

Klasifikasi Kanker Paru pada Citra CT Scan menggunakan *Extreme Learning Machine* dan *Histogram Equalization*

Omar Pahlevi^{1*}, Yopi Handrianto², Dewi Ayu Nur Wulandari³, Henny Leidiyana⁴, Luci Kanti Rahayu⁵

^{1,3,4} Program Studi Sistem Informasi, Fakultas Teknik dan Informatika,

² Program Studi Teknologi Informasi, Fakultas Teknik dan Informatika,

⁵ Program Studi Informatika, Fakultas Teknik dan Informatika,

Universitas Bina Sarana Informatika

*Email: omar.opi@bsi.ac.id

Abstrak

Kanker paru-paru merupakan salah satu jenis kanker dengan tingkat kematian tertinggi di dunia, sehingga deteksi dini menjadi sangat penting untuk meningkatkan peluang kesembuhan. Salah satu metode utama untuk mendeteksi kanker paru adalah melalui citra *Computed Tomography (CT) scan*, namun analisis otomatis citra ini menghadapi tantangan karena kualitas citra yang dipengaruhi oleh noise dan kontras rendah. Penelitian ini bertujuan untuk membangun model klasifikasi kanker paru-paru dari citra CT scan menggunakan algoritma *Extreme Learning Machine (ELM)* dan ekstraksi fitur *Gray Level Co-occurrence Matrix (GLCM)* yang didukung oleh teknik *Histogram Equalization* untuk meningkatkan kualitas citra. *Histogram Equalization* digunakan untuk meningkatkan kontras citra sehingga mempermudah proses ekstraksi fitur tekstur dari GLCM, seperti kontras, homogenitas, energi, dan entropi. ELM dipilih karena kecepatan dan performa akurasi yang baik dalam klasifikasi data medis yang kompleks. Hasil penelitian menunjukkan bahwa model yang diusulkan berhasil meningkatkan kinerja klasifikasi dengan akurasi mencapai 91.06%. Kombinasi antara teknik ELM dan *Histogram Equalization* menghasilkan sistem klasifikasi yang efisien dan akurat untuk mendeteksi kanker paru-paru dari citra CT Scan.

Kata kunci: CT Scan, *Extreme Learning Machine*, GLCM, *Histogram Equalization*, Kanker Paru-paru

Abstract

Lung cancer is one of the deadliest types of cancer worldwide, making early detection crucial to improving patient survival rates. One of the primary methods for detecting lung cancer is through *Computed Tomography (CT) scan* images. However, automated analysis of these images faces challenges due to image quality being affected by noise and low contrast. This study aims to develop a lung cancer classification model from CT scan images using the *Extreme Learning Machine (ELM)* algorithm and *Gray Level Co-occurrence Matrix (GLCM)* feature extraction, supported by *Histogram Equalization* techniques to enhance image quality. *Histogram Equalization* is employed to improve image contrast, facilitating the extraction of texture features from GLCM, such as contrast, homogeneity, energy, and entropy. ELM was chosen for its speed and accuracy in handling complex medical image classification tasks. The study results demonstrate that the proposed model successfully enhances classification performance with an accuracy of 91.06%. The combination of ELM and *Histogram Equalization* techniques produces an efficient and accurate classification system for detecting lung cancer from CT scan images.

Keywords: CT Scan, *Extreme Learning Machine*, GLCM, *Histogram Equalization*, Lung Cancer

PENDAHULUAN

Sebagai salah satu penyakit kanker yang paling mematikan secara global, kanker paru-paru menempati posisi teratas dalam hal angka mortalitas. *World Health Organization (WHO)* mencatat bahwa dari total kematian yang disebabkan oleh kanker per tahun, sekitar 18%

merupakan akibat dari kanker paru-paru (Nugroho et al., 2023). Deteksi dini kanker paru sangat penting untuk meningkatkan peluang kesembuhan pasien, karena kanker ini sering kali ditemukan pada stadium lanjut (Sugiharto et al., 2021). *CT scan (Computed Tomography Scan)* menjadi salah satu teknik pemeriksaan

utama dalam mengidentifikasi kanker paru-paru, dimana teknologi ini mampu menghasilkan citra terperinci dari anatomi paru-paru.

Namun, meskipun citra *CT scan* mampu memberikan gambaran mendetail, tantangan dalam menganalisis gambar tersebut tetap besar. Kualitas citra yang dipengaruhi oleh *noise* dan kontras rendah sering kali menghambat proses diagnosis otomatis (Ulga et al., 2024). Oleh karena itu, diperlukan teknik peningkatan citra yang dapat membantu meningkatkan kualitas visual dari gambar, sehingga proses ekstraksi fitur dan klasifikasi menjadi lebih akurat. Salah satu teknik yang digunakan untuk meningkatkan kualitas gambar adalah *Histogram Equalization*. Teknik ini berfungsi untuk menyeimbangkan distribusi intensitas piksel pada citra, sehingga detail penting yang tersembunyi dalam area kontras rendah dapat ditingkatkan dan lebih mudah dikenali (Winarno et al., 2022).

Untuk proses klasifikasi kanker paru dari citra *CT scan*, algoritma pembelajaran mesin memainkan peran penting dalam mengidentifikasi pola-pola yang relevan. Pada penelitian ini, digunakan *Extreme Learning Machine* (ELM) sebagai algoritma klasifikasi. ELM dipilih karena memiliki keunggulan dalam hal kecepatan dan performa akurasi yang baik (Wahid et al., 2023). Algoritma ini dirancang untuk jaringan saraf tiruan dengan satu lapisan tersembunyi, yang memungkinkannya memproses data dengan sangat cepat (Yel et al., 2023). Kecepatan ini membuat ELM sangat cocok untuk diaplikasikan dalam bidang kesehatan, di mana kebutuhan untuk mendapatkan hasil yang cepat dan akurat sangat penting. Selain itu, untuk ekstraksi fitur dari citra, digunakan metode *Gray Level Co-occurrence Matrix* (GLCM). GLCM mampu mengekstraksi fitur tekstur dengan menganalisis hubungan spasial antara pasangan piksel dengan nilai intensitas tertentu dalam suatu citra (Ihsan & Sunyoto, 2022). Kombinasi antara teknik ekstraksi fitur GLCM dan algoritma klasifikasi ELM diharapkan dapat memberikan hasil klasifikasi yang akurat dan efisien.

Secara keseluruhan, penelitian ini bertujuan untuk mengembangkan sistem klasifikasi kanker paru dari citra *CT scan* melalui pendekatan *Extreme Learning Machine*

(ELM) dan teknik ekstraksi fitur GLCM, yang didukung oleh proses peningkatan citra menggunakan *Histogram Equalization*. Dengan pendekatan ini, dapat diperoleh hasil yang lebih akurat dalam mendeteksi dan mengklasifikasikan kanker paru dari citra *CT scan*, yang pada akhirnya dapat berkontribusi pada peningkatan prognosis pasien.

TINJAUAN PUSTAKA

Dalam dunia medis, pengolahan citra telah menjadi alat penting dalam membantu diagnosis berbagai penyakit, termasuk kanker paru-paru. Teknologi pencitraan seperti *CT scan* memungkinkan visualisasi struktur organ dalam tubuh secara detail, memudahkan deteksi dini dan penilaian tingkat keparahan penyakit (Fendriani et al., 2023). Penelitian terdahulu tentang klasifikasi penyakit paru-paru dari citra *CT scan* menggunakan pengolahan citra telah dilakukan oleh banyak peneliti dengan pendekatan yang beragam. Sebuah studi menggunakan metode *Fuzzy Kernel C-Means* untuk mengklasifikasikan penyakit paru-paru dengan tingkat akurasi sebesar 74% (Rustam et al., 2021). Namun, algoritma ini sangat sensitif terhadap *outlier* dan *noise*, yang menyebabkan hasil pengelompokan kurang representatif. Penelitian lain menggunakan pendekatan *Support Vector Machine* (SVM) dan mencapai akurasi sebesar 83,33% (Firdaus et al., 2020). Akan tetapi, kelemahan dari SVM terletak pada sensitivitasnya terhadap pemilihan parameter, yang dapat mempengaruhi performa secara signifikan. Penelitian selanjutnya memanfaatkan jaringan saraf tiruan menggunakan SOM (*Self-Organizing Map*) yang mencapai tingkat akurasi 87% (Apsari et al., 2020). Namun, pendekatan ini bergantung pada parameter seperti ukuran peta dan fungsi tetangga, yang sulit untuk diatur dan memerlukan eksperimen yang mendalam. Penelitian lain yang menggunakan jaringan saraf tiruan dengan *Backpropagation* menghasilkan akurasi 90% (Putri et al., 2024). Meskipun demikian algoritma *Backpropagation* memiliki masalah konvergensi yang bisa membuat proses pelatihan menjadi lambat. Studi yang dilakukan oleh (Hossain et al., 2025), dimana mengembangkan metode klasifikasi kanker paru multikelas berbasis citra *CT scan* dengan memanfaatkan CNN-SVD *hybrid feature extraction* dan *ridge regression*

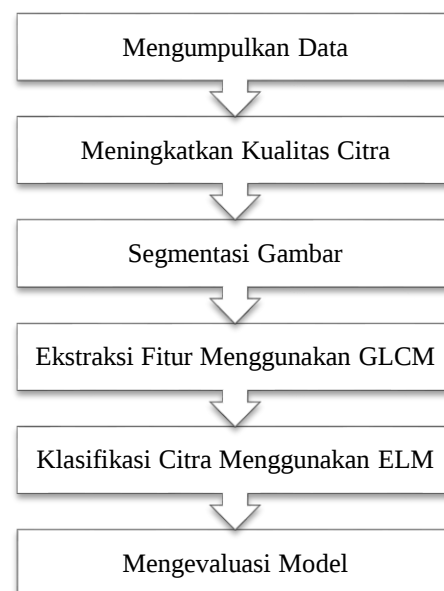
ELM sebagai *classifier*. Pemrosesan citra tersebut menggunakan *adaptive histogram equalization* (AHE) agar meningkatkan kontras. Pendekatan ini difokuskan pada efisiensi arsitektur dan ketepatan klasifikasi multikelas dengan akurasi mencapai 96%. Menggunakan pendekatan ELM tanpa CNN atau SVD, sehingga lebih ringan dan cocok untuk sistem *real-time* dan tahapan preprocessing pada penelitian yang dibuat ini dapat diarahkan ke segmentasi otomatis sebelum *histogram equalization*, dan menambahkan *pre-classification intelligence* ...

Dari penelitian-penelitian tersebut, terlihat bahwa penggunaan jaringan saraf tiruan dalam proses klasifikasi citra memperoleh hasil yang lebih baik. Hal ini disebabkan oleh kemampuan jaringan saraf tiruan yang dianggap efektif dalam mengidentifikasi dan mengklasifikasikan citra, karena bekerja dengan cara meniru fungsi jaringan saraf manusia (Borman, Ahmad, et al., 2022). Namun, penggunaan jaringan syaraf tiruan dengan kemampuan pelatihan dan generalisasi yang cepat dibutuhkan. ELM terkenal dengan kecepatannya dalam generalisasi data yang kompleks (Hizriadi et al., 2023). Untuk mendapatkan informasi tekstur dari citra *CT scan* dibutuhkan pendekatan *Gray Level Co-occurrence Matrix* (GLCM). GLCM memungkinkan pengambilan informasi tekstur penting pada citra medis, yang sangat berguna dalam mengenali pola pada jaringan paru-paru yang terkena kanker. Namun, penggunaan GLCM saja tidak cukup optimal jika kualitas citra rendah atau mengandung *noise* (Chen, 2020). Oleh karena itu, diperlukan teknik peningkatan kualitas citra, seperti *Histogram Equalization*, untuk meningkatkan kontras dan detail pada citra. *Histogram Equalization* merupakan salah satu teknik peningkatan citra yang digunakan untuk memperbaiki kualitas visual dengan menyesuaikan distribusi intensitas piksel (Kurnia & Hidayat, 2023). Metode ini bekerja dengan meratakan distribusi histogram dari intensitas citra, sehingga nilai intensitas piksel tersebar lebih merata di seluruh rentang dinamis yang tersedia (Manalu, 2021). Dalam konteks citra medis seperti *CT scan*, *Histogram Equalization* sangat berguna karena sering kali citra medis memiliki area dengan kontras rendah yang sulit diidentifikasi secara visual.

Sehingga pada penelitian ini, dilakukan pengembangan model dengan menggabungkan ELM sebagai model klasifikasi, GLCM untuk ekstraksi fitur, dan *Histogram Equalization* untuk peningkatan kualitas citra, bertujuan untuk menghasilkan sistem deteksi kanker paru yang lebih akurat dan efisien dibandingkan penelitian sebelumnya.

METODE

Metode penelitian merupakan kerangka kerja yang sistematis dan terstruktur untuk mencapai tujuan dari sebuah studi atau eksperimen (Borman et al., 2021). Pada penelitian ini, metode yang digunakan bertujuan untuk menghasilkan model klasifikasi yang efektif dalam mendeteksi kanker paru pada citra *CT scan*. Tahapan yang diterapkan dalam penelitian ini divisualisasikan pada Gambar 1.



Gambar 1. Tahapan Penelitian

3.1. Mengumpulkan Data

Tahap ini dimulai dengan mengumpulkan dataset yang akan diterapkan dalam melakukan klasifikasi citra. Dalam studi ini, data yang dianalisis bersumber dari Chest CT-Scan Images Dataset (Hany, 2020), sebuah kumpulan data terbuka yang dapat diakses melalui *platform* Kaggle. Dataset ini berisi gambar *CT scan* paru-paru yang akan dianalisis untuk mendeteksi kanker paru. Dataset ini dibagi menjadi data untuk melatih model dan data untuk menguji model dengan rasio 70%:30%. Karena jumlah data yang terbatas, sebagian

besar digunakan untuk pelatihan guna membentuk pola pembelajaran yang lebih akurat dan relevan (Kurniabudi et al., 2020). Total 600 gambar digunakan, dengan 420 gambar untuk pelatihan dan 180 untuk pengujian.

3.2. Meningkatkan Kualitas Citra

Pada tahap ini, citra yang telah dikumpulkan dari dataset melalui proses pra-pemrosesan untuk memperbaiki kualitas visual serta detail penting pada citra. *Histogram Equalization* berguna dalam meningkatkan kontras citra dengan menyebarkan intensitas piksel secara merata pada spektrum intensitas, sehingga perbedaan antara jaringan normal dan jaringan yang terkena kanker menjadi lebih jelas (Winarno et al., 2022). Teknik ini membantu mengatasi variasi pencahayaan dan membuat struktur internal paru-paru lebih terlihat, yang penting untuk segmentasi dan ekstraksi fitur.

Histogram Equalization bekerja dengan menyebarkan distribusi intensitas piksel secara lebih merata pada rentang yang tersedia, sehingga detail dalam citra menjadi lebih terlihat dan kontras antara objek utama dan latar belakang meningkat (Manalu, 2021). Secara matematis, *Histogram Equalization* melibatkan perhitungan histogram kumulatif untuk memetakan nilai intensitas piksel asli ke nilai intensitas yang baru. Untuk mendapatkan nilai tersebut maka digunakan persamaan (1).

$$s_k = \text{floor} \left(\frac{(L-1)}{MN} \sum_{j=0}^k n_j \right) \quad (1)$$

di mana s_k merujuk pada intensitas piksel baru setelah pemetaan, L merupakan jumlah tingkat intensitas yang tersedia dalam citra, M dan N adalah jumlah piksel pada dimensi horisontal dan vertikal dari citra, n_j menunjukkan jumlah piksel pada tingkat intensitas j , dan $\sum_{j=0}^k n_j$ adalah histogram kumulatif pada intensitas k .

3.3. Segmentasi Gambar

Setelah pra-pemrosesan, langkah berikutnya adalah segmentasi gambar, di mana area paru-paru dipisahkan dari bagian lain dalam citra. Tujuan utama dari segmentasi adalah untuk mengenali dan memisahkan area atau objek penting dari latar belakang pada sebuah gambar (Borman, Fernando, et al.,

2022). *Otsu Thresholding* merupakan metode segmentasi gambar yang berfungsi untuk secara otomatis menetapkan ambang batas optimal guna memisahkan objek latar depan dari latar belakang (Veriarinal & Gunawan, 2024). Segmentasi menghasilkan gambar biner yang memisahkan objek utama (*foreground*) dari *background* berdasarkan nilai pikselnya.

3.4. Ekstraksi Fitur Menggunakan GLCM

Setelah segmentasi, fitur tekstur dari area paru-paru yang tersegmentasi diekstraksi menggunakan *Gray Level Co-occurrence Matrix* (GLCM). GLCM adalah pendekatan yang digunakan untuk mendapatkan fitur tekstur yang mengidentifikasi pola spasial antara piksel dengan intensitas tertentu dalam suatu citra (Ahmad et al., 2022). Beberapa fitur utama yang diekstraksi dari GLCM meliputi kontras, energi, homogenitas, dan korelasi. Fitur-fitur ini membantu dalam mengidentifikasi pola tekstur yang dapat membedakan jaringan yang sehat dengan yang terkena kanker, sehingga menjadi representasi penting dalam proses klasifikasi.

3.5. Klasifikasi Citra Menggunakan ELM

Tahap berikutnya adalah klasifikasi gambar dengan menggunakan metode *Extreme Learning Machine* (ELM). ELM adalah pendekatan *artificial neural network* yang memiliki kecepatan pelatihan tinggi dan kemampuan generalisasi yang baik (Najiyah & Topiq, 2021). Dalam penelitian ini, ELM dilatih menggunakan fitur tekstur yang diekstraksi dari GLCM, dan model ini akan mempelajari pola-pola yang khas pada jaringan paru-paru yang sehat dan yang terkena kanker. Setelah pelatihan, ELM kemudian mengklasifikasikan gambar *CT scan* ke dalam kelas yang relevan, misalnya jaringan sehat atau jaringan yang terkena kanker.

Berbeda dengan jaringan saraf tradisional yang memerlukan pembaruan bobot secara iteratif, ELM menggunakan metode *feedforward* dengan satu lapisan tersembunyi yang diinisialisasi secara acak dan hanya bobot keluaran yang dihitung secara eksak (Rumandan et al., 2022). Ini membuat ELM sangat efisien dan cepat, karena hanya memerlukan satu kali pelatihan untuk mengoptimalkan bobot keluaran. Untuk menghitung bobot keluaran dari lapisan

tersembunyi ke lapisan keluaran digunakan persamaan (2).

$$H\beta = T \quad (2)$$

di mana H menunjukkan matriks keluaran dari lapisan tersembunyi, β merupakan bobot keluaran yang menghubungkan neuron-neuron di lapisan tersembunyi ke neuron-neuron di lapisan keluaran, dan T merujuk pada matriks target, yang mewakili nilai keluaran yang diharapkan berdasarkan data input.

Untuk mendapatkan nilai β yang merupakan bobot yang menghubungkan lapisan tersembunyi dengan lapisan keluaran, maka dapat menggunakan persamaan (3).

$$\beta = H^+T \quad (3)$$

di mana H^+ adalah *Moore-Penrose pseudo-inverse* dari matriks keluaran H , T adalah matriks target, dan β adalah bobot yang menghubungkan neuron-neuron di lapisan tersembunyi ke lapisan keluaran.

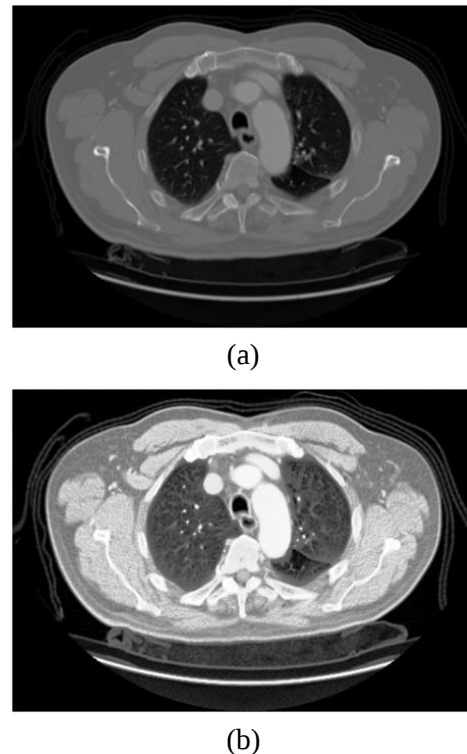
3.6. Mengevaluasi Model

Tahap terakhir adalah mengevaluasi performa model yang telah dibangun. Untuk mengevaluasi model digunakan metrik seperti akurasi, presisi, *recall*, dan *F1-Score* untuk memastikan bahwa model bekerja dengan baik dalam mengklasifikasikan citra kanker paru. Selain itu, *confusion matrix* dapat digunakan untuk melihat tingkat kesalahan klasifikasi yang terjadi (Borman, Rossi, et al., 2022). Dengan melakukan evaluasi model, penelitian ini dapat mengungkap kelebihan dan kekurangan dari pendekatan yang diterapkan, sehingga dapat memberikan saran perbaikan untuk penelitian di masa mendatang.

HASIL DAN PEMBAHASAN

Untuk melakukan klasifikasi penyakit paru-paru dari citra CT scan, langkah awal yang perlu dilakukan adalah mempersiapkan dataset. Dalam penelitian ini menggunakan *Chest CT-Scan Images Dataset*, yang merupakan dataset publik yang tersedia di situs web Kaggle (Hany, 2020). Dataset ini terdiri dari 600 citra, dengan 420 citra digunakan sebagai data pelatihan dan 180 citra sebagai data pengujian. Dataset tersebut mencakup empat kelas, yaitu: *Healthy Lungs*, *Adenocarcinoma*, *Squamous Cell Carcinoma*, dan *Large Cell Carcinoma*. Setelah

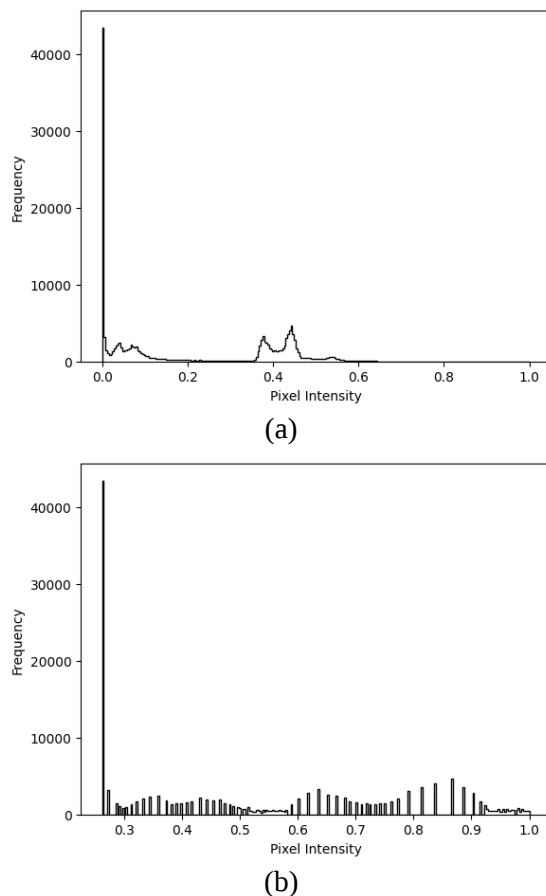
mengumpulkan citra *CT scan*, langkah berikutnya adalah pra-pemrosesan, di mana citra-citra tersebut diubah menjadi citra *grayscale* untuk mengurangi kompleksitas data dan memudahkan pemrosesan. Representasi visual dalam citra *grayscale* kemudian dilakukan peningkatan kualitas citra dengan teknik *Histogram Equalization*. *Histogram Equalization* diterapkan untuk meningkatkan kontras gambar dengan meratakan distribusi intensitas piksel pada spektrum intensitas, sehingga perbedaan antara jaringan normal dan jaringan yang terinfeksi kanker menjadi lebih terlihat jelas. Hasil peningkatan citra dengan menggunakan *Histogram Equalization* divisualisasikan pada Gambar 2.



Gambar 2. (a) Citra Asli; (b) Citra *Histogram Equalization*

Pada Gambar 2, citra asli (Gambar 2. (a)) menunjukkan kontras yang lebih rendah, sehingga beberapa detail pada jaringan paru-paru dan struktur sekitarnya kurang terlihat jelas. Setelah diterapkan *Histogram Equalization* (Gambar 2. (b)), kontras citra meningkat secara signifikan. Teknik ini mendistribusikan intensitas piksel secara lebih merata, sehingga fitur-fitur penting seperti jaringan paru-paru dan tulang menjadi lebih terlihat. Peningkatan kontras ini mempermudah

proses analisis dan interpretasi citra medis, sehingga membantu deteksi yang lebih akurat. Selanjutnya, untuk melihat perbedaan distribusi piksel antara citra asli dan citra *Histogram Equalization*, kedua citra tersebut ditampilkan dalam bentuk histogram, sebagaimana diperlihatkan pada Gambar 3.



Gambar 3. (a) Histogram Asli; (b) Citra *Histogram Equalization*

Pada Gambar 3 menunjukkan histogram dari dua gambar yaitu citra asli dan citra *Histogram Equalization*. Pada histogram citra asli (Gambar 3. (a)), distribusi intensitas tampak terkonsentrasi pada rentang intensitas yang sempit, dengan mayoritas piksel berada pada intensitas rendah. Hal ini menunjukkan bahwa citra memiliki kontras yang terbatas. Ini menyebabkan detail penting mungkin tidak terlalu terlihat. Setelah diterapkan *Histogram Equalization* (Gambar 3. (b)), distribusi intensitas piksel menjadi lebih merata di seluruh rentang intensitas. Hal ini menunjukkan bahwa kontras citra telah meningkat, sehingga

detail pada citra lebih mudah dikenali. Dengan peningkatan distribusi intensitas ini, area yang awalnya kurang terlihat menjadi lebih jelas dan lebih mudah dianalisis.

Proses selanjutnya adalah segmentasi citra, yang bertujuan untuk memisahkan objek utama dari latar belakang dalam citra. Segmentasi ini penting dalam pengolahan citra medis, terutama untuk mengidentifikasi area yang relevan seperti jaringan yang terkena penyakit. Dalam penelitian ini, metode *Otsu Thresholding* digunakan untuk menentukan ambang batas optimal yang secara otomatis memisahkan objek latar depan dari latar belakang berdasarkan distribusi intensitas piksel. Teknik ini bekerja dengan mencari nilai ambang yang meminimalkan variansi intra-kelas dari piksel latar depan dan latar belakang, sehingga menghasilkan citra biner yang menyoroti objek yang penting untuk dipelajari lebih lanjut. Hasil citra yang tersegmentasi diperlihatkan pada Gambar 4.



Gambar 4. Citra Hasil Segmentasi

Gambar 4 adalah visualisasi dari proses segmentasi citra yang dilakukan menggunakan metode *Otsu Thresholding*. Pada gambar tersebut, objek latar depan, seperti paru-paru dan beberapa struktur internal, telah dipisahkan dari latar belakang. Warna putih mewakili latar belakang, sementara objek utama, seperti jaringan paru-paru, diwakili oleh warna hitam.

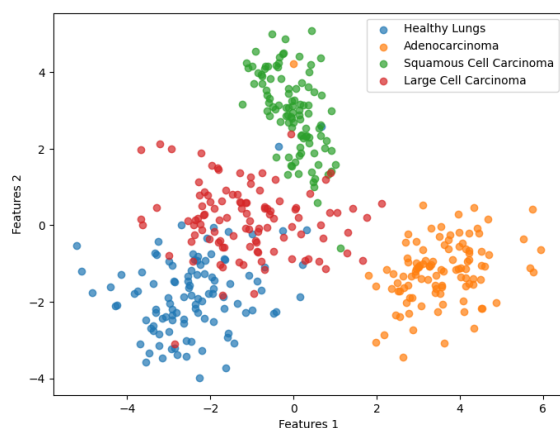
Setelah tahap segmentasi citra, nilai fitur tekstur yang terkandung dalam gambar diekstraksi menggunakan pendekatan *Gray Level Co-occurrence Matrix (GLCM)*. GLCM adalah matriks dua dimensi yang menghitung frekuensi pasangan piksel dengan nilai intensitas tertentu yang terletak bersebelahan dalam gambar. Dengan menggunakan GLCM, fitur-fitur tekstur penting seperti kontras,

energi, homogenitas, dan korelasi dapat diidentifikasi. Sampel nilai GLCM ditampilkan pada Tabel 1.

Tabel 1. Sampel Nilai GLCM

Citra	Paramter	Nilai
	Kontras	723.1433
	Energi	0.7039
	Homogenitas	0.9889
	Korelasi	0.9774

Tabel 1 menampilkan sampel nilai GLCM (Gray Level Co-occurrence Matrix) untuk citra yang telah disegmentasi. Informasi mengenai tekstur yang dihasilkan oleh GLCM kemudian menjadi input bagi algoritma *Extreme Learning Machine* (ELM) untuk melakukan klasifikasi. Fase berikutnya melibatkan pembangunan model pembelajaran mesin menggunakan jaringan saraf tiruan dengan pendekatan *Extreme Learning Machine* (ELM). ELM bekerja dengan menggunakan satu lapisan tersembunyi yang bobotnya diinisialisasi secara acak, sementara bobot keluaran dihitung secara eksak tanpa proses iteratif. Model ELM dilatih dengan 420 gambar untuk mengenali pola-pola tekstur yang dihasilkan dari GLCM, dengan setiap neuron tersembunyi melalui proses sekali iterasi untuk mempelajari pola. Visualisasi yang ditampilkan menunjukkan distribusi sampel data pelatihan, sebagaimana diilustrasikan pada Gambar 5.

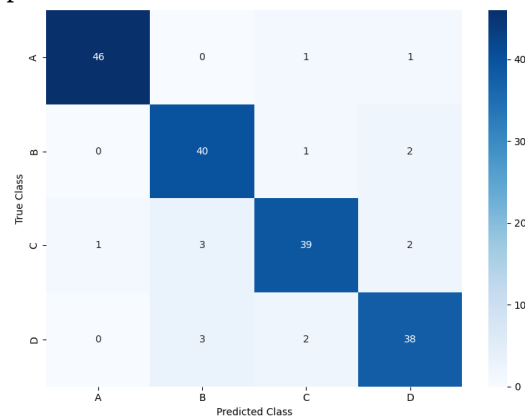


Gambar 5. Visualisasi Distribusi Data Pelatihan

Gambar 5 merupakan visualisasi distribusi data pelatihan untuk empat kelas (*Healthy Lungs*, *Adenocarcinoma*, *Squamous Cell Carcinoma*, dan *Large Cell Carcinoma*) yang digunakan dalam algoritma *Extreme Learning Machine* (ELM). Setiap kelas terlihat cukup terpisah dari yang lain, yang

menunjukkan bahwa model ELM memiliki peluang untuk mengenali dan mengklasifikasikan setiap kelas dengan baik.

Setelah model pelatihan dikembangkan, langkah selanjutnya adalah melakukan evaluasi kinerja model menggunakan *confusion matrix*. *Confusion matrix* digunakan untuk mendapatkan nilai-nilai penting seperti akurasi, presisi, *recall*, dan *F1-Score*, yang memberikan gambaran tentang seberapa baik model dalam melakukan klasifikasi. Evaluasi dilakukan menggunakan 180 gambar uji yang dibagi menjadi empat kategori: *Healthy Lungs*, *Adenocarcinoma*, *Squamous Cell Carcinoma*, dan *Large Cell Carcinoma*. Penilaian model dilakukan dengan menghitung prediksi yang benar dan salah untuk masing-masing kategori, seperti yang terlihat dalam *confusion matrix* pada Gambar 6.



A: Healthy Lungs, B: Adenocarcinoma, C: Squamous Cell Carcinoma, D: Large Cell Carcinoma

Gambar 6. Hasil *Confusion Matrix*

Pada Gambar 6 menunjukkan *confusion matrix* yang berguna untuk mengidentifikasi kelas mana yang paling sulit diprediksi oleh model, serta sejauh mana model mampu membedakan setiap kategori dengan baik. Berdasarkan *confusion matrix* tersebut kemudian diperoleh nilai presisi, *recall*, *F1-Score*, dan akurasi seperti yang ditunjukkan pada Tabel 2.

Tabel 2. Hasil Evaluasi Model

Kelas	Precision	Recall	F1-Score	Accuracy
A	0,9787	0,9583	0,9684	0,9106
B	0,8696	0,9302	0,8989	
C	0,9070	0,8667	0,8864	
D	0,8837	0,8837	0,8837	

A: *Healthy Lungs*, B: *Adenocarcinoma*, C: *Squamous Cell Carcinoma*, D: *Large Cell Carcinoma*

Tabel 2 menunjukkan hasil evaluasi model klasifikasi menggunakan tiga metrik utama: *precision*, *recall*, dan *F1-Score*, serta *accuracy* secara keseluruhan. Kelas A (*Healthy Lungs*) memiliki kinerja terbaik dengan *precision*, *recall*, dan *F1-Score* tertinggi, yang menunjukkan kemampuan model yang sangat baik dalam mengklasifikasikan citra paru-paru sehat. Sebaliknya, kelas B (*Adenocarcinoma*) dan D (*Large Cell Carcinoma*) memiliki *precision* dan *F1-Score* yang lebih rendah, mengindikasikan bahwa model lebih sulit dalam mengklasifikasikan kedua kelas ini dengan akurat. Secara keseluruhan, akurasi model mencapai 0.9106 atau 91.06%, yang menunjukkan bahwa model ini berkinerja baik. Selanjutnya model yang dibangun dibandingkan dengan hasil akurasi dengan algoritma ELM tanpa melalui proses *Histogram Equalization*. Hasil perbandingan tersebut dapat dilihat pada Tabel 3.

Tabel 3. Perbandingan Hasil Akurasi

Metode	Accuracy
ELM	0,8444
ELM + <i>Histogram Equalization</i>	0,9106

Tabel 3 menunjukkan perbandingan akurasi antara model ELM tanpa dan dengan *Histogram Equalization*. Model ELM sendiri menghasilkan akurasi sebesar 0,8444 atau 84,44%, sedangkan setelah ditambahkan proses *Histogram Equalization*, akurasinya meningkat menjadi 0,9106 atau 91,06%. Peningkatan ini menunjukkan bahwa teknik *Histogram Equalization* berhasil meningkatkan kualitas citra dengan memperbaiki kontras, sehingga membantu model dalam mengenali pola yang lebih baik dan menghasilkan prediksi yang lebih akurat. Hasil ini mengindikasikan bahwa metode yang diusulkan, yakni menggabungkan ELM dengan *Histogram Equalization*, memberikan peningkatan kinerja yang signifikan dalam klasifikasi citra medis, khususnya dalam mendeteksi penyakit paru-paru pada citra *CT scan*.

Meskipun akurasi model telah mencapai 91.06%, masih ada ruang untuk perbaikan, terutama dalam mengklasifikasikan *Adenocarcinoma* dan *Large Cell Carcinoma*, yang memiliki *precision* dan *recall* lebih rendah dibandingkan kelas lainnya. Salah satu

pendekatan yang dapat diambil untuk meningkatkan kinerja adalah dengan menambahkan lebih banyak data pelatihan khusus untuk kelas-kelas yang sulit diklasifikasikan, sehingga model dapat belajar lebih baik tentang pola-pola yang membedakan antara kelas-kelas ini. Selain itu, dapat dipertimbangkan untuk menggunakan teknik augmentasi data, seperti rotasi atau pergeseran kontras, untuk memperluas variasi data dan membantu model lebih robust dalam mengenali fitur khas dari masing-masing kategori. Terakhir, tuning *hyperparameter* lebih lanjut, seperti jumlah neuron tersembunyi di ELM, atau optimisasi lebih lanjut pada teknik *Histogram Equalization*, dapat membantu dalam meningkatkan kemampuan klasifikasi model.

SIMPULAN

Penelitian ini telah berhasil membangun model untuk mengklasifikasikan penyakit paru-paru dari citra *CT scan* menggunakan algoritma *Extreme Learning Machine* (ELM) dan ekstraksi fitur GLCM yang dikombinasikan dengan *Histogram Equalization*. Pendekatan ELM memiliki kemampuan pelatihan yang cepat dan generalisasi yang baik, sehingga cocok untuk menangani data citra medis yang kompleks. Model yang diusulkan terbukti efektif dalam meningkatkan kinerja klasifikasi, dengan akurasi mencapai 91.06%. Penggunaan *Histogram Equalization* berkontribusi signifikan terhadap peningkatan akurasi karena kemampuannya dalam meningkatkan kontras citra, sehingga memudahkan proses ekstraksi fitur tekstur yang lebih representatif.

Untuk penelitian selanjutnya dapat mempertimbangkan teknik augmentasi data untuk memperluas variasi data dan membantu model menjadi lebih tangguh dalam mengenali fitur khas dari setiap kategori penyakit. Selain itu, tuning lebih lanjut pada *hyperparameter* seperti jumlah neuron tersembunyi di ELM, atau penggunaan arsitektur jaringan saraf tiruan yang lebih kompleks seperti *Convolutional Neural Networks* (CNN), juga dapat dieksplorasi untuk mendapatkan model yang lebih optimal dan meningkatkan akurasi klasifikasi.

DAFTAR PUSTAKA

- Ahmad, I., Rahmanto, Y., Borman, R. I., Rossi, F., Jusman, Y., & Alexander, A. D. (2022). Identification of Pineapple Disease Based on Image Using Neural Network Self-Organizing Map (SOM) Model. *International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)*, 12–17.
- Apasari, R., Aditya, Y. N., Purwanti, E., & Arof, H. (2020). Development of lung cancer classification system for computed tomography images using artificial neural network. *International Conference on Mathematics, Computational Sciences and Statistics*, 1–9.
- Borman, R. I., Ahmad, I., & Rahmanto, Y. (2022). Klasifikasi Citra Tanaman Perdu Liar Berkhasiat Obat Menggunakan Jaringan Syaraf Tiruan Radial Basis Function. *Bulletin of Informatics and Data Science*, 1(1), 6–13.
- Borman, R. I., Fernando, Y., & Yudoutomo, Y. E. P. (2022). Identification of Vehicle Types Using Learning Vector Quantization Algorithm with Morphological Features. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(2), 339–345.
- Borman, R. I., Napianto, R., Nugroho, N., Pasha, D., Rahmanto, Y., & Yudoutomo, Y. E. P. (2021). Implementation of PCA and KNN Algorithms in the Classification of Indonesian Medicinal Plants. *International Conference on Computer Science, Information Technology and Electrical Engineering (ICOMITEE)*, 46–50.
- Borman, R. I., Rossi, F., Alamsyah, D., Nuraini, R., & Jusman, Y. (2022). Classification of Medicinal Wild Plants Using Radial Basis Function Neural Network with Least Mean Square. *International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)*.
- Chen, Y. (2020). Identification of Tea Leaf Based on Histogram Equalization, Gray-Level Co-Occurrence Matrix and Support Vector Machine Algorithm BT - Multimedia Technology and Enhanced Learning. *Multimedia Technology and Enhanced Learning*, 3–16.
- Fendriani, Y., Kharisma, R., & Peslinof, M. (2023). Analisis Perbandingan Variasi Filter Pada Deteksi Tepi Menggunakan Metode Canny Terhadap Citra CT-Scan Kanker Paru-Paru. *Journal Online of Physics*, 8(2), 77–81.
- Firdaus, Q., Sigit, R., Harsono, T., & Anwar, A. (2020). Lung Cancer Detection Based on CT-Scan Images with Detection Features Using Gray Level Co-Occurrence Matrix (GLCM) and Support Vector Machine (SVM) Methods. *International Electronics Symposium (IES)*, 643–648.
- Hany, M. (2020). *Chest CT-Scan images Dataset*. Kaggle.
- Hizriadi, A., Purnamawati, S., & Angreni, F. (2023). Implementation of Extreme Learning Machine for Classification of Retina Ablasio Results on Retina Fundus Images. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 3(4), 371–376.
- Hossain, M. S., Basak, N., Mollah, M. A., Nahiduzzaman, M., Ahsan, M., & Haider, J. (2025). Ensemble-based multiclass lung cancer classification using hybrid CNN-SVD feature extraction and selection method. *PLoS ONE*, 20(3), 1–30. <https://doi.org/10.1371/journal.pone.0318219>
- Ihsan, M. F., & Sunyoto, A. (2022). Gray Level Co-Occurrence Matrix Algorithm and Backpropagation Neural Networks for Herbal Plants Identification. *International Conference on Information and Communications Technology (ICOIACT)*, 373–378.
- Kurnia, H., & Hidayat, T. (2023). Penajaman Kualitas Citra Digital Menggunakan Histogram Equalization. *Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD*, 6(9), 1–7.
- Kurniabudi, Stiawan, D., Darmawijoyo, Bin Idris, M. Y., Bamhdi, A. M., & Budiarto, R. (2020). CICIDS-2017 Dataset Feature Analysis With Information Gain for Anomaly Detection. *IEEE Access*, 8, 132911–132921.
- Manalu, R. E. (2021). Analisis Metode Histogram Equalization Dalam Proses Perbaikan Gambar Closed Circuit Television (CCTV). *TIN: Terapan Informatika Nusantara*, 2(1), 1–5.
- Najiyah, I., & Topiq, S. (2021). Klasifikasi Jenis Kendaraan Roda Empat Menggunakan Extreme Learning

- Machine. *Jurnal Responsif*, 3(2), 199–206.
- Nugroho, B., Utami, A. S., Ratnaningsih, D. K., & Handis, L. A. (2023). Perbandingan Gambaran CT Scan Paru Perokok dan Non Perokok Pasien Kanker Paru. *Jurnal Kesehatan*, 2, 1–22.
- Putri, E. P., Nurhasanah, N., Wahyuni, D., Hasanuddin, H., Adriat, R., & Arsyad, Y. M. (2024). Klasifikasi Keganasan Kanker Paru Menggunakan Algoritma Propagasi Balik pada Citra CT-Scan. *Jurnal Ilmu Dasar*, 25(2), 111–118.
- Rumandan, R. J., Nuraini, R., Sadikin, N., & Rahmanto, Y. (2022). Klasifikasi Citra Jenis Daun Berkhasiat Obat Menggunakan Algoritma Jaringan Syaraf Tiruan Extreme Learning Machine. *Journal of Computer System and Informatics (JoSYC)*, 4(1), 145–154. <https://doi.org/10.47065/josyc.v4i1.2586>
- Rustam, Z., Purwanto, A., Hartini, S., & Saragih, G. S. (2021). Lung Cancer Classification Using Fuzzy C-Means and Fuzzy Kernel C-Means Based on CT Scan Image. *International Journal of Artificial Intelligence (IJ-AI)*, 10(2), 11591.
- Sugiharto, S., Simanjuntak, R. A. P. S., & Larissa, O. (2021). Kanker Paru, Faktor Risiko dan Pencegahannya. *Seminar Nasional Hasil Penelitian Dan Pengabdian Kepada Masyarakat*, 613–620.
- Ulga, S., Muttaqin, A., & Fitriyani, D. (2024). Pengaruh Filter Lowpass Terhadap Kualitas Citra CT-Scan Paru-Paru. *Jurnal Fisika Unand (JFU)*, 13(4), 587–593.
- Veriarinal, V., & Gunawan, C. (2024). Identifikasi Nilai Mata Uang Logam Menggunakan Metode Otsu Thresholding. *Kohesi: Jurnal Multidisiplin Saintek*, 2(11), 1–9.
- Wahid, R. R., Anggraeni, F. T., & Nugroho, B. (2023). Implementasi Metode Extreme Learning Machine untuk Klasifikasi Tumor Otak pada Citra Magnetic Resonance Imaging. *Seminar Nasional Informatika Bela Negara (SANTIKA)*, 16–20. <https://doi.org/10.33005/santika.v1i0.45>
- Winarno, G., Irsal, M., Karenina, C. A., Sari, G., & Hidayati, R. N. (2022). Metode Histogram Equalization untuk Peningkatan Kualitas Citra dengan Menggunakan Studi Phantom Lumbosacral. *Jurnal Kesehatan Vokasional*, 7(2), 104–110.
- Yel, M. B., Mulyana, D. I., & Hidayat, W. (2023). Klasifikasi Jenis Ikan Neon Dengan Ekstraksi Fitur GLCM dan Algoritma Extreme Learning Machine. *Jurnal Ilmiah Teknologi Informasi Dan Komunikasi (JTIK)*, 14(2), 228–238.