

Model Hibrida DenseNet-161 dan XGBoost untuk Klasifikasi Multi-Kelas Kelainan Paru-Paru Berbasis Citra CT-Scan

Irma Amelia Dewi¹, Faris Akbar Shiddiq², Muhammad Ichwan²

^{1, 2, 3}Program Studi Informatika, Institut Teknologi Nasional Bandung, Indonesia

Email: ¹irma_amelia@itenas.ac.id, ²farisakbaaar@gmail.com, ³ichwan@itenas.ac.id

Abstrak - CT-Scan paru-paru membantu mendeteksi gejala awal, memantau progresi penyakit, dan mengidentifikasi komplikasi paru-paru. Selain itu terdapat kesamaan karakteristik antar penyakit paru-paru lain seperti pneumonia, tumor, kanker. Sehingga CT-Scan paru-paru menjadi alat diagnostik penting dalam penanganan penyakit-penyakit tersebut. Hasil CT-Scan dapat diklasifikasikan menggunakan Deep Learning yaitu Convolutional Neural Network dengan menggunakan arsitektur DenseNet-161 untuk ekstraksi fitur dari gambar CT-Scan paru-paru, yang kemudian diklasifikasikan menggunakan Extreme Gradient Boosting (XGBoost). Hasil eksperimen menunjukkan bahwa kombinasi DenseNet-161 dan XGBoost memberikan akurasi tinggi dalam klasifikasi penyakit paru-paru. DenseNet-161 digunakan untuk mengekstraksi fitur dari dataset citra CT-Scan, yang kemudian digunakan sebagai input untuk XGBoost. Evaluasi menggunakan metrik yang relevan menunjukkan bahwa model ini berhasil mengklasifikasikan 5 jenis penyakit paru-paru dengan rata-rata akurasi mencapai 98%. Analisis mendalam menunjukkan bahwa penggunaan $n_estimators$ yang besar dapat menyebabkan overfitting. Namun, kombinasi DenseNet-161 dengan XGBoost memberikan akurasi training sebesar 98.99% dan akurasi validasi sebesar 99.13%, dengan precision, recall, dan F1-score mendekati 100%. Hasil ini menegaskan pentingnya memilih kombinasi model yang tepat untuk meningkatkan kinerja secara signifikan, terutama dalam analisis data kompleks seperti CT-Scan.

Kata kunci: CT-Scan Paru-Paru, Convolutional Neural Network, Densenet-161, Ekstraksi Fitur, XGBoost.

Abstract - A lung CT scan aids in detecting early symptoms, monitoring disease progression, and identifying pulmonary complications. Additionally, there are common characteristics shared with other lung diseases such as pneumonia, tumors, and cancer. Consequently, a lung CT scan serves as a vital diagnostic tool in the management of these conditions. CT scan results can be classified using Deep Learning, specifically a Convolutional Neural Network employing the DenseNet-161 architecture for feature extraction from lung CT scan images, which are then classified using Extreme Gradient Boosting (XGBoost). Experimental results show that the combination of DenseNet-161 and XGBoost provides high accuracy in classifying lung diseases. DenseNet-161 is used to extract features from the CT scan image dataset, which are then used as input for XGBoost. Evaluation using relevant metrics shows that this model successfully classifies 5 types of lung diseases with an average accuracy of 98%. In-depth analysis indicates that using a large number of $n_estimators$ can lead to overfitting. However, the combination of DenseNet-161 with XGBoost yields a training accuracy of 98.99% and a validation accuracy of 99.13%, with precision, recall, and F1-score approaching 100%. These results underscore the importance of selecting the right model

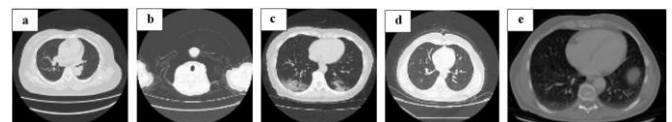
combination to significantly improve performance, particularly in the analysis of complex data such as CT scans.

Keywords - Lung CT-Scan, Convolutional Neural Network, Densenet-161, Feature Extraction, XGBoost.

I. PENDAHULUAN

Penyakit pada sistem pernapasan, mulai dari infeksi paru akut hingga kondisi kronis seperti Lung Adenokarsinoma atau kanker paru-paru yang paling umum ditemukan di Amerika Serikat, termasuk dalam kategori kanker paru-paru non-sel kecil (NSCLC) [1][1]. Computed Tomography (CT) Scan paru-paru menjadi alat diagnostik yang penting dalam penanganan pasien dalam mendeteksi gejala awal, memantau progresi penyakit, dan mengidentifikasi komplikasi paru-paru yang terjadi pada pasien [2].

Kendati demikian, interpretasi citra CT-Scan paru-paru secara manual sering kali dihadapkan pada kendala subjektivitas tinggi dan waktu analisis yang lama. Masalah ini diperumit oleh fakta bahwa terdapat tumpang tindih (*overlap*) karakteristik visual berupa pola opasitas atau densitas jaringan antara pneumonia akibat infeksi virus/bakteri dengan lesi tumor maupun tanda keganasan lainnya. Oleh karena itu, diperlukan metode penapisan komparatif yang lebih mendalam dan objektif guna membedakan anomali tersebut secara multi-kelas.



Gambar 1. Kelas citra CT-Scan paru-paru: (a) LMT (Lung Mass Tumor) (b) NiCT (Non-Informative CT) (c) pCT (Positive CT) (d) nCT (Negative CT) (e) Adenocarcinoma (ADC)

Hasil CT-Scan dapat diklasifikasikan menggunakan Deep Learning untuk menganalisis citra medis dengan menggunakan model Convolutional Neural Network [3]. Salah satu model pada CNN adalah Densenet yang memiliki beragam layer, yaitu Densenet-121, Densenet-161, Densenet-201, dan Densenet-264. Pemilihan layer Densenet perlu diperhatikan karena peningkatan jumlah layer sering kali menyebabkan kompleksitas model yang lebih besar dan meningkatkan jumlah parameter, yang dapat menyebabkan masalah seperti overfitting dan memerlukan lebih banyak sumber daya komputasi untuk pelatihan [4].

Pemilihan model *DenseNet-161* dalam penelitian ini didasarkan pada perbedaan utama dari *Densenet* lainnya yaitu penggunaan *filter* sebesar 48 ($K=48$), sedangkan *layer Densenet* lainnya hanya memiliki *filter* sebesar 32. Nilai “ K ” ini adalah salah satu dari banyak *hyperparameter* yang dapat disesuaikan dalam model *Densenet-161*. Nilai ini biasanya dipilih berdasarkan eksperimen dan penyietelan terhadap *dataset* yang spesifik. Jumlah *filter* yang lebih besar dapat memberikan model lebih banyak kapasitas untuk mempelajari fitur-fitur yang lebih kompleks [5].

Extreme Gradient Boosting atau *XGBoost* merupakan varian dari peningkatan algoritma *tree gradient* dengan kombinasi antara *Gradient Descent* dan *Boosting* yang bekerja dengan metode *boosting* yaitu kumpulan pohon keputusan di mana pengembangan pohon berikutnya akan bergantung pada pohon sebelumnya. Pohon pertama di *XGBoost* akan lemah dalam melakukan klasifikasi dengan inisialisasi probabilitas yang ditentukan oleh peneliti kemudian memperbarui bobot dari setiap pohon yang dibangun untuk menghasilkan kumpulan pohon klasifikasi yang kuat [6].

Penelitian sebelumnya [7] menunjukkan bahwa model *DenseNet161* mampu mengklasifikasikan objek berbeda pada citra medis dengan tingkat akurasi masing-masing sebesar 95,02% dan 95,20%. Sementara itu, penelitian lainnya [3] melakukan perbandingan kinerja sembilan arsitektur CNN serta dua jenis pengoptimalisasi (*optimizer*) yaitu Adam dan SGD untuk klasifikasi multi-kelas penyakit paru-paru (*Pneumonia*, *Lung Opacity*, dan Kondisi Normal) menggunakan citra sinar-X toraks. Hasil akhir menunjukkan bahwa kombinasi arsitektur berbasis *ResNet50v1* dengan pengoptimalisasi Adam menghasilkan kinerja tertinggi dengan nilai rerata akurasi mencapai $92,22 \pm 0,25\%$.

Pendekatan klasifikasi dua tahap pada citra CT-Scan paru-paru juga dikembangkan [8]. Penelitian tersebut menggunakan empat model dasar CNN (*ResNeXt-50*, *DenseNet161*, *Inception-v3*, dan *Wide-ResNet-50*) sebagai pengekstrak fitur, yang dipadukan dengan pengklasifikasi *Support Vector Machine* (SVM) dan *Extreme Gradient Boosting* (*XGBoost*). Hasil pengujian membuktikan bahwa *DenseNet161* dan *Inception-v3* memberikan kinerja akurasi tertinggi, baik saat dipadukan dengan SVM (82,65%) maupun *XGBoost* (hingga 85,71%) [8].

Penelitian sebelumnya mengembangkan model hibrida CNN dan *XGBoost* untuk diagnosis dini COVID-19 berbasis citra X-ray dada. Pada pendekatan tersebut, lapisan *fully connected* pada CNN sederhana digantikan oleh *XGBoost* sebagai pengklasifikasi fitur hasil ekstraksi CNN. Selain itu, *hyperparameter XGBoost* dioptimasi menggunakan *modified Arithmetic Optimization Algorithm*. Model tersebut divalidasi pada dataset X-ray dada seimbang yang terdiri atas 12.000 citra dari tiga kelas, yaitu normal, COVID-19, dan *pneumonia*, serta menghasilkan akurasi sekitar 99,39% [9].

Di samping itu, pengembangan model hibrida CNN-*XGBoost* yang dioptimalkan dengan *Adaptive Swarm Optimization* (APSO) untuk klasifikasi citra terbukti mampu menghasilkan akurasi sebesar 88,20% [10]. Terkait

pemilihan parameter latih yang optimal menegaskan dalam studinya mengenai dampak algoritma optimasi bahwa dari sembilan *optimizer* yang dievaluasi untuk klasifikasi citra, pengoptimalisasi Adam memberikan hasil terbaik disusul oleh Nadam dan Adamax [11].

Berdasarkan penelitian terdahulu, kontribusi penelitian ini adalah menggabungkan dan menyempurnakan keunggulan metodologi dari penelitian-penelitian sebelumnya. Penelitian ini mengimplementasikan model *DenseNet161* untuk ekstraksi fitur spasial secara otomatis yang telah terbukti tangguh pada penelitian [7], [8], [9], [11], [12].

Meskipun beberapa penelitian sebelumnya telah menerapkan *DenseNet-161*, CNN, dan *XGBoost* pada klasifikasi citra medis, masih terdapat beberapa keterbatasan yang dapat diidentifikasi. Pertama, sebagian penelitian masih berfokus pada klasifikasi objek atau penyakit tertentu dengan cakupan kelas yang terbatas, sehingga belum merepresentasikan kompleksitas variasi kelainan paru-paru secara lebih luas. Kedua, beberapa pendekatan hibrida CNN-*XGBoost* telah digunakan pada citra medis, namun umumnya masih diterapkan pada kasus tertentu seperti COVID-19, *pneumonia*, atau klasifikasi citra umum, bukan pada klasifikasi multi-kelas kelainan paru-paru berbasis CT-Scan yang mencakup kondisi infeksi, non-infeksi, citra non-informatif, tumor massa paru, dan adenokarsinoma. Ketiga, penelitian sebelumnya belum secara khusus mengevaluasi kombinasi *DenseNet-161* sebagai ekstraktor fitur dan *XGBoost* sebagai pengklasifikasi untuk membedakan lima kelas kelainan paru-paru dengan karakteristik visual yang saling tumpang tindih.

Perbedaan utama sekaligus kebaruan dalam penelitian ini terletak pada cakupan dataset citra CT-Scan paru-paru yang lebih variatif untuk menangani kompleksitas multi-kelas, yang meliputi: citra dengan kondisi infeksi akut (*positive infection*), kondisi normal/non-infeksi (*negative CT*), *Non-Informative CT*, *Lung Mass Tumor*, serta *Adenokarsinoma* (kanker paru). Pendekatan ini diharapkan dapat memastikan tingkat akurasi dan efisiensi yang lebih tinggi dalam prosedur penapisan komprehensif berbagai kelainan paru-paru.

Melihat permasalahan yang telah dijelaskan, dilakukan eksplorasi dan eksperimen terhadap hal-hal pendukung yang dapat mengatasi permasalahan tersebut, sehingga dapat mengklasifikasikan penyakit paru-paru menggunakan *DenseNet-161* dan *XGBoost*. *DenseNet-161* digunakan untuk mengekstraksi fitur-fitur dari *dataset* citra CT-Scan, kemudian setiap gambar akan direpresentasikan oleh vektor fitur yang dihasilkan *DenseNet-161* dan digunakan sebagai input untuk *XGBoost*. Selanjutnya, dilakukan pelatihan dan penyietelan *hyperparameter* pada model *XGBoost*. Evaluasi kinerja model *XGBoost* yang telah dioptimalkan dilakukan menggunakan metrik evaluasi yang relevan untuk klasifikasi penyakit paru-paru.

II. METODE PENELITIAN

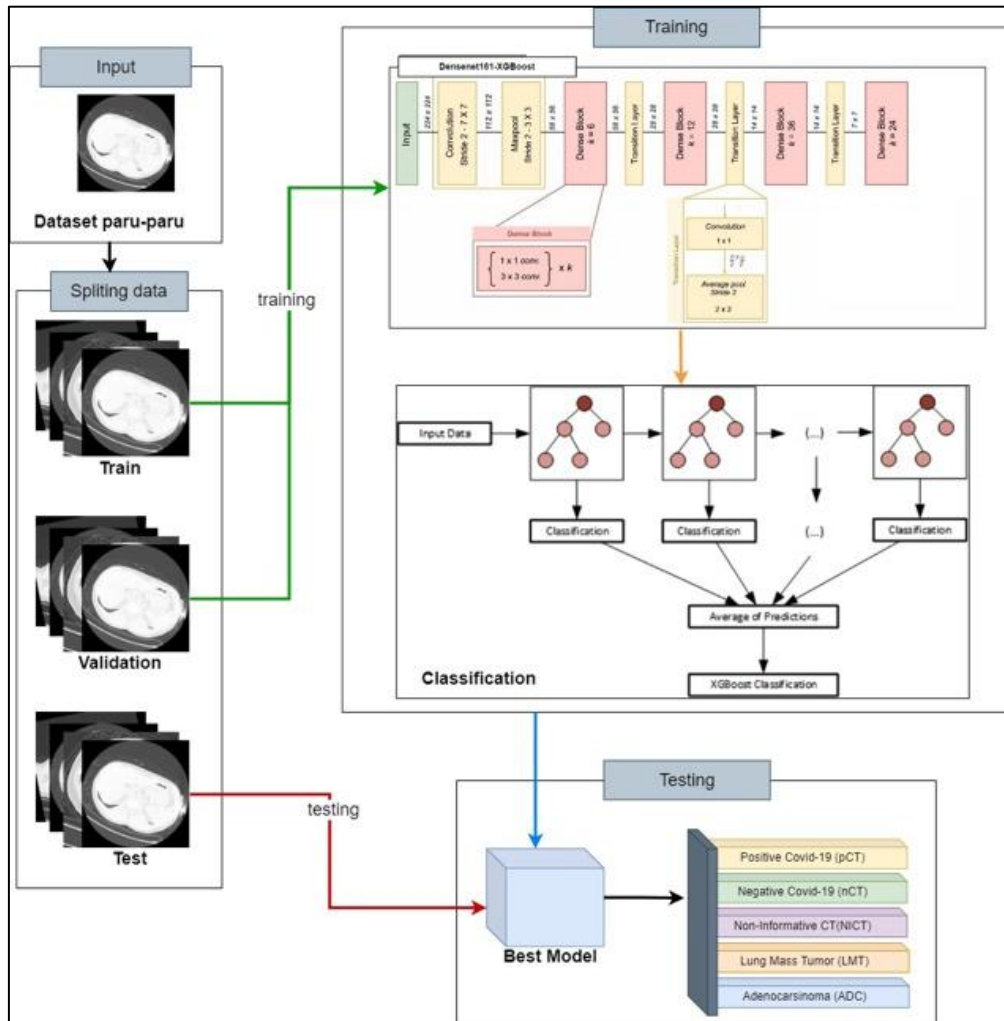
Metode penelitian merupakan rancangan tahapan yang akan dilakukan dalam penelitian. Pada Gambar 2

menunjukkan blok diagram mengenai alur kerja dari metode penelitian ini.

A. Deskripsi Data

Data yang digunakan terdiri atas dataset primer dan dataset sekunder. *Dataset* primer diperoleh dari Rumah Sakit Dustira Cimahi terdapat satu kelas dengan total citra 763 *CT-Scan* paru-paru yang terkena *Lung Mass Tumor*. *Dataset* sekunder diambil dari *open data Kaggle* yang dikumpulkan dari Rumah Sakit Union dan Rumah Sakit

Liyuan dengan berisikan empat kelas hasil citra *CT-Scan* yaitu, *non-informative CT* (NiCT) dimana parenkim paru-paru tidak ditangkap untuk penelitian apa pun, *positive CT* (pCT) dimana citra merepresentasikan manifestasi infeksi paru akut atau pneumonia spesifik seperti COVID-19, *negative CT* (nCT) dimana fitur pencitraan di kedua paru-paru tidak relevan, dan *CT Adenocarcinoma* (ADC) adalah citra kanker paru-paru. Tabel 1 adalah jumlah *dataset* gambar *CT-Scan* paru-paru.



Gambar 2. Blok diagram metode penelitian

B. Preprocessing Data

Preprocessing data augmentasi adalah untuk meningkatkan keberagaman dan mengatasi ketidakseimbangan data. Diantaranya adalah *random rotation*, *flip horizontal*, dan *resize image* [12]. Gambar 3 menunjukkan *flowchart* dari *preprocessing* data.

C. Splitting Data

Dataset dibagi untuk training digunakan untuk melatih model sebesar 70%, dataset validation untuk tuning hyperparameter sebesar 20%, dan testing untuk mengukur evaluasi model sebesar 10%. *Splitting* dataset untuk memastikan bahwa model yang dibangun tidak *overfitting* terhadap data training dan memiliki generalisasi yang baik [13]. Hasil *splitting* dataset diperoleh sebesar 2,408 total

data train, 692 total data validation dan 341 total data test. Hasil *splitting* dataset ditunjukkan pada Tabel 2.

1) DenseNet-161

Densenet (*Dense Convolutional Network*) adalah arsitektur yang berfokus untuk membuat jaringan pembelajaran mendalam menjadi lebih dalam, tetapi pada saat yang sama membuatnya lebih efisien untuk dilatih, dengan menggunakan koneksi yang lebih pendek antar lapisan [14]. Pada Gambar 4 merupakan arsitektur *Densenet-161*. Model *DenseNet-161* memiliki lapisan output terakhir yang umumnya memiliki jumlah neuron yang sama dengan jumlah kelas dalam *dataset* pelatihan. Setiap neuron di lapisan ini menghasilkan probabilitas

bahwa input gambar masuk ke dalam kelas yang sesuai dengan neuron tersebut.

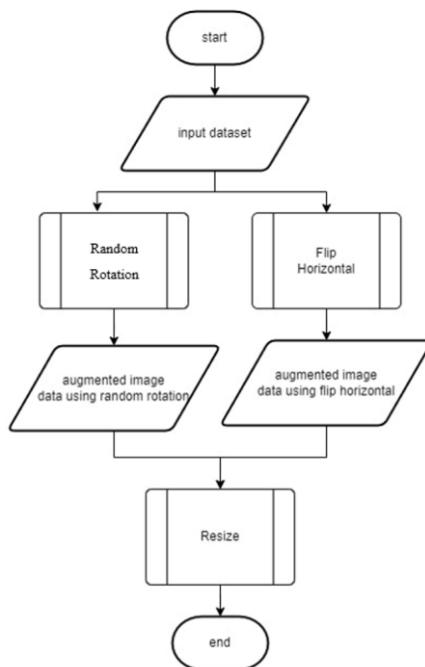
Arsitektur *DenseNet-161* memiliki alur yang terdiri dari beberapa blok yang berulang-ulang. Dalam arsitektur *Densenet-161*, dapat dijelaskan input menerima citra. Selanjutnya terdapat *convolutional block* dilakukan operasi konvolusi, *batch normalization*, dan fungsi aktivasi *ReLU*.

Tabel 1. Jumlah data

No	Nama Kelas	Jumlah Data
1	LMT (Lung Mass Tumor)	763
2	NiCT (Non-Informative CT)	661
3	pCT (Positive CT)	753
4	nCT (Negative CT)	500
5	Adenocarcinoma (ADC)	764
Total Jumlah Data		3441

Tabel 2. Splitting dataset

No	Nama Kelas	Train	Val	Test	Total
1	Lung Mass Tumor (LMT)	534	154	75	763
2	Non-Informative CT (NiCT)	462	133	66	661
3	Negative CT (nCT)	527	151	75	753
4	Positive CT (pCT)	350	101	49	500
5	Adenocarcinoma (ADC)	535	153	76	764
Total Jumlah Data		2408	692	341	3441



Gambar 3. Flowchart preprocessing

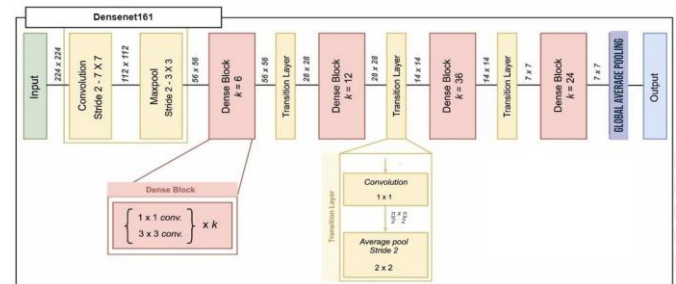
D. DenseNet-161 dan XGBoost

Pada pemodelan arsitektur yang digunakan adalah *DenseNet-161* untuk ekstraksi fitur dan *XGBoost* sebagai pengklasifikasi.

Dense block terdiri dari beberapa blok *dense layer* yang terhubung secara padat dengan operasi *concatenation* yang berfungsi untuk menggabungkan *output* dari setiap *layer*. Terdapat *transition layer* untuk mengurangi jumlah saluran (*channels*) dengan operasi konvolusi 1×1 dan melakukan *average pooling*. Pada *dense block* dan *transition layer* diulang sebanyak empat kali. Kemudian *Global Average Pooling* dilakukan untuk menghasilkan vektor fitur. *Output layer* menghasilkan prediksi probabilitas kelas untuk citra input.

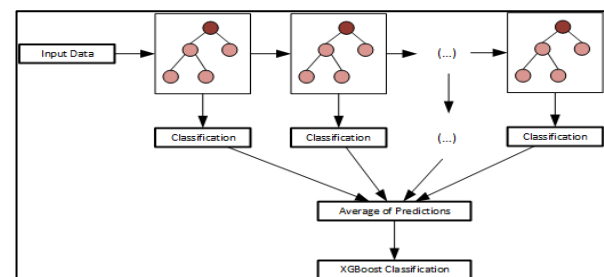
2) XGBOOST

XGBoost merupakan salah satu metode percepatan yaitu kumpulan pohon keputusan, di mana pembangunan pohon berikutnya bergantung pada pohon sebelumnya. *Extreme Gradient Boosting* merupakan algoritma yang dapat menemukan solusi yang optimal untuk berbagai masalah khususnya pada klasifikasi [15]. Gambar 5 merupakan proses klasifikasi pada *XGBoost*.



Gambar 4. Arsitektur Densenet-161

Pohon *XGBoost* pertama menampilkan klasifikasi yang lemah dimulai dengan probabilitas yang ditemukan oleh peneliti dan kemudian memperbarui bobot setiap pohon yang dibangun sehingga membangun kumpulan pohon klasifikasi yang kuat. *XGBoost* digunakan dalam konteks pembelajaran terarah (*supervised learning*) di mana terdapat data pelatihan dengan berbagai fitur X_i yang digunakan untuk memprediksi variabel target Y_i .



Gambar 5. Proses klasifikasi XGBoost

Model dan Parameter dalam *supervised learning* mengacu pada struktur matematika di mana prediksi Y_i dibuat berdasarkan input X_i . Sebagai contoh, model linier

merupakan contoh umum di mana prediksi dinyatakan sebagai $\hat{y}_i = \sum_j \theta_j X_{ij}$, yaitu kombinasi linier dari bobot fitur masukan. Nilai prediksi dapat diinterpretasikan dengan cara yang berbeda tergantung pada tugas yang dilakukan, apakah itu regresi atau klasifikasi. Parameter adalah variabel yang perlu dipelajari dari data. Dalam kasus regresi *linier*, parameter yang dimaksud adalah koefisien θ . Biasanya, simbol θ digunakan untuk merujuk pada parameter tersebut [6].

Dengan memilih variabel target y_i , dapat mengungkapkan berbagai tugas seperti regresi, klasifikasi, dan ranking. Tujuan dari pelatihan model adalah menemukan parameter θ yang optimal yang sesuai dengan data latihan X_i dan label y_i . Untuk melatih model, penting untuk mendefinisikan fungsi objektif yang digunakan untuk mengukur sejauh mana model cocok dengan data latihan. Fungsi objektif memiliki karakteristik penting yang terdiri dari dua bagian: *training loss* dan *regularization* seperti pada Persamaan (1) [10].

$$obj(\theta) = L(\theta) + \Omega(\theta) \quad (1)$$

Di mana L merupakan fungsi *training loss* dan Ω merupakan istilah terkait dengan regulasi. *Training loss* mengukur sejauh mana model dapat memprediksi data pelatihan dengan baik. L adalah kesalahan kuadrat rata-rata, yang diberikan oleh Persamaan (2) [6].

$$L(\theta) = \sum_i (y_i - \hat{y}_i)^2 \quad (2)$$

Training loss yang digunakan seperti pada Persamaan (3) [6].

$$L(\theta) = [y_i \ln(1 + e^{-\hat{y}_i}) + (1 - y_i) \ln(1 + e^{\hat{y}_i})] \quad (3)$$

Regularisasi mengatur tingkat kompleksitas model, yang berguna dalam mencegah *overfitting* [6].

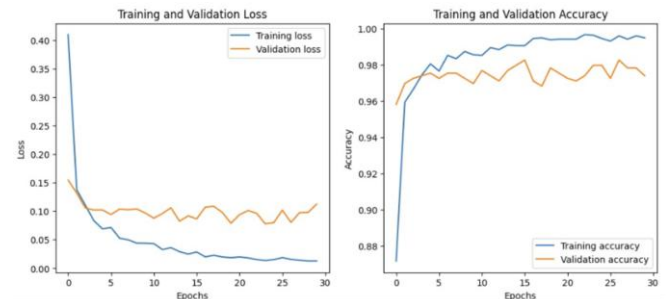
III. HASIL DAN PEMBAHASAN

Pelatihan dimulai dengan input data citra CT-Scan ke base model Densenet-161. Dalam pelatihan dilakukan percobaan sebanyak tujuh kali pelatihan untuk menemukan hyperparameter terbaik pada model DenseNet-161 yang direpresentasikan pada Tabel 3. Dari Tabel 3 dilakukan beberapa tuning hyperparameter model DenseNet-161 dengan optimizer dan epoch yang berbeda. Terlihat akurasi pelatihan tertinggi yang didapatkan adalah pada hyperparameter no 5 dengan optimizer AdamMax, learning rate 0.5, dan epoch 30 memperoleh akurasi pelatihan sebesar 97.98%.

Tabel 3. Perbandingan hyperparameter DenseNet-161

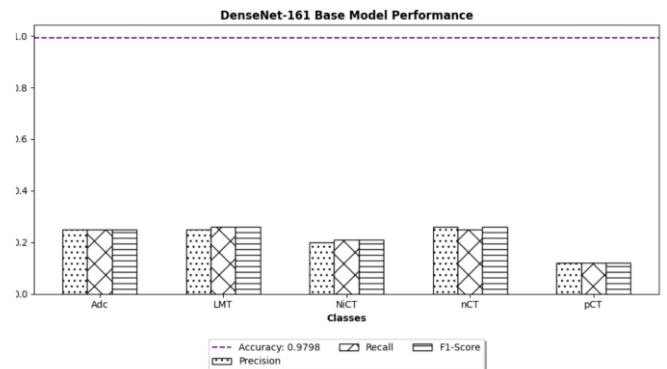
No	Optimizer	Learning rate	Epoch	Akurasi (%)
1	Adam	0.5	15	97.40
2	Adam	0.5	30	97.40
3	Adam	0.5	45	97.69
4	AdamMax	0.5	15	97.40
5	AdamMax	0.5	30	97.98
6	AdamMax	0.5	45	97.26
7	Nadam	0.5	15	22.94

Dapat dilihat dari Gambar 6 training loss menurun secara konsisten seiring bertambahnya jumlah epoch sementara validation loss meningkat, ini mengindikasikan adanya *overfitting* yaitu kemungkinan terjadi karena model sangat cocok dengan data pelatihan (training loss rendah dan akurasi tinggi), tetapi kesulitan untuk menggeneralisasi pada data validasi (validation loss meningkat dan akurasi validasi lebih rendah).



Gambar 6. Grafik pelatihan Densenet-161

Dapat dilihat Gambar 7 adalah hasil evaluasi metrik model dari pelatihan base model DenseNet-161 no 5. Berdasarkan evaluasi model Gambar 8 rata-rata, precision, recall, dan F1-score memiliki nilai sekitar 22%. Hal ini menunjukkan bahwa model memiliki kinerja yang rendah dalam memprediksi kelas yang ada dalam dataset. Hasil diantara ketiga metrik evaluasi tersebut menunjukkan keseimbangan dalam kinerja model di seluruh kelas.



Gambar 7. Evaluasi DenseNet-161 base model

Pelatihan menggunakan kombinasi DenseNet-161 dan XGBoost. Skenario perbandingan penggunaan tuning hyperparameter dalam pelatihan DenseNet-161 dan XGBoost ditunjukkan pada Tabel 4. Dapat dilihat dari Tabel 4 hasil pelatihan Densenet-161-XGBoost memperoleh akurasi yang stabil rata-rata 98% dengan hasil akurasi tertinggi diperoleh dari pelatihan model ke 13 menggunakan $n_estimators = 500$, $max_depth = 8$, $learning_rate = 0.5$, dan $subsample = 0.4$ dengan rata-rata akurasi sebesar 99.20%. Maka dari itu hasil dari pelatihan ke-13 diambil sebagai model terbaik dari kombinasi Densenet-161-XGBoost.

Faktor menggunakan hyperparameter tuning pada pelatihan model Densenet-161-XGBoost sangat berpengaruh, yakni perubahan nilai pada max_depth , $learning_rate$, dan $subsample$. Adapun faktor pada parameter $n_estimator$ yang mempengaruhi nilai akurasi, jika nilai $n_estimator$ semakin besar terdapat beberapa

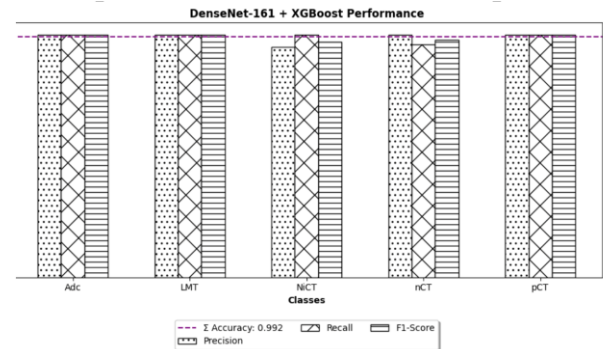
alasan yakni: (1) Overfitting, saat nilai $n_estimator$ semakin besar maka model ensemble menjadi semakin kompleks dan cenderung mengingat data pelatihan atau overfitting, berarti model akan menjadi sangat baik dalam menyesuaikan data pelatihan, tetapi tidak akan generalisasi dengan baik pada data baru; (2) Semakin besar nilai $n_estimator$ akan mengalami komputasi yang lebih lambat, hal ini disebabkan karena dataset yang terlalu besar; (3) Semakin besar nilai $n_estimator$ semakin kecil tingkat peningkatan gain informasi pada setiap langkah, karena model sudah cukup baik dalam melakukan pelatihan.

Penambahan $n_estimator$ tidak memberikan kontribusi signifikan dalam meningkatkan kinerja model. Pada penelitian ini diambil tingkat akurasi validasi dan akurasi tes terbaik dalam pembangunan model DenseNet-161 sebagai fitur ekstraksi dan model XGBoost sebagai klasifikasi yakni akurasi train atau pelatihan 98.99% dan akurasi validasi 99.13%. Gambar 8 adalah hasil evaluasi dari model Densenet-161-XGBoost.

Pada Gambar 7 dan Gambar 8 dapat dilihat perbandingan evaluasi model yang dihasilkan dari masing-masing model. Precision, recall, dan F1-score untuk setiap kelas hampir sempurna, dengan nilai mendekati 100%. Ini menunjukkan bahwa model tidak hanya dapat memprediksi kelas dengan benar tetapi juga dapat melakukannya dengan konsisten di seluruh kelas, sementara DenseNet-161 base model menunjukkan kinerja yang sangat rendah dengan nilai sekitar 22% saja.

Tabel 4. Perbandingan hyperparameter DenseNet-161 dan XGBoost

No	N Estimator	Sub Sample	Akurasi Train	Akurasi Val	Σ Akurasi
1	100	0,4	98.56%	98.55%	98.67%
2	100	0,5	98.85%	98.55%	98.67%
3	100	0,6	98.85%	98.84 %	98.93%
4	200	0,4	98.71%	98.55%	98.67%
5	200	0,5	98.85%	98.55%	98.67%
6	200	0,6	98.85%	98.84%	98.93%
7	300	0,4	98.88%	98.84%	98.93%
8	300	0,5	98.85%	98.55%	98.67%
9	300	0,6	98.85%	98.84%	98.93%
10	400	0,4	98.89%	99.13%	99.20%
11	400	0,5	98.71%	99.13%	99.20%
12	400	0,6	99.42%	98.55%	98.67%
13	500	0,4	98.99%	99.13%	99.20%
14	500	0,5	98.71%	99.13%	99.20%
15	500	0,6	99.42%	98.55%	98.67%



Gambar 8. Evaluasi DenseNet-161 dan XGBoost

Hasil penelitian menunjukkan bahwa penggunaan DenseNet-161 sebagai model klasifikasi langsung belum memberikan kinerja yang optimal, meskipun akurasi pelatihan yang diperoleh cukup tinggi. Kondisi ini terlihat dari adanya penurunan *training loss* yang tidak diikuti oleh perbaikan *validation loss*, sehingga mengindikasikan adanya kecenderungan *overfitting*. Hasil ini menunjukkan bahwa penggunaan DenseNet-161 secara langsung sebagai classifier belum cukup efektif pada dataset CT-Scan paru-paru multi-kelas yang memiliki karakteristik visual saling tumpang tindih.

Sebaliknya, kombinasi DenseNet-161 dan XGBoost menghasilkan kinerja yang jauh lebih baik. DenseNet-161 berperan sebagai ekstraktor fitur spasial dari citra CT-Scan, sedangkan XGBoost berperan sebagai pengklasifikasi berbasis *ensemble learning*. Hasil terbaik diperoleh pada kombinasi parameter $n_estimators = 500$, $max_depth = 8$, $learning_rate = 0,5$, dan $subsample = 0,4$, dengan akurasi pelatihan sebesar 98,99%, akurasi validasi sebesar 99,13%, dan akurasi pengujian sebesar 99,20%. Nilai precision, recall, dan F1-score yang mendekati 100% menunjukkan bahwa model hibrida mampu mengenali pola fitur dari setiap kelas secara lebih stabil dibandingkan model DenseNet-161 dasar.

DenseNet-161 memiliki kemampuan yang baik dalam mengekstraksi fitur visual pada citra medis [7]. Namun, hasil penelitian ini juga menunjukkan bahwa fitur yang dihasilkan oleh DenseNet-161 dapat dimanfaatkan secara lebih efektif ketika dikombinasikan dengan pengklasifikasi machine learning seperti XGBoost. Dengan cakupan kelas yang lebih beragam, hasil penelitian ini menunjukkan bahwa pendekatan hibrida DenseNet-161 dan XGBoost berpotensi digunakan untuk klasifikasi multi-kelas pada citra CT-Scan paru-paru, bukan hanya untuk membedakan kondisi normal dan abnormal atau kasus infeksi tertentu.

Implikasi utama dari penelitian ini bahwa pada kasus dataset citra medis dengan jumlah data terbatas dan karakteristik visual antar kelas yang mirip, penggunaan CNN sebagai ekstraktor fitur dan XGBoost sebagai pengklasifikasi dapat menjadi alternatif yang efektif. Pendekatan ini dapat membantu meningkatkan kinerja klasifikasi karena XGBoost memiliki mekanisme *ensemble* dan regularisasi yang dapat memperkuat proses pengambilan keputusan berdasarkan fitur hasil ekstraksi CNN.

Meskipun demikian, penelitian ini masih terbatas pada pembagian data training, validation, dan testing dari dataset yang digunakan, sehingga generalisasi model terhadap data dari rumah sakit, perangkat CT-Scan, atau populasi pasien yang berbeda masih perlu diuji lebih lanjut. Oleh karena itu, penelitian selanjutnya disarankan untuk melakukan validasi eksternal, pengujian dengan data multi-institusi, serta analisis kesalahan klasifikasi untuk mengetahui kelas-kelas yang masih sulit dibedakan oleh model.

IV. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa penerapan model *DenseNet-161* dan *XGBoost* dalam mengklasifikasikan lima kelas penyakit paru-paru berdasarkan citra *CT-Scan* memberikan hasil yang sangat baik, dengan rata-rata akurasi mencapai 98%. Hasil terbaik diperoleh pada pelatihan model ke-13 dengan *tuning hyperparameter* $n_estimators = 500$, $max_depth = 8$, $learning_rate = 0.5$, dan $subsample = 0.4$ yang menghasilkan akurasi sebesar 99,20%. Hal ini menegaskan pentingnya *tuning hyperparameter* dalam meningkatkan kinerja model, terutama perubahan pada max_depth , $learning_rate$, dan $subsample$. Analisis menunjukkan bahwa penggunaan $n_estimator$ yang besar dapat menyebabkan *overfitting*, memperlambat komputasi, dan memberikan peningkatan *gain* informasi yang semakin kecil. Namun, dalam penelitian ini, kombinasi model *Densenet-161* sebagai fitur ekstraksi dan *XGBoost* sebagai pengklasifikasi menghasilkan akurasi pelatihan sebesar 98,99% dan akurasi validasi sebesar 99,13%. Oleh karena itu, hasil ini menekankan pentingnya memilih kombinasi model yang tepat dan bagaimana penggunaan pengklasifikasi yang kuat seperti *XGBoost* dapat meningkatkan kinerja secara signifikan.

DAFTAR PUSTAKA

- [1] S. Kanchustambham, V. Sharma, "Lung Adenocarcinoma," *StatPearls Publishing*, 2026.
- [2] R. A. Hastari and S. Mulyati, "Rancangan Sistem Informasi Manajemen Monitoring Kebutuhan Isolasi Mandiri Pasien," *J. Autom.*, vol. 2, no. 2, pp. 1–14, 2021.
- [3] D. Gunawan and H. Setiawan, "Convolutional Neural Network dalam Citra Medis," *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 2, no. 2, pp. 376–390, 2022, doi: 10.24002/konstelasi.v2i2.5367
- [4] T. Chauhan, H. Palivela, and S. Tiwari, "Optimization and fine-tuning of DenseNet model for classification of COVID-19 cases in medical imaging," *Int. J. Inf. Manag. Data Insights*, vol. 1, no. 2, p. 100020, 2021, doi: 10.1016/j.jjimei.2021.100020. Available: <https://doi.org/10.1016/j.jjimei.2021.100020>
- [5] M. K. Ramadan, A. A. A. Youssif, and W. H. El-Behaidy, "Detection and Classification of Human-Carrying Baggage Using DenseNet-161 and Fit One Cycle," *Big Data Cogn. Comput.*, vol. 6, no. 4, 2022, doi: 10.3390/bdcc6040108
- [6] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, California, USA: Association for Computing Machinery, 2016, pp. 785–794. doi: <https://doi.org/10.1145/2939672.2939785>
- [7] A. Dhillon and G. K. Verma, "A Multiple Object Recognition Approach via DenseNet-161 Model," *Smart Electr. Mech. Syst.*, pp. 39–64, 2022, doi: 10.1016/B978-0-323-90789-7.00009-9
- [8] F. Bougourzi, R. Contino, C. Distanto, and A. Taleb-Ahmed, "Recognition of COVID-19 from CT scans using two-stage deep-learning-based approach: CNR-IEMN," *Sensors*, vol. 21, no. 17, 2021, doi: 10.3390/s21175878
- [9] M. Zivkovic et al., "Hybrid CNN and XGBoost Model Tuned by Modified Arithmetic Optimization Algorithm for COVID-19 Early Diagnostics from X-ray Images," *Electronics*, vol. 11, p. 3798, 2022, doi: 10.3390/electronics11223798
- [10] W. Jiao, X. Hao, and C. Qin, "The Image Classification Method with CNN-XGBoost Model Based on Adaptive Particle Swarm Optimization," *Information*, vol. 15, p. 156, 2021, doi: 10.3390/info12040156
- [11] J. Uddin, "Attention-Based DenseNet for Lung Cancer Classification Using CT Scan and Histopathological Images," *Designs*, vol. 8, p. 27, 2024, doi: 10.3390/designs8020027
- [12] Z. Gao, Y. Tian, S.-C. Lin, and J. Lin, "A CT Image Classification Network Framework for Lung Tumors Based on Pre-trained MobileNetV2 Model and Transfer learning, And Its Application and Market Analysis in the Medical field," *ArXiv*, 2025, doi: 10.48550/arXiv.2501.04996
- [13] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0197-0. Available: <https://doi.org/10.1186/s40537-019-0197-0>
- [14] J. Heaton, "Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning," *Genet. Program. Evolvable Mach.*, vol. 19, no. 1–2, pp. 305–307, 2018, doi: 10.1007/s10710-017-9314-z
- [15] R. Sarkar, A. Hazra, K. Sadhu, and P. Ghosh, "A Novel Method for Pneumonia Diagnosis from Chest X-Ray Images Using Deep Residual Learning with Separable Convolutional Networks," in *Computer Vision and Machine Intelligence in Medical Image Analysis. Advances in Intelligent Systems and Computing*, Springer, 2020, pp. 1–12. doi: https://doi.org/10.1007/978-981-13-8798-2_1