

Deteksi Pneumonia dengan Ekstraksi Fitur Gray-Level Co-occurrence Matrix (GLCM) dan Support Vector Machine (SVM)

I Gusti Bagus Sutha Arianata Putra^{a1}, Gst. Ayu Vida Matrika Giri^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
¹bagussuthaap@gmail.com
²vida@unud.ac.id

Abstract

Pneumonia, a prevalent lung disease globally, poses significant challenges in accurate diagnosis despite its severity. This paper proposes a novel approach leveraging Support Vector Machine (SVM) classification and Gray-Level Co-occurrence Matrix (GLCM) analysis on chest X-ray images to aid in pneumonia diagnosis. By extracting pneumonia-indicative features from digital X-ray images using Gray-Level Co-occurrence Matrix (GLCM) and employing Support Vector Machine (SVM) for classification, the study aims to enhance pneumonia diagnosis effectiveness, particularly crucial in regions with limited healthcare resources. The proposed method focuses on identifying characteristic patterns indicative of pneumonia in chest X-ray images and distinguishing between normal and pneumonia-affected images based on GLCM-extracted features. Furthermore, the study evaluates the impact of hyperparameter tuning using grid search on the proposed diagnostic system's performance, including accuracy, sensitivity, and specificity. By achieving these objectives, the research aims to contribute significantly to the development of more accurate and effective diagnostic tools for pneumonia, especially in resource-constrained areas.

Keywords: Gray-Level Co-occurrence Matrix (GLCM), Machine Learning, Pneumonia, Support Vector Machine, X-Ray

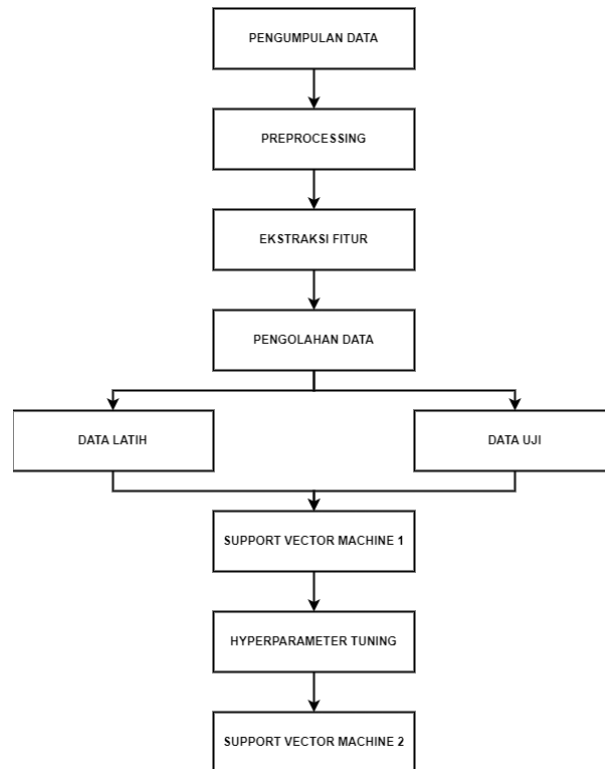
1. Pendahuluan

Penyakit paru-paru banyak terjadi di seluruh dunia, diantaranya berkaitan dengan penyakit paru obstruktif kronis, asma, tuberkulosis, fibrosis, dan pneumonia [1]. Pneumonia, atau dikenal juga dengan istilah paru-paru basah, adalah infeksi yang mengakibatkan peradangan pada kantong-kantong udara di salah satu atau kedua paru-paru. Biasanya pneumonia ini ditandai dengan peningkatan opasitas pada area paru-paru [2]. Pada pasien penderita pneumonia, terjadi peradangan dan pengisian cairan atau nanah pada sekelompok kecil kantong udara di ujung saluran pernapasan paru-paru, yang disebut alveoli. Akibatnya, para penderita pneumonia mengalami gejala batuk berdahak, demam, menggigil, nyeri dada, dan kesulitan bernapas. Bakteri, virus, dan jamur merupakan organisme yang dapat menyebabkan pneumonia atau paru-paru basah. Namun, pada penderita dewasa, kondisi ini lebih sering diakibatkan karena infeksi bakteri, meskipun kondisi tersebut juga dapat disebabkan oleh virus maupun jamur. Pneumonia dapat dikatakan salah satu penyebab kematian pada anak tertinggi di dunia. World Health Organization (WHO) memperkirakan bahwa penyakit ini telah menjadi pemicu 15% kematian anak-anak berusia di bawah 5 tahun [4]. Menurut Ourworldindata. Di Indonesia sendiri, data menunjukkan bahwa pada tahun 2019, estimasi tingkat kematian tahunan akibat pneumonia dan infeksi saluran pernapasan bawah lainnya adalah 28.43 per 100,000 penduduk dan estimasi tingkat kematian tahunan akibat pneumonia dan infeksi saluran pernapasan bawah lainnya untuk anak dibawah umur 5 tahun adalah 45.8 per 100,000 penduduk [3]. Studi ini bertujuan untuk mengusulkan dan mengevaluasi pendekatan dalam diagnosis pneumonia melalui analisis citra digital menggunakan metode Gray-Level Co-occurrence Matrix (GLCM) dan Support Vector

Machine (SVM). Tujuan utama adalah mengembangkan metode ekstraksi fitur berbasis GLCM untuk mengidentifikasi pola karakteristik pada citra rontgen dada yang mengindikasikan adanya pneumonia, dan menerapkan algoritma SVM sebagai classifier untuk membedakan antara citra rontgen normal dan citra rontgen yang menunjukkan tanda-tanda pneumonia berdasarkan fitur yang diekstraksi dari GLCM. Selain itu, studi ini juga bertujuan untuk mengukur pengaruh penyetalan *hyperparameter* menggunakan grid search dalam kinerja sistem diagnosis yang diusulkan dalam mendeteksi pneumonia, termasuk akurasi, sensitivitas, dan spesifisitasnya. Dengan mencapai tujuan ini, diharapkan penelitian ini dapat memberikan kontribusi signifikan dalam pengembangan alat bantu diagnosis yang lebih akurat dan efektif untuk pneumonia, terutama di daerah-daerah dengan sumber daya kesehatan yang terbatas. Beberapa penelitian sebelumnya telah mengusulkan metode-metode berbeda untuk diagnosis pneumonia. Beberapa menggunakan teknologi citra seperti pemrosesan citra digital, machine learning, dan deep learning. Metode-metode ini sering kali memanfaatkan fitur-fitur yang diekstraksi dari citra rontgen dada untuk membedakan antara citra rontgen normal dan yang menunjukkan tanda-tanda pneumonia. Pada beberapa penelitian menggunakan deep learning seperti cnn, telah menunjukkan bahwa deep learning telah sangat mampu dalam mendeteksi pneumonia dimana terutama metode CNN-ELM, yang dimana akurasinya mencapai 93,59% dan F1-Score mencapai nilai 0,9474 [2]. Penggunaan metode GLCM untuk ekstraksi fitur dan SVM sebagai classifier dalam penelitian ini dapat berpotensi memberikan kontribusi yang signifikan dalam memberikan perspektif baru dalam deteksi pneumonia menggunakan machine learning, meningkatkan akurasi dan efektivitas diagnosis pneumonia. Urgensi penelitian ini tak dapat diabaikan. Diagnosis yang cepat dan akurat adalah kunci untuk menangani pneumonia secara efektif. Keterlambatan dalam diagnosis bisa mengakibatkan risiko komplikasi serius, seperti gagal napas atau sepsis, yang berpotensi fatal. Oleh karena itu, pengembangan metode yang dapat mempercepat dan memperbaiki proses diagnosis merupakan prioritas utama dalam upaya penyelamatan nyawa. Penerapan teknologi ini diharapkan dapat menaikkan produktivitas radiolog dengan bantuan algoritma komputer dalam mendeteksi tanda-tanda pneumonia pada citra rontgen. Hal ini memungkinkan radiolog untuk lebih fokus pada kasus-kasus yang membutuhkan perhatian khusus, mengurangi beban kerja mereka dan meningkatkan efisiensi diagnosis secara keseluruhan. Dengan demikian, penelitian ini tidak hanya bertujuan untuk meningkatkan akurasi diagnosis pneumonia, tetapi juga memiliki implikasi yang luas dalam meningkatkan perawatan kesehatan masyarakat, terutama di daerah-daerah dengan sumber daya terbatas.

2. Metode Penelitian

Tahapan-tahapan penelitian yang ditampilkan dalam Gambar 1 yang menjelaskan langkah-langkah secara terperinci, mulai dari tahap awal persiapan data hingga tahap akhir evaluasi hasil. Setiap tahapan memiliki peran krusial dalam memastikan bahwa penelitian berjalan dengan validitas dan efektivitas yang maksimal. Mulai dari mengumpulkan dan mempersiapkan data, melakukan preprocessing untuk mempersiapkan data agar dapat digunakan dalam analisis, mengekstraksi fitur penting dari data, mengolah data dengan membagi dataset dan menjalankan algoritma machine learning, hingga melakukan evaluasi untuk mengukur kinerja model yang telah dikembangkan. Keseluruhan proses penelitian ini mengikuti rangkaian-rangkaian tahapan tersebut dengan tujuan sehingga mencapai pemahaman yang lebih mendalam tentang dataset yang digunakan serta memastikan performa yang optimal dari model yang dikembangkan.



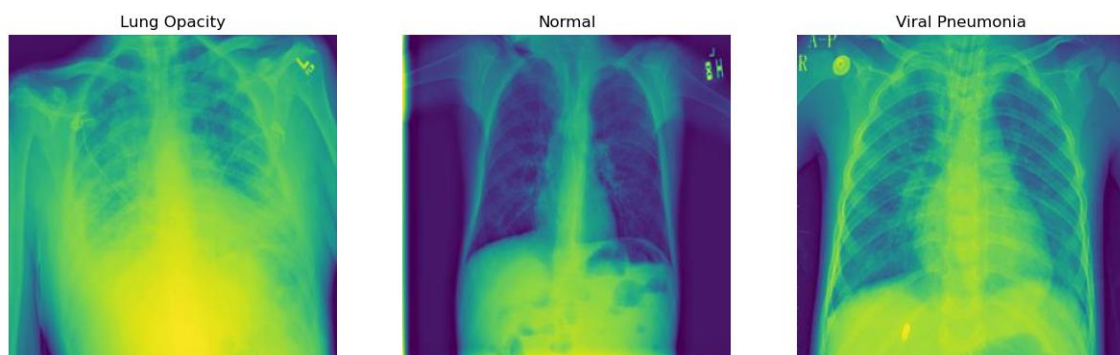
Gambar 1. Tahapan Penelitian

2.1. Persiapan Data

Data yang digunakan dalam penelitian ini diperoleh dari platform Kaggle dan dibuat oleh Fatemeh Mehrparvar dengan ukuran 37 mb yang dapat diakses pada link berikut:

<https://www.kaggle.com/datasets/fatemehmehrparvar/lung-disease> [10]

Pada dataset tersebut terdapat 3475 file jpg dibagi menjadi 3 klasifikasi citra yang berbeda. Berikut adalah contoh gambar dataset pada gambar 2.



Gambar 2. Dataset

2.2. Preprocessing

a. Data Balancing

Menyeimbangkan data dari setiap kelas adalah langkah penting untuk memastikan distribusi yang merata dalam dataset. Dalam kasus ini, total terdapat 3475 file JPG dengan

pembagian yang tidak merata antar kelas. Oleh karena itu, pendekatan yang diambil adalah dengan menghapus sebagian data agar setiap kelas memiliki jumlah sampel yang seimbang, yaitu sekitar 1000 file JPG per kelas.

b. Normalisasi Label

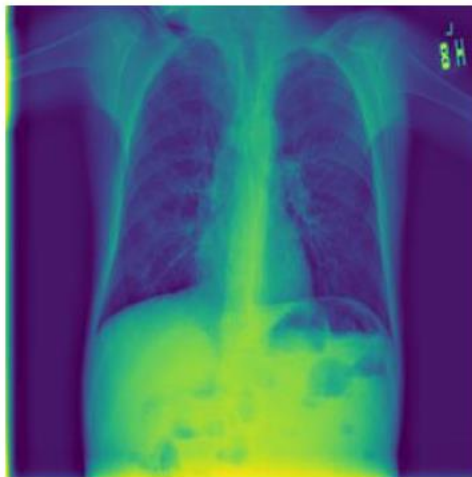
Normalisasi label digunakan untuk mengubah label dari format yang diberikan dalam nama folder menjadi format yang lebih sederhana dan standar. Proses ini menghapus spasi, tanda kurung, dan mengambil dua kata pertama dari label untuk menghasilkan label yang dinormalisasi.

c. Normalisasi Deskripsi

Normalisasi deskripsi digunakan untuk menghasilkan deskripsi yang dinormalisasi dari folder dan sub-folder. Proses ini menghapus angka, tanda titik, dan spasi berlebih, serta menggabungkan nama folder dan subfolder menjadi satu string.

d. CLAHE (Contrast Limited Adaptive Histogram Equalization)

CLAHE (Contrast Limited Adaptive Histogram Equalization) digunakan untuk meningkatkan kontras gambar dengan memperhatikan distribusi histogram lokal, sehingga membuat detail gambar yang rendah lebih jelas.



Original Image

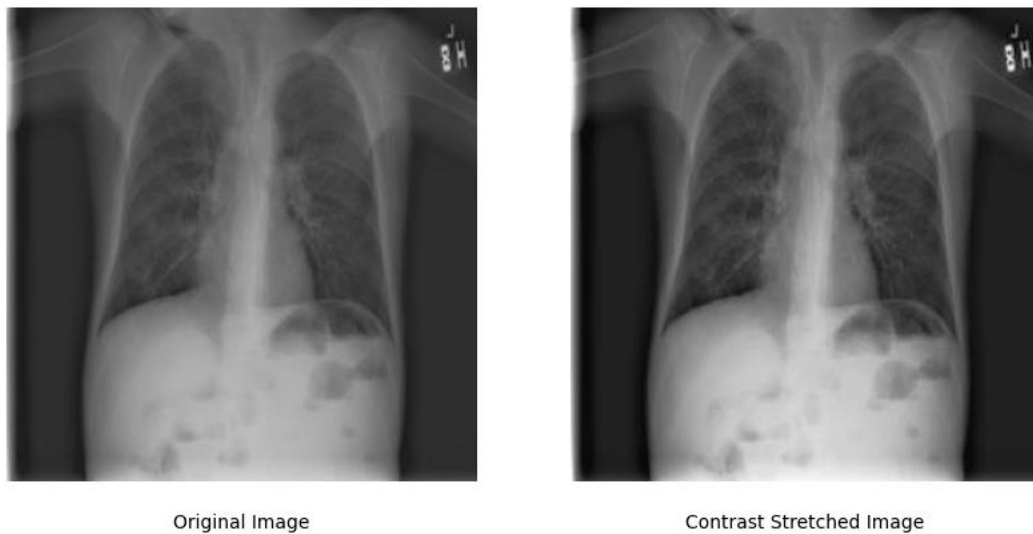


CLAHE Image

Gambar 3. Perbandingan Gambar Asli dan Contrast Stretched Image

e. Kontrast Stretching

Kontrast stretching adalah teknik yang digunakan untuk meningkatkan kontras gambar dengan meregangkan rentang intensitas piksel sehingga seluruh rentang intensitas dapat digunakan secara optimal.



Gambar 4. Perbandingan Gambar Asli dan Contrast Stretched Image

2.3. Ekstraksi Fitur

Ekstraksi fitur pada penelitian ini akan menggunakan pendekatan GLCM untuk melakukan analisis citra dari sudut pandang tekstur pada sebuah citra digital. Metode GLCM menggunakan hasil kalkulasi dari matriks co-occurrence yang didapatkan melalui proses menghitung hubungan antara nilai-nilai piksel yang saling berdekatan dalam gambar, dengan mempertimbangkan dua parameter: sudut (angle) piksel citra dan jarak ketetanggaan antar pixel. Dalam metode GLCM, nilai sudut yang umum digunakan adalah 0, 45, 90, dan 135 derajat, sedangkan untuk jarak ketetanggaan antar pixel dapat menggunakan nilai 1, 2, 3, dan seterusnya [5]. Setelah matriks GLCM terbentuk, nilai-nilai fitur statistik seperti *contrast*, *dissimilarity*, *homogeneity*, *ASM*, *energy* dan *correlation* dapat dihitung berdasarkan matriks ini. Berikut merupakan persamaan yang digunakan untuk menghasilkan nilai-nilai statistik pada metode GLCM [6].

$$a. \text{ Contrast} = \sum_{i,j=0}^{levels-1} p_{i,j} (i - j)^2 \quad (1)$$

$$b. \text{ Dissimilarity} = \sum_{i,j=0}^{levels-1} p_{i,j} |i - j| \quad (2)$$

$$c. \text{ Homogeneity} = \sum_{i,j=0}^{levels-1} \frac{p_{i,j}}{1+(i-j)^2} \quad (3)$$

$$d. \text{ ASM} = \sum_{i,j=0}^{levels-1} p_{i,j}^2 \quad (4)$$

$$e. \text{ Energy} = \sqrt{ASM} \quad (5)$$

$$f. \text{ Correlation} = \sum_{i,j=0}^{levels-1} p_{i,j} \left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right] \quad (6)$$

2.4. Pengolahan Data

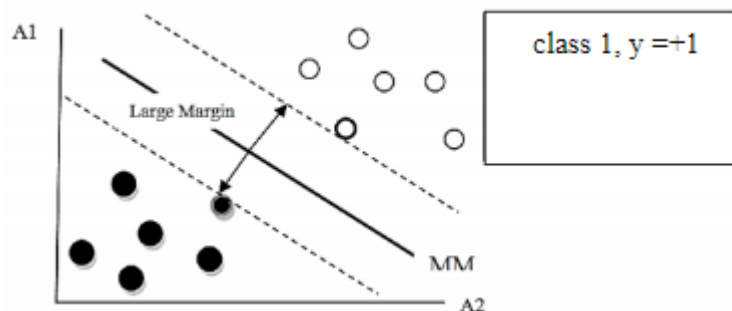
Pengolahan data dilakukan dengan membagi dataset yang telah melewati tahap sebelumnya menjadi 2 (dua) yaitu data uji dan data dengan perbandingan 80% untuk data latih dan 20% untuk data uji. Setelah pembagian, data-data tersebut lalu dijalankan pada algoritma Support Vector Machine.

2.5. Grid Search

Grid search merupakan pencarian yang menyeluruh berdasarkan subset yang telah ditentukan dari ruang hiper-parameter. Hiper-parameter tersebut ditentukan dengan menggunakan nilai minimal (batas bawah), nilai maksimal (batas atas), dan jumlah langkah [8]. Grid Search akan membagi rentang parameter yang akan dioptimalkan menjadi sebuah grid dan melintasi semua titik untuk mendapatkan parameter optimal. Grid Search mengoptimalkan parameter SVM menggunakan teknik validasi silang sebagai metrik kinerja. Menurut Lin et al., menerapkan teknik validasi silang mampu mencegah overfitting [9]. Grid search bertujuan untuk mengidentifikasi kombinasi *hyperparameter* terbaik sehingga *klasifier* mampu memprediksi data-data yang tidak diketahui dengan benar. Sebelum melakukan penyetelan *hyperparameter* menggunakan Grid Search, model dilatih pada set data latih menggunakan algoritma Support Vector Machine (SVM). Prediksi kemudian dibuat dengan menggunakan set data uji. Setelah itu, untuk penyetelan *hyperparameter*, pendekatan Grid Search digunakan. Grid Search mencari kombinasi *hyperparameter* terbaik untuk model SVM dengan membagi rentang parameter yang akan dioptimalkan menjadi sebuah grid, dan akan melintasi semua titik sehingga mendapatkan parameter paling optimal. Dalam konteks penelitian ini, variabel yang diubah-ubah adalah parameter C dan gamma. Nilai-nilai yang digunakan untuk parameter 'C' adalah [0.1, 1, 10, 100, 1000] dan untuk parameter 'gamma' adalah [1, 0.1, 0.01, 0.001, 0.0001]. Rentang nilai ini dipilih untuk mencakup sejumlah besar kemungkinan nilai, sehingga memungkinkan Grid Search untuk mengeksplorasi ruang hiper-parameter dengan baik.

2.6. Support Vector Machine

Support Vector Machine (SVM) adalah sebuah sistem pembelajaran yang memanfaatkan ruang hipotesis berupa fungsi-fungsi linier dalam fitur yang memiliki dimensi tinggi [7]. SVM dilatih dengan menggunakan algoritma pembelajaran yang berbasis teori optimasi. Tingkat keakuratan yang dihasilkan oleh model SVM sangat bergantung pada pengaturan parameter dan jenis fungsi kernel yang digunakan. SVM dapat diklasifikasikan menjadi dua tipe, yaitu SVM linear dan SVM non-linear. SVM linear digunakan ketika data dapat dipisahkan secara linier, sementara SVM non-linear digunakan untuk data yang tidak bisa dipisahkan secara linier dan biasanya menggunakan kernel untuk memungkinkan pemisahan tersebut. Gambar 5 mengilustrasikan bagaimana SVM bekerja dengan menemukan *hyperplane* dengan margin maksimum yang memisahkan dua kelas.



Gambar 5. Ilustrasi Support Vector Machine

2.7. Evaluasi

Performa diukur dengan cara memeriksa berapa banyak pengujian yang berhasil diprediksi secara tepat dan berapa yang tidak tepat oleh model. Informasi ini biasanya disusun dalam sebuah tabel yang dikenal sebagai confusion matrix.

Tabel 1. Confusion Matrix

Kelas Sebenarnya	Kelas Prediksi			
	Positif		Negatif	
	TP	FN	FP	TN

Berikut merupakan persamaan yang digunakan untuk mengukur performa model dalam bentuk akurasi.

$$Akurasi = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \% \tag{7}$$

Dengan keterangan sebagai berikut:

- True Positive (TP) = Jumlah data positif yang berhasil diklasifikasikan oleh model
- False Negative (FN) = Jumlah data positif yang diklasifikasikan salah oleh model
- False Positive (FP) = Jumlah data negatif yang diklasifikasikan salah oleh model
- True Negative (TN) = Jumlah data negatif yang diklasifikasikan benar oleh model

3. Hasil dan Diskusi

Pada pengujian yang dilakukan, kombinasi parameter yang digunakan adalah sudut yang digunakan pada GLCM dan hasil parameter C dan gamma setelah grid search. Kemudian arah GLCM dalam bentuk sudut yang digunakan adalah sudut 0°, 45°, 90° dan 135°. Data hasil GLCM sebelum dan sete tersebut akan dilatih pada model SVM yang kemudian akan diujikan untuk mencari akurasi model dalam proses klasifikasi. Akurasi dari hasil pengujian setelah dan sebelum penyetelan *hyperparameter* dapat dilihat pada tabel 2.

Tabel 2. Hasil Uji Coba

	Sebelum Hyperparameter Tuning				Setelah Hyperparameter Tuning				
	Precision	Recall	F1-Score	Support	Precision	Recall	F1-Score	Support	
GLCM 0	Lung Opacity	0.52	0.29	0.37	203	0.58	0.51	0.54	203
	Normal	0.47	0.62	0.53	206	0.64	0.51	0.57	206
	Viral Pneumonia	0.57	0.64	0.6	191	0.64	0.86	0.74	191
	Accuracy	-	-	0.51	600	-	-	0.62	600
	Macro Avg	0.52	0.52	0.5	600	0.62	0.63	0.62	600
	Weighted Avg	0.52	0.52	0.5	600	0.62	0.62	0.61	600
GLCM 45	Lung Opacity	0.35	1	0.52	203	0.6	0.69	0.64	203
	Normal	0	0	0	206	0.47	0.48	0.48	206

		Sebelum Hyperparameter Tuning				Setelah Hyperparameter Tuning			
		Precision	Recall	F1-Score	Support	Precision	Recall	F1-Score	Support
	Viral Pneumonia	1	0.08	0.15	191	0.6	0.5	0.54	191
	Accuracy	-	-	0.36	600	-	-	0.56	600
	Macro Avg	0.45	0.36	0.22	600	0.56	0.56	0.55	600
	Weighted Avg	0.44	0.36	0.22	600	0.56	0.56	0.55	600
	Lung Opacity	0.35	1	0.52	203	0.68	0.72	0.7	203
GLCM 90	Normal	0	0	0	206	0.62	0.44	0.52	206
	Viral Pneumonia	1	0.08	0.15	191	0.69	0.85	0.76	191
	Accuracy	-	-	0.36	600	-	-	0.67	600
	Macro Avg	0.45	0.36	0.22	600	0.66	0.67	0.66	600
	Weighted Avg	0.44	0.36	0.22	600	0.66	0.67	0.66	600
GLCM 135	Lung Opacity	0.35	1	0.52	203	0.61	0.65	0.63	203
	Normal	0	0	0	206	0.5	0.5	0.5	206
	Viral Pneumonia	0.94	0.08	0.14	191	0.59	0.54	0.57	191
	Accuracy	-	-	0.36	600	-	-	0.56	600
	Macro Avg	0.43	0.36	0.22	600	0.57	0.56	0.56	600
GLCM COMBINED	Weighted Avg	0.42	0.36	0.22	600	0.57	0.56	0.56	600
	Lung Opacity	0.35	1	0.52	203	0.73	0.67	0.7	203
	Normal	0	0	0	206	0.7	0.74	0.72	206
	Viral Pneumonia	1	0.08	0.15	191	0.9	0.92	0.91	191
	Accuracy	-	-	0.36	600	-	-	0.77	600
	Macro Avg	0.45	0.36	0.22	600	0.78	0.78	0.78	600

	Sebelum Hyperparameter Tuning				Setelah Hyperparameter Tuning			
	Precision	Recall	F1-Score	Support	Precision	Recall	F1-Score	Support
Weighted Avg	0.44	0.36	0.22	600	0.77	0.77	0.77	600

Berdasarkan hasil pengujian, ditemukan bahwa model dengan konfigurasi GLCM dengan sudut yang dikombinasikan menghasilkan akurasi tertinggi, mencapai 0.77 setelah dilakukan penyetelan *hyperparameter* menggunakan metode grid search dimana parameternya adalah {'C': 10000, 'gamma': 0.0001, 'kernel': 'rbf'} dan score sebesar 0.7895833333333333. Sebaliknya, sebelum dilakukan penyetelan *hyperparameter* model awal adalah {'C': 1.0, 'gamma': 'scale', 'kernel': 'rbf'} dengan hasil akurasi mencapai 0.36. Dimana berdasarkan data tersebut didapat bahwa rata-rata akurasi sebelum penyetelan *hyperparameter* mendapatkan rata rata akurasi sebesar 0.3975 dan setelah 0.6025 menunjukkan bahwa grid search cukup efektif dalam mengoptimalkan kinerja model SVM. Perbaikan signifikan dari rata-rata akurasi sebesar 0.205 menunjukkan bahwa penyetelan *hyperparameter* menggunakan metode grid search berhasil meningkatkan kinerja model secara substansial. Hal ini menunjukkan pentingnya penyetelan *hyperparameter* dalam proses pengembangan model untuk mencapai kinerja yang optimal. Dengan menggunakan konfigurasi {'C': 10000, 'gamma': 0.0001, 'kernel': 'rbf'} setelah penyetelan *hyperparameter*, model SVM mampu mencapai akurasi sebesar 0.67, yang merupakan peningkatan signifikan dari akurasi sebelumnya sebesar 0.36. Konfigurasi ini menunjukkan bahwa nilai C yang tinggi (10000) mengindikasikan bahwa model SVM cenderung untuk memisahkan data dengan batas keputusan yang sangat ketat, sementara nilai gamma yang rendah (0.0001) menunjukkan bahwa pengaruh dari satu sampel terbatas pada jarak yang cukup jauh dari batas keputusan. Dengan demikian, hasil ini menegaskan bahwa pemilihan *hyperparameter* yang tepat sangat penting dalam meningkatkan kinerja model SVM, dan metode grid search adalah pendekatan yang efektif dalam menemukan kombinasi *hyperparameter* yang optimal.

4. Kesimpulan

Dalam penelitian ini, langkah-langkah metodologi yang komprehensif telah dijalankan untuk menganalisis dataset gambar. Tahapan awal melibatkan persiapan data dengan melakukan normalisasi label dan deskripsi, serta penerapan teknik CLAHE dan kontras stretching untuk meningkatkan kualitas gambar. Selanjutnya, fitur-fitur tekstur diekstraksi menggunakan metode GLCM, yang memungkinkan analisis citra digital dari sudut pandang tekstur. Proses ini menghasilkan serangkaian fitur seperti kontras, dissimilarity, homogeneity, ASM, energy, dan correlation. Setelah data diproses, Data tersebut akan dibagi dua bagian yaitu data uji 20% dan data latih 80%, yang selanjutnya dianalisis menggunakan algoritma Support Vector Machine (SVM). Pentingnya tuning *hyperparameter* diakui dengan menerapkan metode grid search, yang memungkinkan identifikasi kombinasi *hyperparameter* terbaik untuk meningkatkan kinerja model dan mencegah overfitting. Evaluasi performa model dilakukan melalui confusion matrix dan mengukur akurasi. Hasil pengujian menunjukkan bahwa konfigurasi GLCM pada sudut yang dikombinasikan menghasilkan akurasi tertinggi setelah penyetelan *hyperparameter* menggunakan grid search, sementara model SVM sebelum penyetelan menunjukkan akurasi rendah. Dengan demikian, penelitian ini menegaskan pentingnya proses preprocessing, ekstraksi fitur, tuning *hyperparameter*, dan evaluasi model dalam mencapai hasil yang optimal dalam klasifikasi gambar menggunakan SVM.

Daftar Pustaka

- [1] S. Bharati, P. Podder, and R. Hossain, "Hybrid deep learning for detecting lung diseases from X-ray images," *Journal of Informatics in Medicine Unlocked*, vol. 20, 2020. DOI: 10.1016/j.imu.2020.100391.
- [2] B. Nugroho and E. Y. Puspaningrum, "Kinerja Metode CNN untuk Klasifikasi Pneumonia dengan Variasi Ukuran Citra Input," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*,

- vol. 8, no. 3, pp. 533-538, 2021. DOI: 10.25126/jtiik.202184515.
- [3] "Pneumonia death rates, age-standardized," IHME, Global Burden of Disease, 2019. [Online]. Available: <https://ourworldindata.org/grapher/pneumonia-death-rates-age-standardized?tab=table&time=2019..latest>.
- [4] "Pneumonia Fact Sheet," World Health Organization (WHO), 2020. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/pneumonia#:~:text=Pneumonia%20accounts%20for%2014%25%20of, and %20by%20addressing%20environmental%20factors>.
- [5] S. S. Panna, A. Alhamad, and K. C. Pelangi, "Pengenalan Ekspresi Wajah Pengemudi Berbasis Fitur Eigenface dan Gray Level Co-Occurance Matrice," Jurnal Edukasi dan Penelitian Informatika, vol. 9, no. 2, 2023.
- [6] M. Yunus, "Feature Extraction: Gray Level Co-occurrence Matrix (GLCM)," [Online]. Available: <https://yunusmuhammad007.medium.com/feature-extraction-gray-level-co-occurrence-matrix-g lcm-10c45b6d46a1>.
- [7] I. M. Sulistiana and M. A. Muslim, "Optimasi Support Vector Machine (SVM) Menggunakan Grid Search dan Unigram untuk Meningkatkan Akurasi Ulasan E-Commerce," Journal of Self-Computing Exploration (JOSTex), vol. 11, no. 3, pp. 1234-1245, 2020. [Online]. Available: <https://jurnal.uns.ac.id/verstek>.
- [8] I. Syarif, A. Prugel-Bennett, and G. Wills, "SVM parameter optimization using grid search and genetic algorithm to improve classification performance," Telkomnika, vol. 14, no. 4, pp. 1502-1509, 2016.
- [9] S. W. Lin, K. C. Ying, S. C. Chen, and Z. J. Lee, "Particle swarm optimization for parameter determination and feature selection of support vector machines," Expert Systems with Applications, vol. 35, no. 4, pp. 1817-1824, 2009.
- [10] F. Mehrparvar, "Lung Disease Dataset." [Online]. Available: <https://www.kaggle.com/datasets/fatemehmehparvar/lung-disease>.