan program Matlab dengan metode *Gray Level Cooccurrence Matrix (GLCM)*.

Metode GLCM adalah metode ekstraksi fitur dalam fungsi statistik orde kedua, dan co-occurrence didefinisikan sebagai distribusi gabungan tingkat keabuan dari dua piksel yang dipisahkan oleh jarak dan arah tertentu (Safitri & Abadi, 2016). Implementasi fungsi *graycomatrix* digunakan untuk menghitung matriks ko-eksistensi keabuan (GLCM) dari citra *grayscale* (Fahrurozi, 2014).

Ditinjau dari Tabel 1 didapatkan perbandingan rata-rata hasil dari ekstraksi fitur menggunakan program Matlab. Pada fitur contrast kelas 1 didapatkan rata-rata sebesar 1,802622993 dan pada kelas 2 didapatkan rata-rata sebesar 2,282232233, pada fitur correlation kelas 1 didapatkan rata-rata sebesar 0,736429582 dan pada kelas 2 didapatkan rata-rata sebesar 0,767615394, pada fitur homogeneity kelas 1 didapatkan rata-rata sebesar 0,967810304 dan pada kelas 2 didapatkan rata-rata sebesar 0,959245853, pada fitur energy kelas 1 didapatkan rata-rata sebesar 0,658146744 dan pada kelas 2 didapatkan rata-rata sebesar 0,599795747, pada fitur

mean kelas 1 didapatkan rata-rata sebesar 33,80922879 dan pada kelas 2 didapatkan rata-rata sebesar 32,97928415, pada fitur skewness kelas 1 didapatkan rata-rata sebesar 2,437992612 dan pada kelas 2 didapatkan rata-rata sebesar 2,439967274, pada fitur kurtosis kelas 1 didapatkan rata-rata sebesar 9,114105459 dan pada kelas 2 didapatkan rata-rata sebesar 8,631666234, pada fitur variance kelas 1 didapatkan rata-rata sebesar 3196,35115 dan pada kelas 2 didapatkan rata-rata sebesar 3817,255096.

Dapat dilihat pada tabel rata-rata fitur variance jauh lebih besar dibandingkan tujuh fitur lainnya. Hal ini mungkin terjadi karena variance merupakan ukuran statistik yang menggambarkan tingkat variasi atau perbedaan nilai intensitas piksel dalam suatu daerah atau objek dalam citra. Jika varian fitur sangat besar, ini menunjukkan adanya variasi tekstur yang signifikan dalam daerah tersebut.

Dalam konteks ekstraksi fitur berbasis tekstur, ini dapat berarti adanya perubahan tekstur yang mencolok dalam objek atau daerah tersebut dibandingkan dengan lingkungan sekitarnya. Oleh karena itu, fitur variace yang tinggi dapat digunakan untuk mengidentifikasi objek atau daerah kanker payudara dengan tekstur yang lebih kompleks atau unik.

## **SIMPULAN**

Pengembangan sebuah sistem ekstraksi fitur citra mammogram berbasis tekstur dikembangkan sebagai hasil penelitian untuk mendukung diagnosis kanker payudara. Pada penelitian ini didapatkan total 8 fitur mammogram, pada penelitian sebelumnya hanya menggunakan 4 fitur. 8 fitur tersebut terdiri dari contrast, correlation, homogeneity, energy, mean, skewness, kurtosis, dan variance dengan metode Gray Level Co-Occurrence Matrix (GLCM).

## **DAFTAR PUSTAKA**

Ainani Shabrina Febrianti, Tri Arief Sardjono, & Atar Fuady Babgei. (2020). Klasifikasi Tumor Otak pada Citra Magnetic Resonance Image dengan Menggunakan Metode Support Vector Machine. *Jurnal Teknik ITS*, Vol. 9, No(1).

- Cahyanti, D., Rahmayani, A., & Husniar, S. A. (2020). Analisis Performa Metode KNN pada Dataset Pasien Pengidap Kanker Payudara. *Indonesian Journal of Data and Science*, 1(2), 39–43. <https://doi.org/10.33096/ijodas.v1i2.13>.
- Fahrurozi, A. (2014). Klasifikasi Kayu Dengan Menggunakan Naïve Bayes-Classifier.
- Firma, A., & Oktamuliani, S. (2022). Pengolahan Filtering Dan Contrast Enhancement untuk. 8(1), 51–54.
- Fuad, N., Studi, P., Informatika, T., & Islam, U. (2016). Klasifikasi Kanker pada Citra Mammogram Berdasarkan Fitur Moment Invariant Dan. *Semantikom*, 30–37.
- Ginantra, N. L. W. S. R. (2016). Deteksi Batik Parang Menggunakan Fitur Co-Occurrence Matrix dan Geometric Moment Invariant Dengan Klasifikasi KNN. Lontar Komputer: *Jurnal Ilmiah Teknologi Informasi*, 7(1), 40. <https://doi.org/10.24843/lkjiti.2016.v07.i01.p05>.
- Handayani, L., Septa, I. I., Teknik, J., & Fakultas, I. (2014). Segmentasi Mamografi Kanker Payudara dengan Algoritma Expectation Maximization Segmentation ( Em-Segmentation ). *Jurnal Sains*, 11(2), 251–258.
- Hari Widowati. (2019). Kasus Kanker Payudara Paling Banyak Terjadi di Indonesia. Databoks. [https://databoks.katadata.co.id/datapublish/2019/06/03/kasus-kanker-payudara-paling-banyak-terjadi-di-indonesia#:~:text=Kementerian Kesehatan %28Kemenkes%29 menyatakan%2C angka kanker payudara di,Indonesia mencapai 23%2C4 orang per 100 ribu penduduk.](https://databoks.katadata.co.id/datapublish/2019/06/03/kasus-kanker-payudara-paling-banyak-terjadi-di-indonesia#:~:text=Kementerian Kesehatan%28Kemenkes%29%20menyatakan%2C%20angka%20kanker%20payudara%20di%20Indonesia%20mencapai%2023%2C4%20orang%20per%20100%20ribu%20penduduk.)
- Hermawan, E., Pendidikan, F., & Pengetahuan, I. (2019). Active Contour Lankton untuk Segmentasi Kanker Payudara pada Citra Mammogram. 28–37. <https://doi.org/10.30864/eksplora.v9i1.258>
- Junita, B. D. (2017). Ekstraksi Fitur dan Klasifikasi Menggunakan Metode GLCM dan SVM pada Citra Mammogram untuk Identifikasi Kanker Payudara. *Jurnal Teknologi Rekayasa*, 22(1), 18–26.
- Maulida, A., Nurhidayah, N., Fendriani, Y., & Haryono, H. (2022). Segmentasi Citra Mammogram untuk Deteksi Dini Kanker Payudara dengan Menggunakan Metode Otsu Thresholding. *Jurnal Fisika Unand*, 11(2), 180–186. <https://doi.org/10.25077/jfu.11.2.180-186.2022>.
- Novar Setiawan, K., & Suwija Putra, I. M. (2018). Klasifikasi Citra Mammogram Menggunakan Metode K-Means, GLCM, dan Support Vector Machine (SVM). *Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi)*, 6(1), 13. <https://doi.org/10.24843/jim.2018.v06.i01.p02>.

- Retno Paras Rasmi. (2020). Peningkatan Hasil Diagnosis Kanker Payudara dari Hasil Citra Mammogram Menggunakan Metode Ekstraksi Ciri dan Klasifikasi.
- Robert Fosbinder, D. O. (2012). *Essentials of Radiologic Science*.
- Rokom. (2022). Kanker Payudara Paling Banyak di Indonesia, Kemenkes Targetkan Pemerataan Layanan Kesehatan. Kementerian Kesehatan Republik Indonesia Sehat Negeriku, Sehatlah Bangsa. <https://sehatnegeriku.kemkes.go.id/baca/umum/20220202/1639254/kanker-payudaya-paling-banyak-di-indonesia-kemenkes-targetkan-pemerataan-layanan-kesehatan/>.
- Safitri, W., & Abadi, A. M. (2016). Implementasi Sistem Fuzzy Sugeno Orde Satu pada Diagnosis Implementation of First-Order Sugeno Fuzzy System on Diagnosing Breast. *Jurnal Pendidikan Matematika Dan Sains, 2009*.
- Setiawan, W. (2021). Klasifikasi Citra Histopatologi Kanker Payudara menggunakan Data Resampling Random dan Residual Network. *Jurnal Sistem Informasi Bisnis, 11*(1), 70–77. <https://doi.org/10.21456/vol11iss1pp70-79>.
- Sholihin, M., & Rohman, M. G. (2018). Klasifikasi Kualitas Mutu Telur Ayam Ras Berdasarkan Fitur Warna dan Tekstur. *Jurnal Teknika, 10*(2), 1056–1059. <https://doi.org/10.30736/teknika.v10i2.244>.
- Sunardi, S., Yudhana, A., & Wijaya, S. A. (2022). Penerapan Metode Median Filtering untuk Optimasi Deteksi Wajah pada Foto Digital. *Journal of Innovation Information Technology and Application (JINITA), 4*(1), 51–60. <https://doi.org/10.35970/jinita.v4i1.1214>.
- Syaputri, S. (2019). Komunikasi Fisika Indonesia Segmentasi Citra Paru-Paru Menggunakan Metode. *16*(2), 91–95.
- Varuna Shree, N., & Kumar, T. N. R. (2018). Identification and Classification of Brain Tumor MRI Images with Feature Extraction Using DWT and Probabilistic Neural Network. *Brain Informatics, 5*(1), 23–30. <https://doi.org/10.1007/s40708-017-0075-5>.
- WHO, W. H. O. (2021). Breast Cancer. <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>.