

Klasifikasi Penyakit Paru-Paru Berdasarkan Peningkatan Kualitas Kontras dan EfficientNet Menggunakan Gambar X-Ray

Asfa Dhevi Azzumzumi^{1*}, Muhammad Hanafi², Windha Mega Pradnya Dhuhi³

^{1,2,3}Program Studi Informatika, Universitas Amikom Yogyakarta, DI Yogyakarta
Email: ^{1*}asfa.az@students.amikom.ac.id, ²hanafi@amikom.ac.id, ³windha@amikom.ac.id

(Naskah masuk: 4 Jun 2024, direvisi: 21 Jun 2024, 2 Jul 2024, diterima: 3 Jul 2024)

Abstrak

COVID-19 dan penyakit paru-paru telah menjadi faktor utama penyebab kematian manusia di seluruh dunia. Kematian pasien dipengaruhi oleh keterlambatan deteksi dini. Sebagian besar profesional medis menggunakan gambar untuk mengidentifikasi kondisi paru-paru. Namun, para ahli yang dapat me-diagnosis dengan gambar sangat terbatas. Diagnosis gambar mendiagnosa menggunakan penglihatan manusia secara konvensional. Klasifikasi penyakit paru-paru sangat bervariasi. Masalah yang disebutkan di atas menunjukkan bahwa deteksi penyakit paru-paru dengan *Artificial Intelligence* (AI) yang efektif telah ditetapkan. Namun, sebagian besar hasil penyakit paru-paru salah didiagnosis. Bagi pasien, masalah ini menjadi masalah besar. Bertujuan untuk menangani klasifikasi penyakit paru-paru dengan deteksi kesalahan yang tinggi, kami menggunakan beberapa teknik *pre-processing* gambar dan menerapkan model pembelajaran mendalam dalam *EfficientNet*. Model *Pre-processing* termasuk augmentasi, peningkatan *white balance*, dan peningkatan kontras. Berdasarkan penelitian sebelumnya, mayoritas proses analisa gambar medis mengalami kualitas gambar yang rendah. Berdasarkan laporan eksperimen, model yang kami usulkan mencapai hasil yang signifikan dalam mengurangi kesalahan deteksi pada klasifikasi penyakit paru-paru. Dimana hasil *F1 score*-nya 0,97, *recall*-nya 0,98, presisinya 0,96, dan akurasi 0,97. Kami mempertimbangkan untuk menggunakan model yang kami usulkan dalam klasifikasi *multi-class*. Kami mengevaluasi model yang kami usulkan menggunakan *evaluation metric* dan *AUC Curve*.

Kata Kunci: Penyakit Paru-paru, *CLAHE*, *EfficientNet*, *Covid-19*, *White Balance*

Advanced Lung Diseases Classification Based on Contrast Enhancement and EfficientNet Using X-Ray Images

Abstract

COVID-19 and other lung illnesses have become a prominent factor because of the death toll of humans worldwide. The patient's death was being affected by late, early detection. The majority of medical professionals use visual material to identify lung conditions. Experts in medical imaging are scarce, nevertheless. When using conventional human eyesight, image diagnostics is used for diagnosis. The classification of lung disorders is highly variable. The aforementioned issue indicates that effective artificial intelligence (AI) lung disease detection has been established. Nonetheless, a significant majority of lung disease results are misdiagnosed. For patients, this issue becomes a major issue. In order to address the issue of high error detection when classifying lung disorders, we utilized multiple image pre-processing techniques and implemented a deep learning model within EfficientNet. The pre-processing model, which includes contrast, white balance, and augmentation. Previous studies have shown that the majority of medical imaging procedures result in images with poor quality. The experiment report indicates that our suggested model greatly reduced the number of false positives for lung illness classification. The results showed that the F1-score was 0,97, recall was 0,98, precision was 0,96, and accuracy was 0.97. We take into account using our suggested methodology for multi-class categorization. We evaluated our proposed model using evaluation metric and AUC Curve.

Keywords: Lung Disease, *CLAHE*, *EfficientNet*, *Covid-19*, *White Balance*

I. PENDAHULUAN

Penyakit pernafasan mempunyai dampak yang signifikan terhadap banyak orang dan masih menjadi penyebab utama kematian secara global. Penelitian telah menunjukkan bahwa individu dengan penyakit paru-paru yang sudah ada sebelumnya, termasuk bronkitis persisten, emfisema, asma bronkial, dan pneumonia, memiliki kemungkinan lebih tinggi terkena kanker paru-paru, meskipun mereka tidak merokok. Kemungkinan tertular penyakit paru-paru meningkat secara signifikan, terutama di negara-negara terbelakang atau kelas menengah ke bawah, karena banyak orang menderita kekurangan dan terpapar kontaminasi dunia luar. *World Health Organisation* (WHO) melaporkan bahwa hampir 4 juta kematian dini terjadi setiap tahunnya akibat penyakit yang disebabkan oleh polusi udara dalam ruangan, termasuk bronkitis dan tuberkulosis. Oleh karena itu, sangat penting untuk menggunakan teknik diagnostik yang sangat efektif untuk membantu identifikasi kelainan paru secara cepat [1] [2].

Deteksi dini dan pengobatan gangguan paru-paru sangat penting untuk mencapai hasil terbaik. Pemeriksaan dan diagnosis yang tepat waktu dapat membantu mengurangi sifat penyakit yang mengancam jiwa dan meningkatkan kualitas hidup orang yang terkena dampak. Rontgen dada/*Chest X-rays* (CXR) adalah teknik diagnostik umum yang digunakan untuk mengidentifikasi gangguan paru, menilai tingkat keparahannya, dan mendeteksi potensi akibatnya. Sejumlah penelitian terbaru menunjukkan kemanjuran metode segmentasi paru-paru dalam mengotomatisasi pemrosesan gambar rontgen dada. Rontgen dada dapat mengungkapkan beberapa kelainan secara bersamaan. Mendeteksi kelainan dalam konteks toraks yang kompleks hanya melalui observasi visual adalah proses yang sangat memakan waktu, dan pendekatannya rumit dan rentan terhadap subjektivitas pengguna [3]. Pelabelan manual oleh ahli radiologi memerlukan sumber daya kesehatan yang signifikan. Oleh karena itu, teknologi komputer dapat dimanfaatkan untuk menganalisis radiografi dada secara efektif, sebanding dengan ahli radiologi, dengan tujuan meningkatkan prioritas alur kerja dan memberikan bantuan dalam keputusan klinis. Hal ini dapat bermanfaat khususnya dalam inisiatif besar dan program kesehatan masyarakat global. Pembelajaran mendalam serta pembelajaran mesin memiliki potensi untuk berkontribusi besar pada diagnosis klinis yang tepat [4].

Algoritma yang memanfaatkan pembelajaran mendalam telah menunjukkan kinerja yang sangat baik dalam beberapa aplikasi komputer, seperti klasifikasi gambar, diagnosis medis, identifikasi adegan, prediksi penyakit, dan analisis layanan kesehatan [3], [4], [5], [6]. Kemajuan pesat algoritma pembelajaran mendalam telah dipercepat dengan pengembangan beberapa kumpulan data gambar yang telah diberi anotasi. Anotasi memiliki peran penting dalam mengatasi tantangan di berbagai bidang pemrosesan citra medis, seperti mengidentifikasi fitur anatomi dan patologi dalam pemindaian radiologi. Metode pembelajaran mendalam telah digunakan untuk mengidentifikasi dan mengkategorikan berbagai kondisi medis, termasuk limfoma, pneumonia

interstitial, deteksi pendarahan mikro fungsi otak, kategorisasi tumor di usus besar, perkiraan hasil radiologis neurologis, klasifikasi hati yang terkomputerisasi, dan identifikasi benjolan pada saluran pernapasan.

Negara-negara terbelakang menghadapi kelangkaan ahli radiologi yang mahir, terutama di daerah terpencil. Selain itu, masalah mengidentifikasi dan mengkategorikan penyakit paru-paru menggunakan pencitraan *Chest X-rays* merupakan tantangan berat bagi ahli radiologi. Oleh karena itu, penting bagi para peneliti untuk mengembangkan sistem otomatis untuk mendeteksi penyakit paru-paru. Dalam kasus seperti ini, teknologi pengujian otomatis dapat digunakan untuk melakukan penilaian ekstensif terhadap penyakit paru-paru dengan menganalisis pemindaian CXR. Ketepatan dalam mengidentifikasi gambar telah ditingkatkan dengan kombinasi kemajuan substansial dalam ketepatan pemrosesan dan aksesibilitas kumpulan data ekstensif yang berisi rontgen dada berlabel. Berbagai pendekatan kontemporer telah diusulkan untuk diagnosis otomatis penyakit paru-paru melalui gambar rontgen dada (CXR). Pada tahun 2017, Beberapa peneliti memperkenalkan koleksi gambar rontgen dada terbesar yang tersedia untuk umum, yang dikenal sebagai "*Chest X-rays 14*", yang mencakup 14 kelainan paru-paru yang paling umum. Beberapa investigasi dilakukan pada kumpulan data ekstensif ini. Wang dkk. mengusulkan kerangka klasifikasi multi-label dengan pengawasan lemah yang komprehensif yang menggabungkan beberapa fungsi kerugian DCNN multi-label dan teknik pengumpulan yang berbeda. Yao dkk. mengembangkan metode yang meningkatkan kinerja analisis rontgen dada dengan memanfaatkan korelasi label statistik. Metode ini efektif karena rontgen dada tunggal dapat menunjukkan beberapa pola yang menyimpang. Dengan cara yang sama, Kumar dkk. [1] menggunakan metode pembelajaran multi-label untuk menyelidiki potensi hubungan antar label. Rajpur et al., mengembangkan algoritma untuk pembelajaran mendalam bernama *CheXNet*, yang menjadi lebih mudah dikelola dengan penerapan normalisasi batch dan koneksi padat. Berdasarkan permasalahan yang disebutkan sebelumnya, penulis memutuskan untuk meningkatkan adopsi pembelajaran mendalam untuk mendeteksi penyakit paru-paru dengan beberapa skenario termasuk:

1. Tingkatkan kualitas gambar x-ray menggunakan aplikasi peningkatan *white balance*.
2. Adopsi peningkatan kontras menggunakan CLAHE untuk menangani masalah analisa gambar medis pada gambar.
3. Penerapan model *deep learning* menggunakan *EfficientNet* bertujuan untuk mengurangi kesalahan klasifikasi penyakit paru-paru.

II. TINJAUAN PUSTAKA

CNN, atau *Convolutional Neural Networks*, kini menjadi salah satu pendekatan paling efektif yang digunakan untuk analisis gambar medis. Mereka menunjukkan efisiensi luar biasa dalam mengidentifikasi gambar secara akurat. Bagian

selanjutnya akan membahas beberapa model CNN yang ada, termasuk model Pra-Terlatih, Praktis, dan *Sequencing*.

Liu dkk. menyarankan tiga kategori berbeda untuk pemanfaatan *CNN-trained algorithm* untuk deteksi TB. Dalam ketiga metode ini, arsitektur CNN bertanggung jawab untuk mengekstraksi fitur, yang kemudian dikembangkan menggunakan *support vector machine* (SVM) [7]. Pada saran selanjutnya, karakteristik dikumpulkan menggunakan *Coreference Resolution* (CR) dan dipelajari menggunakan algoritma SVM. Pendekatan ketiga melibatkan penggabungan dua gagasan ini untuk membentuk suatu kelompok peserta didik. Kumpulan data *Montgomery* berisi total 138 gambar sinar-X, sedangkan *database Shenzhen* berisi 662 gambar sinar-X. Model yang dikembangkan ini efektif untuk mengurangi waktu pemrosesan, namun memiliki akurasi yang terbatas, sehingga tidak cocok untuk digunakan dalam diagnosis klinis.

Penelitian yang dilakukan oleh Hanafi dkk [3] telah menghasilkan hasil klasifikasi yang efektif dengan tingkat akurasi mencapai 0.98. Penelitian ini mengidentifikasi paru-paru dari pasien yang terinfeksi COVID-19 dan kontrol normal. Untuk mendapatkan klasifikasi yang efektif mereka mengkombinasikan dua metode yaitu *convolutional neural networks* (CNNs) dan *autoencoders* (AEs). A CNN merupakan algoritma *machine learning* yang optimal untuk pemrosesan gambar dan merupakan cikal bakal dari *deep learning*. Selain itu, AE memiliki kemampuan untuk mengekstrak fitur melalui algoritma *machine learning*. Kombinasi dari kedua metode tersebut merupakan kunci keberhasilan kinerja model CAE dan akan disebut sebagai algoritma yang dihasilkan dalam penelitian ini.

Penggunaan teknologi CNN merupakan hal yang umum. Penelitian ini memiliki 3 kategori: normal, COVID-19, dan pneumonia. Pada penelitian ini, mereka menggunakan beberapa dataset yang didapatkan dari sumber yang berbeda. Beberapa algoritma modern *deep learning* seperti VGG, *MobileNet*, *RestNet*, *Inception* dan *EfficientNet* digunakan. Mereka menemukan bahwa penggunaan CNN secara tradisional memiliki kinerja yang baik bahkan dibandingkan dengan beberapa algoritma yang dirancang khusus, seperti VGG16 [3] [8] [9].

Penggunaan teknologi *deep learning* telah dilakukan oleh para peneliti sebelumnya [10]. Ada tiga kelas yang digunakan yaitu COVID-19, pneumonia, dan normal. Algoritma CNN digunakan untuk mengklasifikasikan ketiga penyakit yang telah disebutkan. Para peneliti mempertimbangkan pengujian dengan banyak kelas dan kelas biner. Hasil evaluasi menunjukkan akurasi mencapai 98%, presisi 97,3%, *recall* 99,50%, dan F1-measure 98,3%. Akurasi metode *biner* adalah 99,37% dan 98,7% untuk *recall*, 100% untuk *recall*, dan 99,3% untuk *F1-measure*. Model ini sangat efektif dalam mendeteksi COVID-19. Hal ini sangat dipengaruhi oleh penggunaan model terlatih yang dibangun menggunakan VGG16. Hal ini sangat berguna dalam membantu tim medis pada tahap pelaksanaan di rumah sakit atau klinik kesehatan.

Model hybrid yang menggunakan teknologi CNN dan SVM telah digunakan untuk klasifikasi COVID-19 dalam

penelitian yang dilakukan oleh Ayalev dkk. [11]. Hasil evaluasi model ini termasuk kuat dengan akurasi sebesar 99,8% pada tes pelatihan, sedangkan pada tes evaluasi akurasinya mencapai 99,1%. Hasil ini tidak terlepas dari proses *preprocessing* yaitu proses *augmentasi*, proses *cropping*, dan proses segmentasi. Kombinasi CNN dan SVM sangat cocok. Namun penelitian tersebut hanya menggunakan klasifikasi *biner* yaitu COVID-19 dan normal. Secara umum, uji coba pada kelas *biner* memiliki performa yang lebih baik dibandingkan uji coba pada banyak kelas.

Kualitas dataset untuk tugas penggambaran medis terutama penyakit paru-paru menjadi isu penting. Permasalahan tersebut antara lain kualitas dataset gambar, perbedaan jumlah tiap kelas sehingga menyebabkan dataset tidak seimbang, jumlah dataset kurang dan data biasa. Beberapa peneliti mencoba menangani masalah tersebut [12] [13]. Penerapan CLAHE bertujuan untuk meningkatkan kualitas peningkatan kontras pada gambar paru. Pada penelitian ini, peningkatan kualitas gambar mempengaruhi efektivitas penyakit paru-paru. Masalah lain dalam pendeteksian penyakit paru-paru adalah kumpulan data yang tidak seimbang. Untuk mengatasi masalah tersebut, sebuah penelitian mengadopsi GAN untuk menghasilkan dataset imitasi. Adopsi informasi sampingan dari informasi medis diadopsi oleh beberapa peneliti dengan menggunakan informasi suara [14]. Model adopsi multimodal melibatkan GAN, VGG dan *RestNet* [14]. Penerapan informasi sampingan dapat meningkatkan kinerja deteksi penyakit paru-paru. Namun, hasil deteksi paru memerlukan biaya komputasi yang tinggi, masih menghadapi kesalahan deteksi dan pengumpulan informasi sampingan dari informasi medis sangat sulit.

Strategi penting untuk mengurangi kesalahan klasifikasi dilakukan sebagian besar peneliti dengan model *deep learning* yang terintegrasi atau *hybrid*. Misalnya model *hybrid* menggunakan *Multi Layer Perceptron* (MLP) dan Bi-LSTM [15]. Penelitian tersebut berhasil meningkatkan efektivitas penyakit paru-paru dengan uji akurasi sebesar 0,9811. Pencapaian MLP dan Bi-LSTM yang luar biasa harus dibayar dengan biaya komputasi yang tinggi. Kekurangan model ini masih menghadapi kesalahan deteksi mendekati 2%. Bertujuan untuk meningkatkan kinerja akurasi, sebuah penelitian mencoba memodifikasi lapisan model CNN [16]. Hasilnya menunjukkan bahwa modifikasi CNN memperoleh uji akurasi sebesar 0,9496. Modifikasi pendekatan lapisan CNN memberikan hasil yang menjanjikan untuk kedepannya. Memungkinkan untuk mengurangi biaya komputasi yang tinggi dibandingkan dengan mekanisme *deep learning hybrid*. Memang, penelitian ini menghadapi kekurangan hasil kesalahan klasifikasi.

III. METODE PENELITIAN

A. Datasets

Penelitian ini melibatkan 3 kelas dataset yaitu Normal, Covid, dan Virus. Kumpulan data diambil dari data publik *Kaggle* dataset *Covid CXR Image* yang dikumpulkan oleh

Manu Siddharta pada tahun 2021 [17]. Karakteristik detail dataset dapat dilihat pada Tabel 1 di bawah ini.

Tabel 1. *Detail Lung Diseases Datasets*

Classes	Total	Training	Testing	Validation
Normal	689	551	69	69
COVID	536	429	53	53
Virus	626	502	62	62

B. Pre-processing

Dataset gambar yang dikumpulkan dari repositori publik yang terletak pada kualitas buruk meliputi ukuran gambar tidak standar, kontras buruk, dan lokasi fokus gambar paru-paru. Kami mempertimbangkan untuk meningkatkan kualitas gambar dengan beberapa *Pre-processing* umum yang mencakup, *zoom range* dengan parameter nilai 0.2, nilai parameter *shear range* 0.2, parameter *horizontal flip* dengan nilai *True*, dan nilai dari parameter *rescale* ialah 1./255. Kami juga melakukan peningkatan *Pre-processing* dengan *opencv*. Detail *Pre-processing* yang ditingkatkan dijelaskan pada bagian di bawah.

CLAHE, juga dikenal sebagai *Contrast Limited Adaptive Histogram Equalization*, berfungsi sebagai metode yang sangat efektif untuk meningkatkan kecerahan grafik. Ini merupakan peningkatan yang lebih baik dibandingkan *Adaptive Histogram Equalization* (AHE) konvensional. Ekualisasi histogram merupakan teknik langsung yang digunakan untuk meningkatkan kontras visual dengan menggeser spektrum kecerahan atau memperbesar tingkat kecerahan yang sering terjadi. Hal ini, dapat dilakukan untuk mengurangi hasil noise pada gambar sehingga gambar akan terlihat lebih jelas.

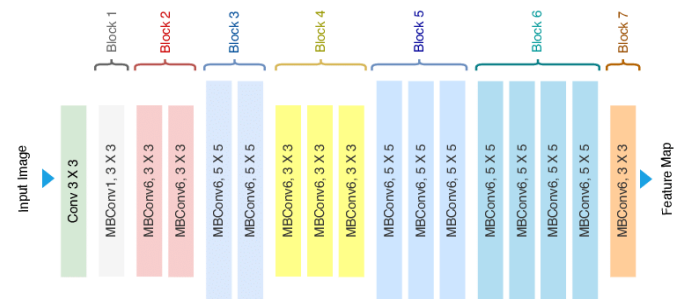
White balancing adalah metode yang digunakan untuk memperbaiki foto yang memiliki latar belakang gelap. Masalah ini biasanya muncul dalam pemindaian radiologi medis, dimana peralatan fotografi mengalami kesulitan dalam menangkap kecerahan yang cukup, sehingga menghasilkan gambar yang lebih gelap. Alat *white balance* memperbaiki masalah ini dengan secara mandiri mengubah corak lapisan saat ini dalam foto, yaitu memanjangkan masing-masing jalur merah, biru, dan hijau. Hasil dari *white balance* sendiri yakni menyeimbangkan warna gambar, agar setiap sudut warna pada gambar terlihat sama.

C. Pembelajaran Mendalam Dengan EfficientNet

EfficientNet adalah kumpulan model CNN yang dirancang khusus untuk mencapai tingkat keandalan yang lebih tinggi sekaligus meminimalkan penggunaan sumber daya operasional. Algoritma tersebut diciptakan oleh para ilmuwan di divisi kecerdasan buatan Google. *EfficientNet* memperkenalkan teknik baru yang disebut penskalaan gabungan, yang melibatkan cara sistematis untuk meningkatkan ukuran CNN.

Pendekatan konvensional dalam penskalaan CNN biasanya memprioritaskan perluasan kedalaman jaringan (dengan memasukkan tingkatan tambahan), lebar (dengan

memasukkan entitas tambahan dalam tingkatan), atau meningkatkan detail visual yang disediakan. Namun demikian, penggunaan teknik penskalaan yang sederhana dapat mengakibatkan efisiensi yang tidak memadai sebagai konsekuensi dari kenaikan biaya operasional dan kesulitan yang terkait dengan mempelajari lebih banyak jaringan. Gambar 1 merupakan desain arsitektur dalam penelitian ini.



Gambar 1. Arsitektur Kerangka *EfficientNet*

D. Evaluations Metrics

Untuk menilai keakuratan pendekatan yang disarankan, beberapa indikator kinerja dipilih, seperti perolehan, presisi, *F1-Score*, akurasi, dan area tipikal di bawah kurva (AUC). Pertama, *matrix of confusion* dibuat menggunakan prediksi yang diberikan oleh kerangka yang disarankan untuk menilai keakuratan klasifikasi, membedakan antara yang benar dan yang salah. Matriks ini terdiri dari empat komponen penting: *True Positives* (TP), *True Negatives* (TN), *False Positives* (FP), dan *False Negatives* (FN). TP mewakili identifikasi kasus penyakit yang tepat berdasarkan model, sedangkan TN mewakili klasifikasi kasus normal yang tepat. FP menunjukkan kejadian ketika pasien normal salah dikategorikan sebagai sakit, sedangkan FN mewakili kasus penyakit yang salah diklasifikasikan sebagai normal.

$$\text{Accuracy test} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Accuracy merupakan nilai perbandingan yang diperoleh dari data yang diprediksi benar terhadap jumlah seluruh data.

$$\text{Recall test} = \frac{TP}{TP+FN} \quad (2)$$

Recall merupakan hasil dari nilai total data positif yang dibagi dengan data yang benar dan data yang salah.

$$\text{Precision test} = \frac{TP}{TP+FP} \quad (3)$$

Precision merupakan perbandingan dari nilai total data positif yang dibagi dengan data yang benar.

$$\text{F1 - measure test} = \frac{2*TP(\text{True Positive})}{(2*TP)+FP+FN} \quad (4)$$

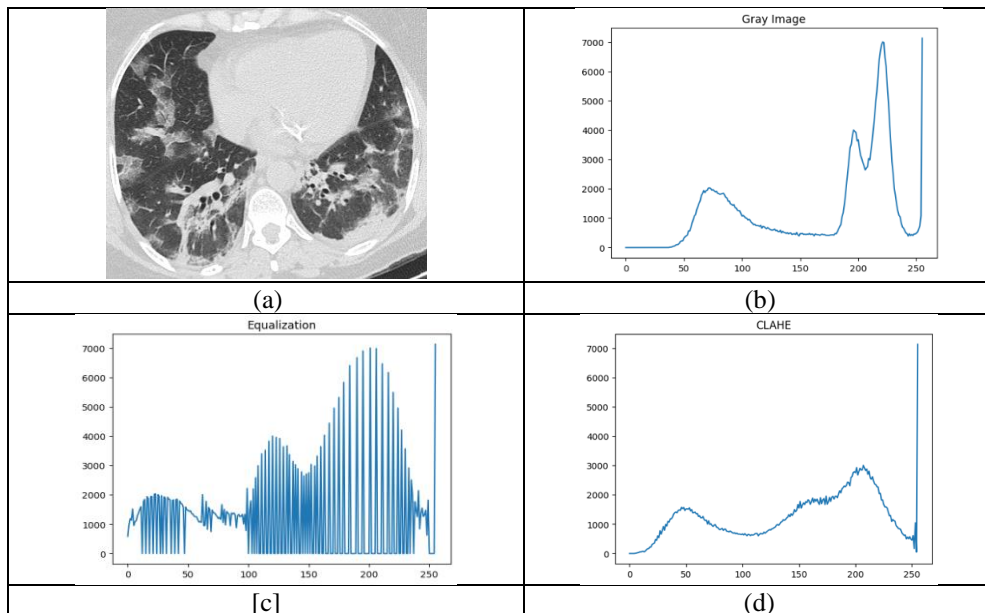
F1-Score merupakan jumlah keseluruhan dari gabungan recall dan presisi.

IV. HASIL DAN DISKUSI

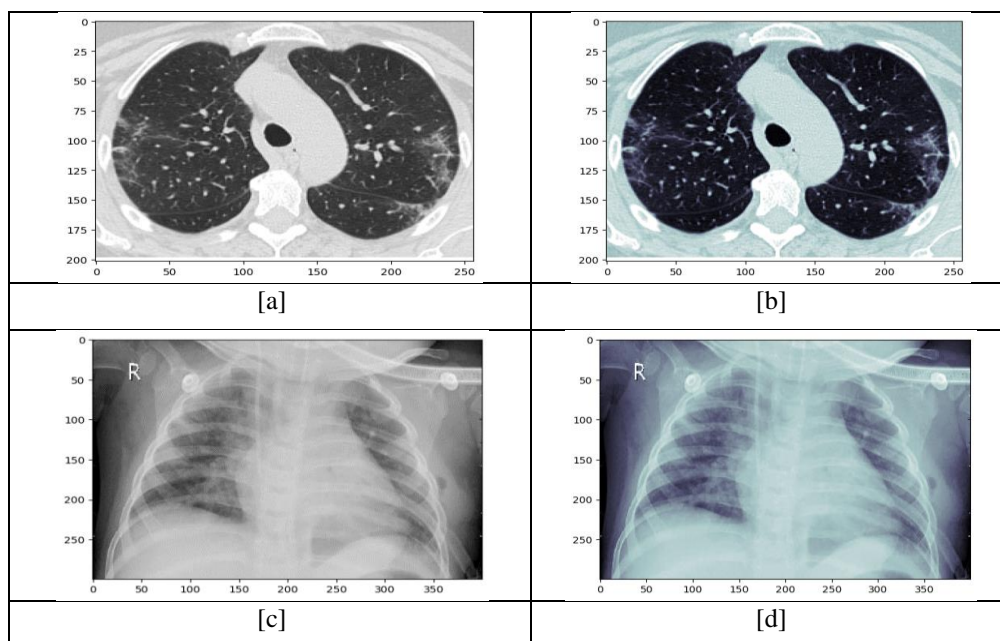
Dalam tahap ini penulis melakukan beberapa proses percobaan. Proses ini bertujuan untuk mencapai beberapa tujuan termasuk kualitas gambar, mengubah ukuran gambar, memperbesar gambar, *white balance*, dan peningkatan kontras. Hasil dari beberapa tahap *Pre-processing* dapat dilihat pada bagian di bawah ini. Peningkatan *white balance* dan kontras bertujuan untuk meningkatkan kinerja mesin *classifier*. Gambar 2 di bawah menunjukkan hasil peningkatan kontras. Gambar tersebut menjelaskan transformasi histogram dari histogram asli yang diambil dari gambar asli (lihat

Gambar b), setelah dilakukan proses *equalizer histogram*, yang berfungsi untuk melakukan perbaikan dengan meratakan histogram dari sebuah citra. Maka hasilnya dapat dilihat pada Gambar c, dan Terakhir, hasil dari peningkatan kontras menggunakan CLAHE dapat dilihat pada Gambar d.

Gambar sebenarnya dari *white balance* dan peningkatan kontras ditunjukkan pada Gambar 3 di bawah. Sebelah kanan adalah gambar asli dari *dataset* gambar, sedangkan sebelah kiri adalah gambar hasil peningkatan kontras menggunakan CLAHE. Terlihat bahwa hasil dari proses CLAHE, gambar tampak lebih jernih dan kontras.



Gambar 2. Output Histogram White Balance dan Peningkatan Kontras

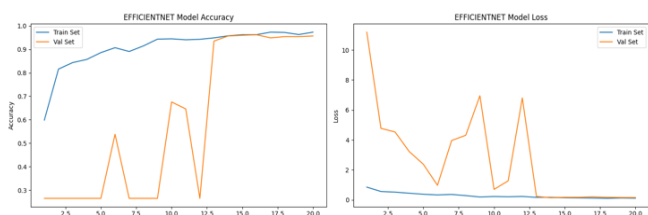


Gambar 3. Output Gambar Peningkatan White Balance dan Kontras

Penelitian ini mengevaluasi menggunakan beberapa pendekatan matrik seperti grafik pengujian akurasi, *Confusion Matrix* dan uji kurva AUC untuk melakukan pengamatan pada 20 epoch, 50 epoch dan 100 epoch.

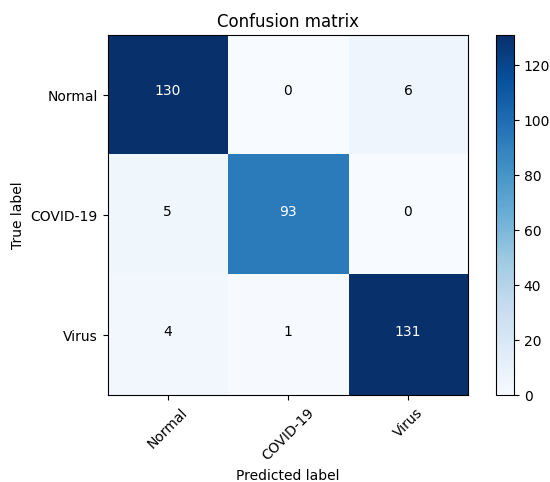
Pengujian akurasi dan *loss* mempertimbangkan untuk mengamati dampak masing-masing epoch. Sedangkan *confusion matrix* digunakan untuk memprediksi hasil menyeluruh dari data yang digunakan, dengan mengklasifikasi kelas yang benar dan salah. Hasil model klasifikasi untuk deteksi penyakit paru dapat dianalisis dengan Kurva AUC. Kurva ROC dengan nilai sekitar 1,00 dapat disimpulkan sebagai model yang memadai untuk diadopsi dalam aplikasi nyata diagnosis medis.

Representasi hasil pelatihan selama 20 epoch ditunjukkan pada Gambar 4. Uji akurasi dan kerugian menunjukkan hasil yang tidak berhasil dengan nilai yang berbeda antara garis akurasi dan garis validasi, serta uji kerugian dan uji validasi.



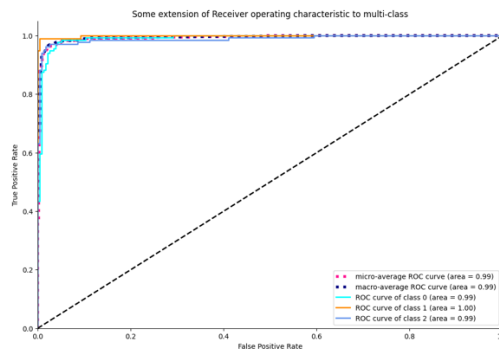
Gambar 4. Proses Pengujian Dengan 20 Epoch

Hasil uji *confusion matrix* menunjukkan bahwa uji klasifikasi mencapai hasil efektivitas yang lebih baik. Hanya sejumlah kecil deteksi kesalahan antar kelas. Yang dimana nilai hasil untuk Normal 130, Covid-19 93, dan Virus 131. Detail pengujian matriks konfusi ditunjukkan pada Gambar 5.



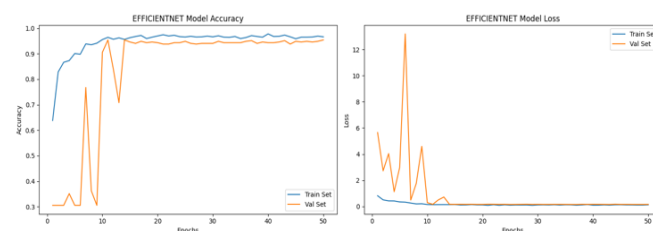
Gambar 5. Uji Matriks Konfusi Dengan 20 Epoch

Hasil dari kurva AUC 20 epoch menghasilkan nilai klasifikasi untuk kelas normal 0,99, kelas covid 1,00, dan kelas virus 0,99. Menunjukkan hasil yang dikatakan sempurna, dimana setiap kelas mencapai di atas 0,99. Hasil kurva AUC untuk 20 epoch ditunjukkan pada Gambar 6.



Gambar 6. Uji Kurva AUC Pada 20 Epoch

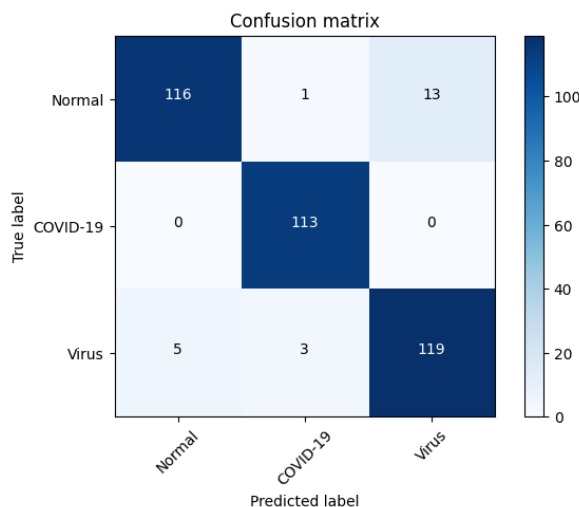
Grafik pengujian akurasi dan *loss* 50 epoch berbeda dengan hasil 20 epoch. Untuk 50 epoch mencapai hasil yang sempurna dimana garis akurasi, *loss* dan validasi memiliki nilai yang sama seperti terlihat pada Gambar 7.



Gambar 7. Proses Pengujian Dengan 50 Epoch

Hasil uji *confusion matrix* menunjukkan bahwa uji klasifikasi mencapai hasil efektivitas yang lebih baik. Hanya sejumlah kecil deteksi kesalahan antar kelas. Detail uji konfusi matriks ditunjukkan pada Gambar 8.

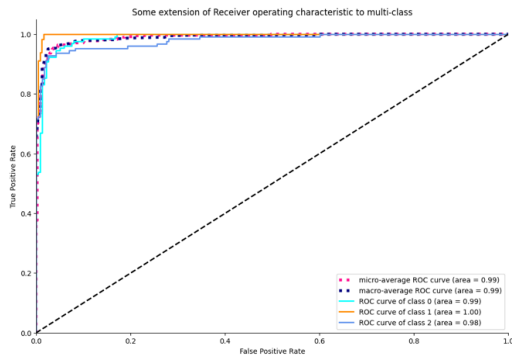
Deteksi kesalahan menunjukkan bahwa 6 kelas normal terdeteksi sebagai kelas virus, 5 kelas Covid terdeteksi sebagai kelas normal.



Gambar 8. Uji Matriks Konfusi Dengan 50 Epoch

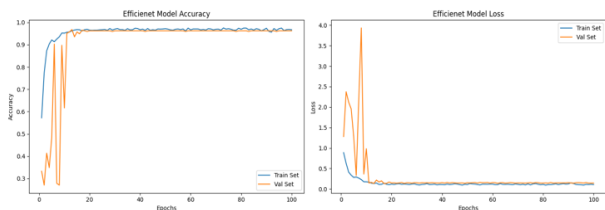
Kurva AUC selama 50 epoch menunjukkan model tugas klasifikasi penyakit paru mencapai hasil yang baik dengan

pencapaian setiap kelas di atas 0,99 yaitu kelas normal 0,99, kelas covid 1,00 dan untuk kelas virus 0,98.



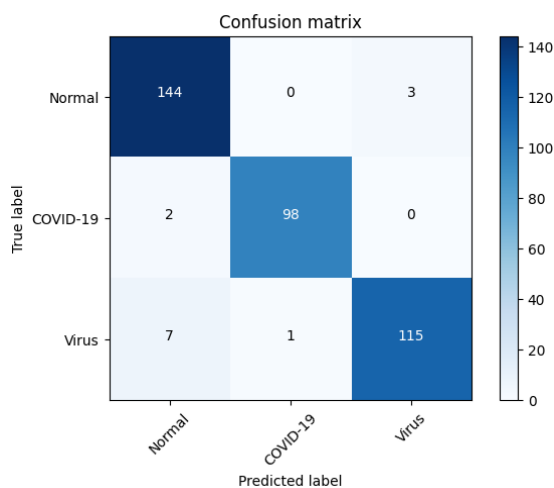
Gambar 9. Pengujian Kurva AUC Pada 50 Epoch

Hasil grafik yang hampir sempurna juga dicapai selama 100 epoch. Dimana hasil pelatihan mencapai kinerja yang lebih baik selama 20 epoch. Terlihat grafik akurasi dan loss test hampir mirip dengan garis validasi. Pelatihan pada 100 epoch ditunjukkan pada Gambar 10 di bawah ini.



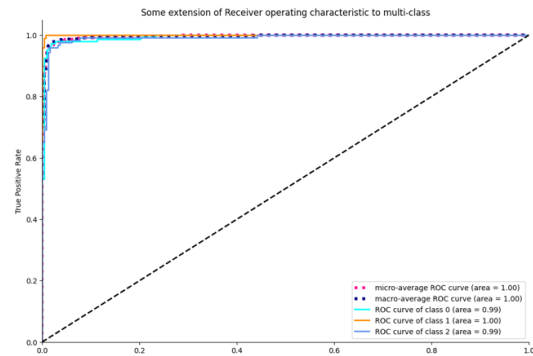
Gambar 10. Proses Pengujian Dengan 100 Epoch

Hasil confusion matrix menunjukkan bahwa uji klasifikasi mencapai hasil efektivitas yang lebih baik pada 100 epoch. Dimana hanya sejumlah kecil deteksi kesalahan antar kelas. Detail pengujian matriks konfusi ditunjukkan pada Gambar 11. Deteksi kesalahan mayoritas terletak pada kelas virus yang terdeteksi pada kelas normal.



Gambar 11. Uji Matriks Konfusi Dengan 100 Epoch

Hasil kurva AUC untuk 100 epoch yang ditunjukkan pada Gambar 12 juga menunjukkan hasil yang sempurna dan hampir sama dengan hasil dari 20 epoch, dimana setiap kelas mencapai diatas 0,99 dengan nilai kelas normal 0,99, kelas covid 1,00 dan kelas virus mencapai 0,99.



Gambar 12. Uji Kurva AUC Pada 100 Epoch

Precision berfungsi untuk mengukur model dengan prediksi positif yang salah. Recall berfungsi untuk mengukur model dengan hasil yang positif. F1-Score memiliki fungsi keseimbangan antara precision dan recall.

Matriks evaluasi pada setiap epoch ditampilkan pada tabel di bawah ini. Tabel 2 tersebut menunjukkan hasil akurasi pada epoch 20 rata-rata sebesar 0,96. Hasil nilai tertinggi terdapat pada precision covid yakni sebesar 0,99 dan yang terendah 0,94.

Tabel 2. Hasil Matriks Evaluasi Pada 20 Epoch

Classes	Accuracy	Precision	Recall	F1
Normal		0,94	0,96	0,95
Covid	0,96	0,99	0,95	0,97
Virus		0,96	0,96	0,96

Tabel 3 menunjukkan hasil akurasi pada epoch 50 rata-rata sebesar 0,95. untuk rata-rata tertinggi diraih oleh hasil dari recall covid dengan nilai 0,99 dan yang terendah 0,91. Tidak ada hasil yang signifikan antara 20 dan 50 epoch.

Tabel 3. Hasil Matriks Evaluasi Pada 50 Epoch

Classes	Accuracy	Precision	Recall	F1
Normal		0,93	0,95	0,94
Covid	0,95	0,97	0,99	0,98
Virus		0,95	0,91	0,93

Tabel 4 menunjukkan hasil akurasi pada epoch 100 rata-rata sebesar 0,97. Dengan capaian rata-rata tertinggi terdapat di precision covid nilai 0,99 dan yang terendah terdapat pada hasil recall virus yakni 0,93. Terdapat hasil yang signifikan antara 20 epoch, 50 epoch dibandingkan 100 epoch. Jumlah epoch mempengaruhi kinerja yang signifikan dalam efektivitas klasifikasi.

Tabel 4. Hasil Matriks Evaluasi Pada 100 Epoch

Classes	accuracy	precision	recall	F1
Normal		0,94	0,98	0,96
Covid	0,97	0,99	0,98	0,98
Virus		0,97	0,93	0,95

V. KESIMPULAN

Berdasarkan uji coba penelitian ini, hasil terbaik yang dicapai dengan akurasi tertinggi terdapat pada pengujian 100 epoch dengan capaian akurasi sebesar 0,97. Penerapan model ini bertujuan untuk meningkatkan keberhasilan deteksi penyakit paru-paru. Algoritma mencapai kinerja yang lebih baik berdasarkan hasil evaluasi. Tingkat efektivitas model ini dengan menerapkan beberapa peningkatan proses termasuk *Pre-processing*, *White balance*, peningkatan kontras, dan pembelajaran mendalam modern dengan *EfficientNet*. Jumlah epoch juga memiliki peran penting dalam meningkatkan efektivitas. Peningkatan kontras dan *White balance* berhasil memecahkan masalah dalam gambar medis dimana masalah yang umum terjadi adalah kontras, tidak ada gambar x-ray standar. Kami yakin peningkatan model ini disebabkan oleh 3 aspek penting yang disebutkan di atas.

REFERENSI

- [1] S. Kumar, O. Ivanova, A. Melyokhin, and P. Tiwari, "Deep-learning-enabled multimodal data fusion for lung disease classification," *Inform Med Unlocked*, vol. 42, Jan. 2023, doi: 10.1016/j.imu.2023.101367.
- [2] M. Jasmine Pemeena Priyadarsini *et al.*, "Lung Diseases Detection Using Various Deep Learning Algorithms," *J Healthc Eng*, vol. 2023, 2023, doi: 10.1155/2023/3563696.
- [3] Hanafi, A. Pranolo, and Y. Mao, "Cae-covidx: Automatic covid-19 disease detection based on x-ray images using enhanced deep convolutional and autoencoder," *International Journal of Advances in Intelligent Informatics*, vol. 7, no. 1, pp. 49–62, 2021, doi: 10.26555/ijain.v7i1.577.
- [4] Hanafi and B. M. Aboobaider, "Word Sequential Using Deep LSTM and Matrix Factorization to Handle Rating Sparse Data for E-Commerce Recommender System," *Comput Intell Neurosci*, vol. 2021, no. 1, 2021, doi: https://doi.org/10.1155/2021/8751173 Research.
- [5] Hanafi, N. Suryana, and A. S. B. H. Basari, "Convolutional-NN and word embedding for making an effective product recommendation based on enhanced contextual understanding of a product review," *Int J Adv Sci Eng Inf Technol*, vol. 9, no. 3, 2019, doi: 10.18517/ijaseit.9.3.8843.
- [6] Hanafi, N. Suryana, and A. Samad, "Dynamic convolutional neural network for eliminating item sparse data on recommender system," *IJAIN*, vol. 4, no. 3, pp. 226–237, 2018.
- [7] V. Ravi, H. Narasimhan, C. Chakraborty, and T. D. Pham, "Deep learning-based meta-classifier approach for COVID-19 classification using CT scan and chest X-ray images," *Multimed Syst*, vol. 28, no. 4, pp. 1401–1415, 2022.
- [8] S. T. H. Kieu, A. Bade, M. H. A. Hijazi, and H. Kolivand, "A survey of deep learning for lung disease detection on medical images: state-of-the-art, taxonomy, issues and future directions," *J Imaging*, vol. 6, no. 12, p. 131, 2020.
- [9] Z. A. Varzaneh, A. Orooji, L. Erfannia, and M. Shanbehzadeh, "A new COVID-19 intubation prediction strategy using an intelligent feature selection and K-NN method," *Inform Med Unlocked*, vol. 28, p. 100825, 2022, doi: 10.1016/j.imu.2021.100825.
- [10] H. Malik and T. Anees, "BDCNet: Multi-classification convolutional neural network model for classification of COVID-19, pneumonia, and lung cancer from chest radiographs," *Multimed Syst*, vol. 28, no. 3, pp. 815–829, 2022.
- [11] A. M. Ayalew, A. O. Salau, Y. Tamyalew, B. T. Abeje, and N. Woreta, "X-Ray image-based COVID-19 detection using deep learning," *Multimed Tools Appl*, pp. 44507–44525, 2023, doi: 10.1007/s11042-023-15389-8.
- [12] R. Prince, Z. Niu, Z. Y. Khan, M. Emmanuel, and N. Patrick, "COVID-19 detection from chest X-ray images using CLAHE-YCrCb, LBP, and machine learning algorithms," *BMC Bioinformatics*, vol. 25, no. 1, pp. 1–25, 2024, doi: 10.1186/s12859-023-05427-5.
- [13] R. Gulakala, B. Markert, and M. Stoffel, "Rapid diagnosis of Covid-19 infections by a progressively growing GAN and CNN optimisation," *Comput Methods Programs Biomed*, vol. 229, p. 107262, 2023.
- [14] A. S. Althenayan *et al.*, "COVID-19 Hierarchical Classification Using a Deep Learning Multi-Modal," *Sensors*, vol. 24, no. 8, 2024, doi: 10.3390/s24082641.
- [15] Y. Liu, W. Xing, M. Zhao, and M. Lin, "An end-to-end framework for diagnosing COVID-19 pneumonia via Parallel Recursive MLP module and Bi-LTSM correlation,"
- [16] 2023M. Nishio *et al.*, "Deep learning model for the automatic classification of COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy: a multi-center retrospective study," *Sci Rep*, vol. 12, no. 1, p. 8214, 2022.
- [17] M. Siddharta, "Covid CXR Image | Kaggle," [Online]. Available: <https://www.kaggle.com/datasets/sid321axn/covid-cxr-image-dataset-research/discussion?sort=undefined> [Accessed April 2024]