

Klasifikasi Rontgen Citra Paru untuk COVID-19 Menggunakan Convolutional Neural Network Berbasis Arsitektur EfficientNetV2

Hutomo Try Wibowo¹

¹Program Studi Manajemen Sistem Informasi, Universitas Gunadarma, Indonesia.

Artikel Info

Kata Kunci:

Convolutional Neural Network;
Covid-19;
EfficientNetV2;
Rontgen Dada.

Keywords:

CNN;
COVID-19;
EfficientNetV2;
Chest X-Ray.

Riwayat Artikel:

Submitted: 6 November 2023
Accepted: 23 Januari 2024
Published: 24 Februari 2024

Abstrak: Pada 30 Januari 2020, WHO mengumumkan wabah COVID-19 sebagai Darurat Kesehatan Masyarakat Internasional, memicu kebutuhan akan pendekatan cepat dalam diagnosis dan penilaian penyakit paru-paru. Meskipun pengamatan citra rontgen paru-paru menjadi metode yang menjanjikan, sistem penilaian manual memakan waktu, mendorong perlunya integrasi teknologi. Dalam konteks ini, perkembangan pesat visi komputer dan pembelajaran mesin memungkinkan pengembangan perangkat lunak untuk mengklasifikasikan citra rontgen secara efisien. Sejumlah penelitian telah mencatat keberhasilan berbagai arsitektur Convolutional Neural Networks (CNN) dalam pengklasifikasian penyakit paru-paru dengan hasil akurasi yang mengesankan. Studi ini memfokuskan pada eksplorasi arsitektur CNN, khususnya EfficientNetV2, dalam mengklasifikasikan citra rontgen paru ke dalam empat kelas: COVID-19, opasitas paru, Normal, dan pneumonia viral. Pada penelitian ini, proses pelatihan model menggunakan dataset COVID-19 Radiography dengan menerapkan teknik pensampelan ulang dan augmentasi beserta skema pembagian data sebagai teknik pemrosesan data, penggunaan metrik evaluasi seperti akurasi, perolehan, presisi, dan F1 juga dilakukan untuk mendapatkan model klasifikasi yang paling optimal. Hasil dari penelitian ini menunjukkan nilai metrics tertinggi yang didapatkan pada 97,05% untuk akurasi, 97,38% untuk presisi, 96,88% untuk perolehan, dan 97,13% untuk F1.

Abstract: On January 30, 2020, WHO declared the COVID-19 outbreak an International Public Health Emergency, triggering the need for rapid approaches in the diagnosis and assessment of lung disease. Although lung x-ray image observation is a promising method, manual scoring systems are time-consuming, prompting the need for technology integration. In this context, the rapid development of computer vision and machine learning allows the development of software to efficiently classify x-ray images. A number of studies have noted the success of various Convolutional Neural Networks (CNN) architectures in lung disease classification with impressive accuracy results. This study focuses on exploring CNN architectures, especially EfficientNetV2, in classifying lung x-ray images into four classes: COVID-19, lung opacity, Normal, and viral pneumonia. In this research, the model training process uses the COVID-19 Radiography dataset by applying resampling and augmentation techniques along with data split schemes as data processing techniques, the use of evaluation metrics such as accuracy, precision, recall and F1 is also carried out to obtain the most optimal classification model. The results of this research show that the highest metric values were obtained at 97.05% for accuracy, 97.38% for precision, 96.88% for retrieval, and 97.13% for F1.

Corresponding Author:

Hutomo Try Wibowo

Email: hutomogundar@gmail.com

PENDAHULUAN

Pada tanggal 30 Januari 2020, Organisasi Kesehatan Dunia (WHO) mengumumkan kemunculan wabah virus corona baru sebagai “Darurat Kesehatan Masyarakat yang Menjadi Perhatian Internasional”. Penunjukan ini menandakan krisis kesehatan global yang memerlukan perhatian semua negara, khususnya Indonesia (Arriani dkk., 2020; Luthfi & Septiyanti, 2023). Virus COVID-19 berpotensi menginfeksi dan membahayakan paru-paru manusia (Rousan dkk., 2020; Singhal, 2020), sehingga salah satu pendekatan yang dapat dilakukan untuk mendeteksi infeksi COVID-19 pada pasien adalah melalui pengamatan citra rontgen paru-paru (Yasin & Gouda, 2020). Teknik pencitraan ini tidak hanya berfungsi untuk diagnosis tetapi juga untuk mengevaluasi tingkat keparahan penyakit. Namun, sistem penilaian manual yang dilakukan dalam mengklasifikasikan penyakit paru-paru (COVID-19, opasitas paru, dan pneumonia viral) memakan waktu dan dapat memperpanjang proses diagnosis pasien tersebut (Warren dkk., 2018).

Proses penilaian secara manual yang memakan waktu dapat diselesaikan dengan bantuan teknologi, dikarenakan pesatnya perkembangan teknologi informasi baik itu dalam visi komputer, pembelajaran mesin, dan teknologi informasi. Perangkat lunak yang mampu mengklasifikasikan citra masukan kedalam kelas-kelasnya memungkinkan untuk diciptakan (Abadi dkk., 2005; Albawi dkk., 2018; Antonio Gulli & Sujit Pal, 2017; Lee & Song, 2019; Ma & Lu, 2017; Putra dkk., 2023). Kemajuan ini membuka pintu bagi pengembangan perangkat lunak yang dapat mengklasifikasikan jenis penyakit secara efektif (Chowdhury dkk., 2020), termasuk COVID-19, opasitas paru-paru, dan pneumonia viral berdasarkan citra sinar-X. Beberapa penelitian terdahulu terkait penerapan Convolutional Neural Networks (CNN) dalam mengklasifikasikan citra sinar-X. Pemanfaatan teknik peningkatan citra dan penggunaan berbagai jenis arsitektur CNN termasuk InceptionV3, DenseNet201, ChexNet (Rahman, Khandakar, dkk., 2021), pelatihan dilakukan menggunakan dataset COVQU dengan akurasi paling besar yang didapatkan sebesar 96,26%. Studi dari (Ozturk dkk., 2020) mengembangkan arsitektur DarkCovidNet untuk mengklasifikasikan citra x-ray, dan menggunakan dataset COVID-chestxray untuk melatih model memperoleh hasil akurasi untuk kelas biner (COVID, No-Findings) sebesar 98,08% dan kelas jamak (COVID, Tidak Ada Temuan, pneumonia) sebesar 87,02%. Studi dari (S. Yadav dkk., 2020), penerapan teknik augmentasi dan pemanfaatan arsitektur InceptionV3 serta VGG16, pelatihan dilakukan pada dataset Database Radiografi COVID19 dan memperoleh akurasi tertinggi sebesar 98,84% untuk kelas jamak (normal, COVID-19, pneumonia). Dalam penelitian (Jia, Lam, & Xu, 2021), pengembangan CNN dinamis dengan memanfaatkan arsitektur dasar MobileNetv3_small dan dilatih pada dataset COVIDx, memperoleh hasil akurasi 99,7% untuk 3 kelas (COVID-19, infeksi non-COVID-19, sehat), 99,9 % untuk 4 kelas (COVID-19, pneumonia non-COVID-19, TBC, sehat), dan 99,6% untuk 5 kelas (COVID-19, pneumonia bakterial, pneumonia virus non-COVID, TBC, sehat).

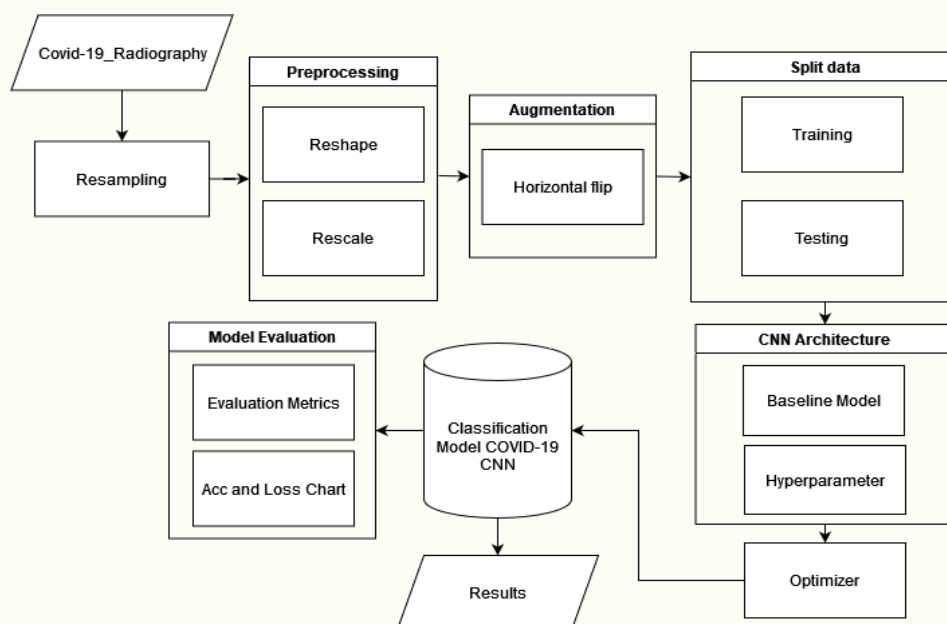
Dalam ruang lingkup arsitektur CNN, EfficientNetV2 (Tan & Le, 2019, 2021) muncul sebagai pilihan yang menarik untuk dieksplorasi, dikarenakan hasil akurasi pengujian ImageNet yang cukup mengesankan dibandingkan dengan arsitektur CNN lainnya. Berdasarkan penelitian tersebut, model EfficientNetV2 dipilih sebagai studi komparatif pada penelitian ini untuk melihat seberapa baik arsitektur ini dapat bersaing dengan arsitektur CNN lainnya khususnya dalam pengklasifikasian citra sinar-X, mengingat hasil akurasi yang didapatkan pada penelitian DarkCovidNet (Ozturk dkk., 2020), VGG, InceptionV3 (S. Yadav dkk., 2020), dan Modified Mobilenetv3 (Jia dkk., 2021) dalam pendeteksian covid-19 yang cukup baik. Diharapkan hasil penelitian ini dapat memberikan gambaran mengenai potensi arsitektur CNN EfficientNetV2 khususnya dalam klasifikasi citra sinar-X serta pemanfaatannya didalam bidang medis.

Penelitian ini mengimplementasikan arsitektur EfficientNetV2 sebagai dasar Convolutional Neural Network (CNN) yang akan digunakan untuk membentuk model klasifikasi terhadap empat

kelas target: COVID-19, opasitas paru, Normal, dan pneumonia viral. Penelitian ini menggunakan kumpulan data citra x-ray yang diperoleh dari Database Radiografi COVID-19 (Rahman, Dr. Chowdhury, dkk., 2021), yang terdiri dari 5845 citra yang dihasilkan dari proses undersampling. Penerapan teknik undersampling bertujuan untuk mencapai keseimbangan data antar kelas (Khaldy, 2018; Zhang dkk., 2019). Selain itu, teknik augmentasi (Perez & Wang, 2017; Sutanty & Kusuma Astuti, 2023; S. S. Yadav & Jadhav, 2019) digunakan untuk memperkenalkan variasi ke dalam data pelatihan, sehingga mengurangi risiko terjadinya *overfitting*. Penggunaan algoritma pengoptimal Adam menjadi pilihan berdasarkan hasil yang didapatkan pada penelitian (Bock dkk., 2018; Kingma & Ba, 2014; Reddi dkk., 2019) dan eva luasi model CNN menggunakan metrik evaluasi (Powers & Ailab, 2020) dilakukan sebagai tolak ukur dari kualitas sebuah model. Hasil yang diharapkan dari penelitian ini adalah pengembangan model klasifikasi yang mampu secara akurat mengkategorikan citra rontgen paru ke dalam empat kelas (COVID-19, opasitas paru, normal, dan pneumonia virus) menggunakan CNN berbasis arsitektur EfficientNetV2, dan studi komperatif mengenai pengaruh metode pelatihan model klasifikasi terhadap berbagai macam kondisi pelatihan dalam menghasilkan model yang optimal.

METODE

Metode penelitian klasifikasi COVID-19 pada citra sinar-X paru terdiri dari beberapa tahapan, dimulai dengan pengumpulan data yang mencakup empat kelas citra rontgen paru: normal, COVID-19, opasitas paru dan pneumonia viral. Untuk mengatasi ketidakseimbangan kelas, dilakukan pengurangan sampel data pada tiap-tiap kelas, untuk memastikan keterwakilan kelas yang adil. Selanjutnya dilakukan pra-pemrosesan, yang melibatkan pembentukan ulang dan penskalaan ulang citra agar sesuai dengan persyaratan dimensi citra masukan pada model CNN. Teknik augmentasi data juga dilakukan karena dinilai secara sistematis dapat meningkatkan keragaman data pada proses pelatihan. Kumpulan data tersebut kemudian dibagi menjadi subset pelatihan dan pengujian, yang dievaluasi berdasarkan rasio yang berbeda (6:4, 7:3, 8:2, 9:1). Arsitektur CNN disiapkan dengan menggabungkan EfficientNetV2_b0 dan mengonfigurasi hyperparameter demi mendapatkan hasil yang optimal. Evaluasi model menggunakan metrik statistik dan analisis grafis, dan hasil akhir adalah sebuah model visualisasi hasil melibatkan penilaian kinerja klasifikasi model pada data tabular dan citra sinar-X. Tahapan tersebut dapat dilihat pada Gambar 1.



Gambar 1. Diagram Flowchart Penelitian

Data rontgen paru-paru yang digunakan merupakan Database Radiografi COVID-19 versi ke-4 yang merupakan dataset publik yang disediakan oleh Kaggle. Dataset terbentuk atas kerjasama Qatar University dan Dhaka University bekerjasama dengan Pakistan dan Malaysia, dataset ini terdiri dari 4 kelas yaitu COVID-19, Normal, Lung Opacity, dan Viral Pneumonia dimana data yang terdapat pada setiap kelas merupakan data rontgen. citra paru-paru dalam format PNG dan resolusi 299 x 299 piksel dengan jumlah data yang terdapat di setiap kelas memiliki rincian, 3615 citra rontgen paru dengan indikasi COVID-19, 10192 citra rontgen normal paru-paru, 6012 citra rontgen Opasitas Paru-paru, dan 1345 citra rontgen paru-paru Pneumonia Viral.

Data yang didapat akan di sampel ulang untuk mendapatkan data pelatihan yang seimbang pada setiap kelasnya. Berdasarkan dataset yang didapatkan, pendekatan *undersampling* dipilih sebagai teknik sampel ulang pada penelitian ini dengan menimbang lamanya waktu pelatihan dan keunikan data citra sinar X yang didapatkan, proses *undersampling* dilakukan dengan mengambil secara acak data dari kelas mayoritas sehingga membentuk kelas baru yang memiliki data sama banyak atau hampir sama dengan kelas minoritas, dengan rincian: 1500 data COVID-19, 1500 data normal, 1500 data opasitas paru, dan 1345 data pneumonia viral.

Dengan dilakukannya proses sampel ulang dan didapatkannya data pelatihan yang seimbang dilanjutkan dengan pra-pemrosesan, proses ini terbagi menjadi 2 yaitu *reshape* dan *rescale* dengan tujuan untuk menyamakan spesifikasi antara citra masukan dengan arsitektur CNN yang digunakan dan meningkatkan stabilitas dan performa pelatihan pada jaringan CNN. Dimensi citra masukan yang diperlukan pada penelitian ini adalah 224 x 224 dan untuk *rescale* poin faktor yang dilakukan pada citra masukan adalah 1./255 sehingga citra sinar masukan yang awalnya bernilai 0-255 yang menunjukkan bahwa citra tersebut adalah RGB menjadi bernilai 0–1 setelah difaktorisasi.

Data yang dihasilkan melalui pra-pemrosesan kemudian diproses dengan teknik augmentasi data, dengan tujuan untuk mendapatkan variasi data pelatihan baru sehingga diharapkan mendapatkan hasil model klasifikasi yang lebih stabil, pada penerapan teknik augmentasi pada penelitian ini akan dilakukan beberapa skenario percobaan dengan mengeksplorasi berbagai macam teknik augmentasi dalam penerapannya, hal ini dilakukan untuk mendapatkan hasil yang paling optimal sehingga penggunaan teknik augmentasi secara linear dapat meningkatkan performa dari model yang didapatkan.

Pembagian data pelatihan merupakan proses yang dilakukan setelah proses augmentasi data, pada penelitian ini akan dilakukan studi komperatif untuk melihat korelasi antara perbandingan data pelatihan dan data validasi terhadap tingkat akurasi yang didapatkan oleh model klasifikasi, skema pelatihan pembagian data dengan perbandingan antara data pelatihan dan data validasi pada penelitian ini diantaranya: (a) 6:4 , (b) 7:3, (c) 8:2, (d) 9:1.

EfficientNetV2_b0 menjadi dasar dari arsitektur model CNN yang dipergunakan pada penelitian ini, EfficientNetV2 merupakan pengembangan dari arsitektur sebelumnya yaitu EfficientNet sehingga memiliki efisiensi parameter yang lebih baik dengan waktu pelatihan yang lebih cepat, hal ini terjadi dikarenakan penggantian struktur jaringan yang sebelumnya MBConv menjadi Fused-MBConv (Tan & Le, 2021). Untuk mendapatkan hasil model yang optimal, dilakukan penambahan jaringan seperti *dropout layer*, *dense layer*. Pemilihan dan penyesuaian parameter algoritma optimasi pelatihan juga nantinya diamati dan di evaluasi, sehingga mendapatkan hasil yang lebih maksimal.

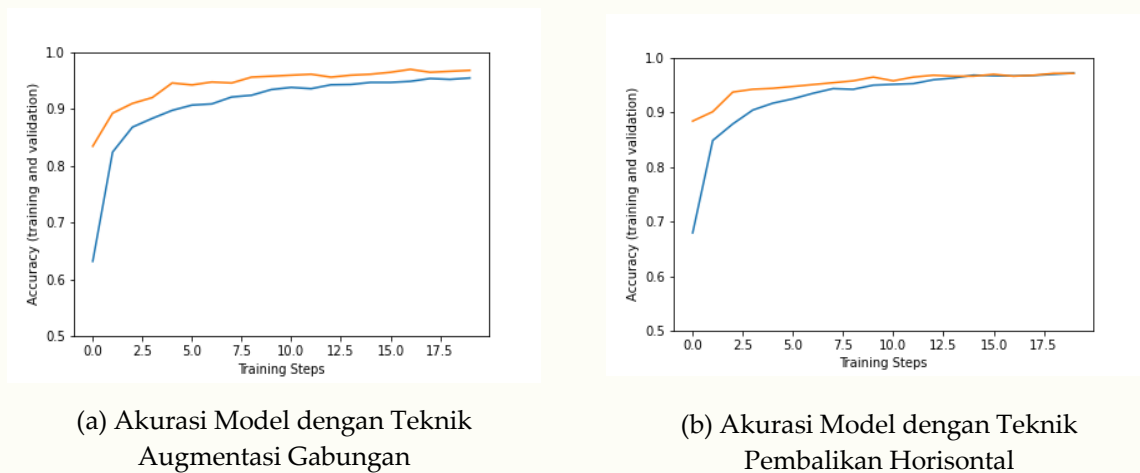
Tahapan evaluasi model, merupakan tahapan yang dilakukan untuk mengevaluasi kualitas dari proses dan hasil pelatihan yang didapatkan terhadap penerapan teknik augmentasi, parameter algoritma optimasasi, dan pembagian data pelatihan. Pada penelitian ini proses evaluasi model terbagi menjadi 2, yang pertama yaitu menggunakan matriks evaluasi yang didapatkan dengan menghitung tingkat akurasi validasi pelatihan model terhadap nilai akurasi, perolehan, presisi, dan juga skor F1 beserta dengan matriks konfusi, dan yang kedua evaluasi berdasarkan grafik akurasi pelatihan dan validasi diantaranya grafik akurasi dan grafik kerugian.

Tahapan akhir yaitu visualiasi prediksi dari model hasil pelatihan, yang bertujuan untuk memvisualisasikan hasil prediksi dari pengujian model yang didapatkan pada penelitian ini. Proses visualisasi dilakukan terhadap 2 skema, yang pertama visualisasi terhadap citra dengan cara memilih

25 citra acak diluar dari data pelatihan yang berfokus dalam melihat bagaimana secara umum model dapat memprediksi citra dengan benar, dan yang kedua secara tabular dengan memprediksi 10 citra masukan dengan memberikan hasil akurasi prediksi yang dilakukan oleh model yang berfokus untuk melihat bagaimana akurasi prediksi model bekerja dalam memprediksi citra masukan dengan benar.

HASIL DAN PEMBAHASAN

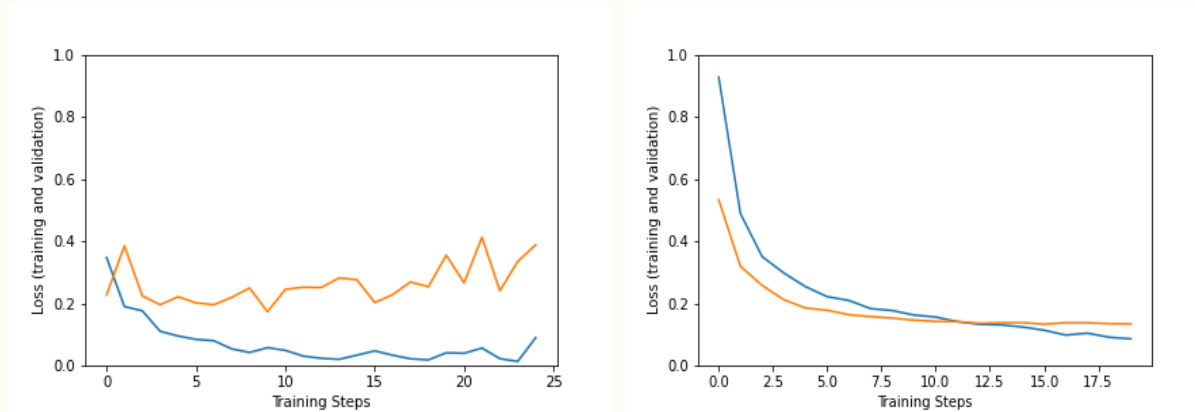
Hasil evaluasi model terhadap penerapan teknik augmentasi didapatkan bahwa penggunaan teknik augmentasi paling optimal yaitu dengan hanya memanfaatkan teknik pembalikan horisontal, dibuktikan berdasarkan grafik pelatihan model CNN yang lebih baik dibandingkan dengan penggunaan teknik augmentasi campuran. Berdasarkan pengamatan penerapan augmentasi rotasi acak, rentang zoom lebar, atau bahkan pembalikan vertikal dapat menimbulkan bias pada model klasifikasi citra sinar-X, sehingga model memiliki grafik pelatihan yang kurang optimal. Hasil tersebut dapat terlihat pada Gambar 2.



Gambar 2. Perbandingan Teknik Augmentasi

Gambar 2, memperlihatkan bahwa titik tertinggi yang dicapai pada model (b) sedikit lebih tinggi dengan keterhubungan kedua garis yang saling bertemu diakhir puncak dibandingkan dengan keterhubungan yang terjadi antara kedua garis yang terdapat pada model (a). Pengamatan ini menggarisbawahi dampak yang cukup signifikan mengenai kesesuaian penggunaan teknik augmentasi terhadap performa model secara keseluruhan. Teknik augmentasi yang tepat dapat memainkan peran penting dalam meningkatkan kemampuan model untuk menggeneralisasi dan membuat prediksi yang akurat.

Hasil evaluasi model terhadap pengaturan parameter pelatihan khususnya dalam algoritma pengoptimalan menunjukan hasil yang signifikan, penggunaan algoritma pengoptimalan Adam dengan hyperparameter: kecepatan pembelajaran $9e-5$, nilai beta 1 sebesar 0,19, dan nilai beta 2 sebesar 0,25. Selain itu, mendapatkan hasil yang terbaik dibandingkan dengan uji coba parameter lainnya, Perbandingan grafik kerugian yang menonjolkan konfigurasi akhir terhadap konfigurasi uji coba dapat dilihat pada Gambar 3.



(a) Grafik Kerugian pada Konfigurasi Uji Coba

(b) Grafik Kerugian pada Konfigurasi Akhir

Gambar 3. Perbandingan Konfigurasi Algoritma Pengoptimalan

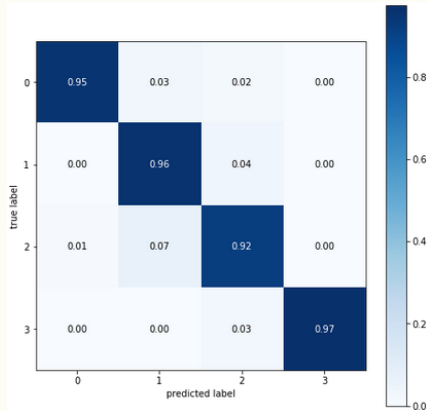
Eksperimen yang digambarkan pada Gambar 3, menggunakan skema pelatihan data dengan rasio 8:2. Perbedaan substansial terlihat pada konfigurasi parameter pengoptimalan yang digunakan dalam masing-masing penelitian. Dalam konfigurasi (a) menunjukkan ketidakstabilan dan tingkat *overfitting* yang signifikan, sehingga menunjukkan perbedaan yang cukup besar antara kerugian pelatihan dan validasi. Sebaliknya, konfigurasi (b) menampilkan grafik yang lebih stabil, menunjukkan regularisasi yang lebih baik dan tingkat *overfitting* yang kecil. Temuan ini menyoroti pengaruh penting dari konfigurasi parameter algoritma pengoptimalan terhadap kinerja dalam mencapai peningkatan stabilitas dan generalisasi dalam model.

Evaluasi metrik yang digunakan dalam mengevaluasi model CNN pada penelitian ini adalah validasi akurasi, presisi, perolehan, dan F1. Hasil evaluasi metrik dari masing-masing model yang telah dilatih dapat dilihat pada Tabel 1.

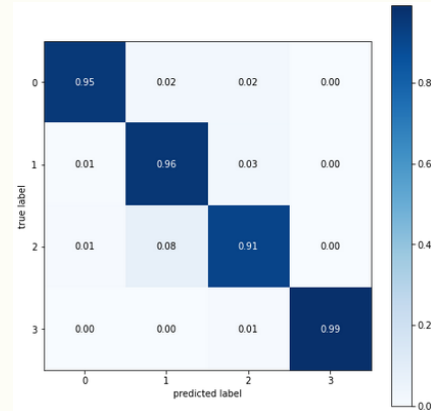
Tabel 1. Hasil Evaluasi Metrik

Model	Akurasi	Presisi	Perolehan	F1
Model 6:4	0.9514	0.95.13	0.9497	0.9505
Model 7:3	0.9566	0.95.81	0.9531	0.9556
Model 8:2	0.9583	0.96.17	0.9583	0.96
Model 9:1	0.9705	0.97.38	0.9688	0.9713

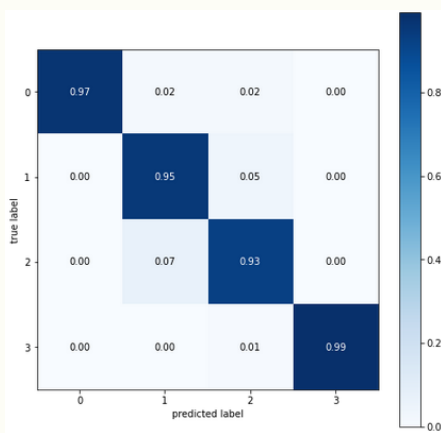
Berdasarkan Tabel 1, pemanfaatan arsitektur EfficientNetV2_b0 sebagai model dasar dan Adam sebagai pengoptimal telah menunjukkan kinerja yang cukup baik. Terlihat dari akurasi tertinggi yang dicapai mencapai 97,05%, disertai dengan skor f1 yang mengesankan sebesar 97,13%. Selanjutnya merupakan hasil evaluasi menggunakan matriks konfusi dengan skema percobaan prediksi pada model 6:4, 7:3, 8:2, dan 9:1 untuk 500 gambar validasi divisualisasikan pada Gambar 4. Matriks konfusi berisi nilai numerik yang sesuai dengan label kelas pengujian, dengan 0 mewakili COVID-19, 1 untuk Opasitas Paru, 2 untuk Normal, dan 3 untuk label kelas Viral Pneumonia. Setiap kelas direpresentasikan dalam matriks konfusi, memungkinkan penilaian komprehensif terhadap kinerja model di berbagai kategori. Ini menunjukkan keefektifan arsitektur dan pengoptimal yang dipilih, menyoroti kemampuannya dalam mengklasifikasikan gambar tertentu secara akurat dan berkontribusi pada performa model secara keseluruhan.



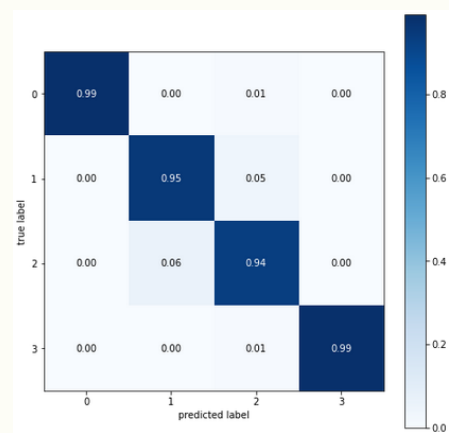
(a) Model 6:4



(b) Model 7:3



(c) Model 8:2

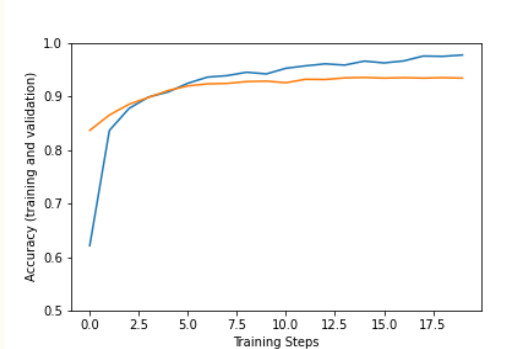


(d) Model 9:1

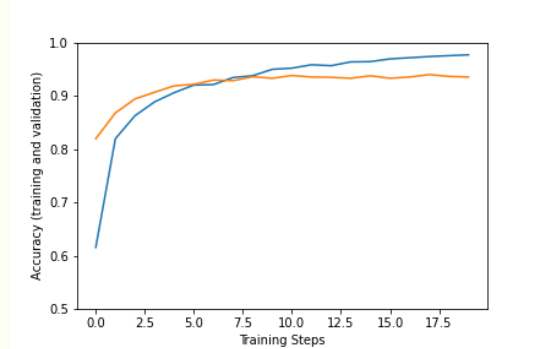
Gambar 4. Perbandingan Hasil Matriks Konfusi

Berdasarkan Gambar 4, terlihat model (d) memiliki tingkat akurasi yang tertinggi pada tiap kelasnya dibandingkan dengan model (c), (b), maupun (a). Ini menunjukkan adanya pengaruh secara linier antara banyaknya data pelatihan dengan tingkat akurasi yang didapatkan. Dan perlu ditekankan lagi bahwa akurasi yang didapatkan pada Gambar 4 dilakukan pada data validasi, diluar dari data pelatihan model.

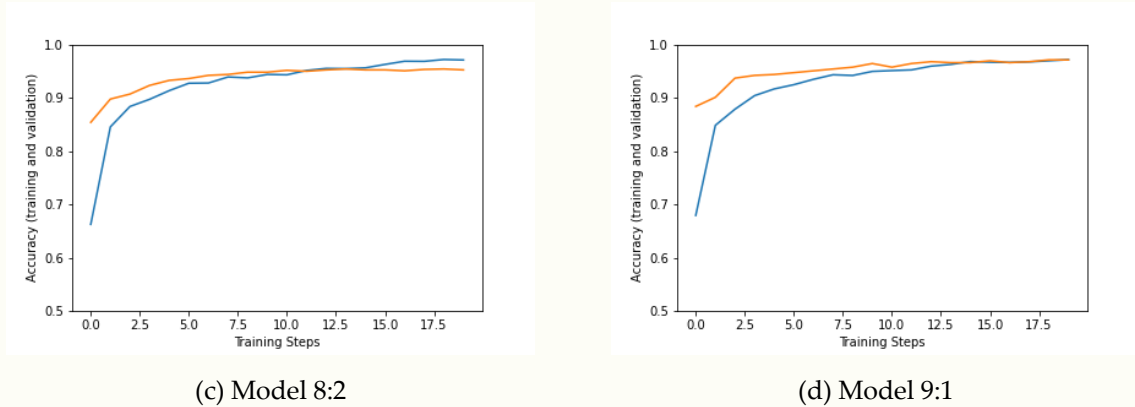
Selanjutnya, Hasil evaluasi diagram grafis menggunakan grafik akurasi dan kerugian yang telah dibuat menggunakan fungsi pyplot dapat dilihat pada Gambar 5 dan 6.



(a) Model 6:4

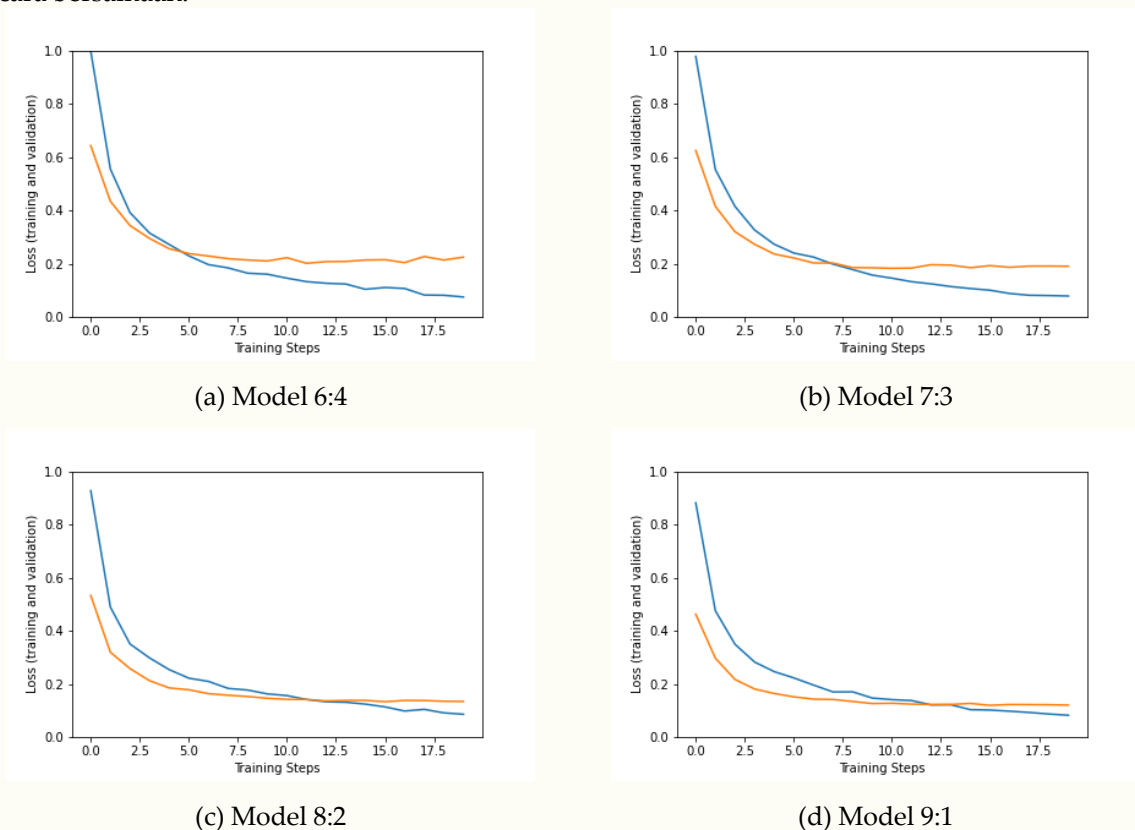


(b) Model 7:3



Gambar 5. Grafik Akurasi Pelatihan dan Validasi

Pada Gambar 5 dan Gambar 6, garis biru mewakili akurasi pelatihan, sedangkan garis oranye mewakili akurasi validasi. Gambar 5 menunjukkan bahwa seiring dengan meningkatnya rasio data pelatihan dalam model, akurasi validasi juga meningkat. Seperti yang terlihat pada model (d) kedua garis bertemu di akhir langkah, menunjukkan didapatkannya tingkat akurasi yang sama antara akurasi pelatihan dan akurasi validasi. Namun, pada model (c), *overfitting* mulai terjadi sekitar langkah ke-13. Demikian pula, untuk model (b), *overfitting* dimulai sekitar langkah ke-7.5, dan untuk model (a), *overfitting* terjadi ketika akurasi pelatihan terus meningkat, tetapi akurasi validasi gagal meningkat secara bersamaan.

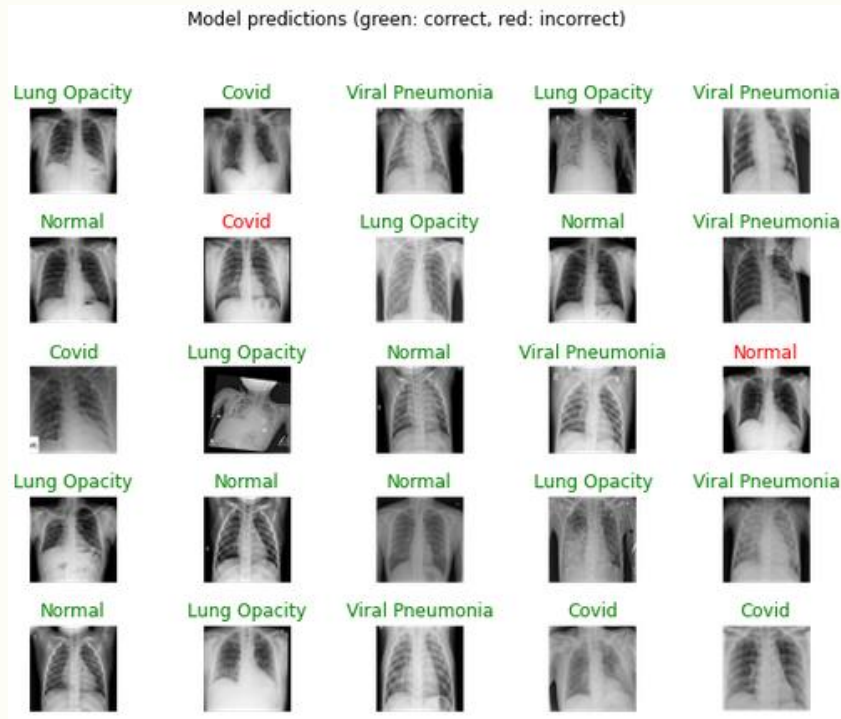


Gambar 6. Grafik Kerugian Pelatihan dan Validasi

Untuk Gambar 6, tingkat *overfitting* yang terjadi pada grafik kerugian model mengalami penurunan sesuai dengan banyaknya rasio data pelatihan yang dilakukan pada model, dengan rincian *overfitting* yang terjadi pada masing-masing model sebagai berikut, untuk model (d) *overfitting* dimulai terjadi pada langkah ke-13, sedangkan untuk model (c) *overfitting* terjadi pada langkah ke-12, dan untuk model (b) dan (a) *overfitting* terjadi pada langkah ke-7 dan ke-4. Berdasarkan hasil evaluasi

komprehensif dari evaluasi metrik dan pelatihan model grafik, terbukti bahwa jumlah data pelatihan secara linear berperan penting dalam memperoleh model CNN yang lebih baik. Hubungan ini ditunjukkan dengan jelas oleh grafik akurasi dan kerugian, khususnya pada model (d), yang menunjukkan akurasi tertinggi, tingkat kerugian rendah, dan tingkat *overfitting* terendah. Temuan ini menggarisbawahi pentingnya mempertimbangkan kuantitas dan distribusi data pelatihan ketika bertujuan untuk mengoptimalkan kinerja dan kemampuan generalisasi model CNN.

Hasil visualisasi pengujian prediksi model, model yang digunakan dalam pengujian prediksi merupakan model 9:1 yang merupakan model dengan akurasi terbaik dan tingkat kerugian terkecil yang didapatkan pada penelitian ini. Hasil pengujian prediksi dapat terlihat pada Gambar 7 dan Gambar 8.



Gambar 7. Visualisasi Hasil Prediksi terhadap Citra

Gambar 7 menunjukkan bahwa dari 25 citra yang dipilih secara acak, model 9:1 dapat memprediksi 23 citra secara akurat. Visualisasi hasil dalam format data tabel dapat terlihat pada Gambar 8 di bawah.

	Covid	Lung Opacity	Normal	Viral Pneumonia
0	0.00010680	0.99966109	0.00020873	0.00002337
1	0.99918574	0.00021412	0.00044138	0.00015865
2	0.00115421	0.00079533	0.00161931	0.99643111
3	0.00003043	0.99948478	0.00035067	0.00013417
4	0.00004577	0.00010617	0.00008209	0.99976593
5	0.00000374	0.00099773	0.99896920	0.00002934
6	0.19033352	0.35951722	0.44038692	0.00976236
7	0.03207203	0.85101181	0.09493526	0.02198096
8	0.00017757	0.02519360	0.97456723	0.00006161
9	0.00002946	0.00017692	0.00007950	0.99971408
10	0.80425459	0.05881646	0.01532831	0.12160057

Gambar 8. Visualisasi Hasil Prediksi terhadap Citra

Gambar 8 menampilkan prediksi model yang dibuat pada 10 data acak. Pada gambar ini, prediksi setiap gambar ditampilkan sebagai nilai probabilitas untuk setiap kelas. Mengingat fungsi aktivasi yang digunakan dalam penelitian ini adalah *softmax*, maka total probabilitas yang diprediksi oleh model untuk setiap kelas harus berjumlah 1 atau 100%. Misalnya, pada baris ke-0, probabilitas gambar yang termasuk dalam kelas Opasitas Paru diperkirakan sebesar 99,9%, dan sisa probabilitas untuk kelas lainnya berjumlah 0,1%. Demikian pula, pada data baris ke-6 total dari probabilitas tiap-tiap kelas akan berjumlah 1 atau 100%, namun hal menarik terjadi pada baris ke-6 dimana tingkat probabilitas antara kelas opasitas paru dan normal memiliki nilai probabilitas yang tidak terpaut jauh, yaitu 35,95% dan 44,03%. Kasus ini mungkin terjadi dikarenakan adanya kesamaan ciri antara kedua kelas tersebut atau opasitas paru yang terjadi pada citra sinar-X dari data validasi tersebut masih tergolong ringan, sehingga mengakibatkan kemungkinan hasil tingkat akurasi yang relatif dekat antara 2 kelas tersebut.

KESIMPULAN

Penelitian ini berhasil membentuk model klasifikasi terhadap 4 jenis penyakit: COVID-19, normal, opasitas paru, dan pneumomonia viral menggunakan arsitektur EfficientNetV2 sebagai dasar model CNN dengan akurasi tertinggi sebesar 97,05%. Keberhasilan ini didapat berdasarkan pengujian penggunaan teknik sampel ulang dan augmentasi data yang cukup efektif mengatasi ketidakseimbangan banyaknya data pada setiap kelas yang ada dalam basis data Radiografi COVID-19. Hasil evaluasi mengenai penkonfigurasi algoritme pengoptimalan Adam beserta periode pelatihan juga berkontribusi dalam menghasilkan tingkat *overfitting* yang minimal dan grafik pembelajaran yang lebih stabil. Evaluasi komprehensif menggunakan metrik akurasi dan grafik pelatihan pada penelitian ini, menunjukkan bahwa model 9:1 sebagai model klasifikasi yang paling baik diantara ke-3 model lainnya dengan akurasi 97,05%, presisi 97,38%, perolehan 96,88%, dan skor F1 97,13% beserta dengan tingkat kerugian terkecil dan minim *overfitting*.

Penelitian ini menggarisbawahi terdapatnya potensi penggunaan CNN sebagai alat bantu analisis citra medis. Penelitian lebih lanjut tentu bisa dilakukan, penambahan kategori mengenai tingkat keparahan penyakit yang diderita berdasarkan durasi infeksi dapat ditambahkan kedalam kumpulan data, mengingat pada penelitian ini data pelatihan hanya berfokus pada pengklasifikasian jenis penyakit berdasarkan kelas yang ada. Pendekatan yang diperluas ini memberdayakan model untuk mengkategorikan kasus COVID-19 berdasarkan perkembangan infeksi, sehingga memberikan pemahaman mendetail tentang kondisi pasien. Model seperti ini berpotensi meningkatkan pemahaman kita tentang penyakit dan mendukung profesional kesehatan dalam menyusun rencana perawatan yang dipersonalisasi, dengan mempertimbangkan stadium dan durasi infeksi COVID-19 pada pasien.

DAFTAR PUSTAKA

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., ... Zheng, X. (2015). *TensorFlow: A System for Large-Scale Machine Learning*. USENIX Association.
- Albawi, S., Mohammed, T. A., & Al-zawi, S. (2018). Understanding of a Convolutional Neural Network. *2017 International Conference on Engineering and Technology (ICET)*. <https://doi.org/10.1109/ICEngTechnol.2017.8308186>
- Antonio Gulli, & Sujit Pal. (2017). *Deep learning with Keras*. Packt Publishing Ltd.
- Arriani, A., Pradityas, H., Sofiani, L., Iwisara, N. G., Anandita, Y., Utami, E. S., Rahardja, M., Yenuarizki, & Dewi, N. (2020, Juli). The Community's Perception Of Covid-19. *Suara Komunitas*, 1-27.
- Bock, S., Goppold, J., & Weiß, M. (2018). *An improvement of the convergence proof of the ADAM-Optimizer*. <http://arxiv.org/abs/1804.10587>

- Chowdhury, M. E. H., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M. A., Mahbub, Z. Bin, Islam, K. R., Khan, M. S., Iqbal, A., Emadi, N. Al, Reaz, M. B. I., & Islam, M. T. (2020). Can AI Help in Screening Viral and COVID-19 Pneumonia? *IEEE Access*, 8, 132665–132676. <https://doi.org/10.1109/ACCESS.2020.3010287>
- Jia, G., Lam, H. K., & Xu, Y. (2021). Classification of COVID-19 chest X-Ray and CT images using a type of dynamic CNN modification method. *Computers in Biology and Medicine*, 134. <https://doi.org/10.1016/j.compbiomed.2021.104425>
- Khalidy, M. Al. (2018). Resampling Imbalanced Class and the Effectiveness of Feature Selection Methods for Heart Failure Dataset. *International Robotics & Automation Journal*, 4(1). <https://doi.org/10.15406/iratj.2018.04.00090>
- Kingma, D. P., & Ba, J. (2014). *Adam: A Method for Stochastic Optimization*. <http://arxiv.org/abs/1412.6980>
- Lee, H., & Song, J. (2019). Introduction to convolutional neural network using Keras; An understanding from a statistician. *Communications for Statistical Applications and Methods*, 26(6), 591–610. <https://doi.org/10.29220/CSAM.2019.26.6.591>
- Luthfi, M. I., & Septiyanti, N. D. (2023). Design Thinking Untuk Analisis Masalah Pembelajaran Daring Pada Masa Pandemi Covid-19 Di Indonesia. *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(2), 193–204. <https://doi.org/10.51454/decode.v3i2.146>
- Ma, W., & Lu, J. (2017). *An Equivalence of Fully Connected Layer and Convolutional Layer*. <http://arxiv.org/abs/1712.01252>
- Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., & Rajendra Acharya, U. (2020). Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Computers in Biology and Medicine*, 121. <https://doi.org/10.1016/j.compbiomed.2020.103792>
- Perez, L., & Wang, J. (2017). *The Effectiveness of Data Augmentation in Image Classification using Deep Learning*. <http://arxiv.org/abs/1712.04621>
- Powers, D. M. W., & Ailab. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *International Journal of Machine Learning Technology* 2, 37–63. <https://doi.org/https://doi.org/10.48550/arXiv.2010.16061>
- Putra, N. S., Hutabarat, B. F., & Khaira, U. (2023). Implementasi Algoritma Convolutional Neural Network Untuk Identifikasi Jenis Kelamin Dan Ras. *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(1), 82–93. <https://doi.org/10.51454/decode.v3i1.123>
- Rahman, T., Dr. Chowdhury, M., & Khandarkar, A. (2021). *COVID-19 Radiography (COVID-19 Chest X-ray Database) Dataset*. <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>
- Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Abul Kashem, S. Bin, Islam, M. T., Al Maadeed, S., Zughair, S. M., Khan, M. S., & Chowdhury, M. E. H. (2021). Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. *Computers in Biology and Medicine*, 132. <https://doi.org/10.1016/j.compbiomed.2021.104319>
- Reddi, S. J., Kale, S., & Kumar, S. (2019). *ON THE CONVERGENCE OF ADAM AND BEYOND*. <https://doi.org/https://doi.org/10.48550/arXiv.1904.09237>
- Rousan, L. A., Elobeid, E., Karrar, M., & Khader, Y. (2020). Chest x-ray findings and temporal lung changes in patients with COVID-19 pneumonia. *BMC Pulmonary Medicine*, 20(1). <https://doi.org/10.1186/s12890-020-01286-5>
- Singhal, T. (2020). A Review of Coronavirus Disease-2019 (COVID-19). Dalam *Indian Journal of Pediatrics* (Vol. 87, Nomor 4, hlm. 281–286). Springer. <https://doi.org/10.1007/s12098-020-03263-6>

- Sutanty, E., & Kusuma Astuti, D. (2023). Penerapan Model Arsitektur Vgg16 Untuk Klasifikasi Jenis Sampah. *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(2), 407-419. <https://doi.org/10.51454/decode.v3i2.331>
- Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *International Conference on Machine Learning*. <https://doi.org/10.48550/arXiv.1905.11946>
- Tan, M., & Le, Q. V. (2021). EfficientNetV2: Smaller Models and Faster Training. *Computer Vision and Pattern Recognition*. <https://doi.org/10.48550/arXiv.2104.00298>
- Warren, M. A., Zhao, Z., Koyama, T., Bastarache, J. A., Shaver, C. M., Semler, M. W., Rice, T. W., Matthay, M. A., Calfee, C. S., & Ware, L. B. (2018). Severity scoring of lung oedema on the chest radiograph is associated with clinical outcomes in ARDS. *Thorax*, 73(9), 840-846. <https://doi.org/10.1136/thoraxjnl-2017-211280>
- Yadav, S., Kaur Sandhu, J., Pathak, Y., Jadhav Babasaheb, S., Jadhav, S., & Yadav Assistant Professor, S. (2020). Chest X-ray scanning based detection of COVID-19 using deepconvolutional neural network. *Chest X-ray scanning based detection of COVID-19 using deep convolutional neural network*. <https://doi.org/10.21203/rs.3.rs-58833/v1>
- Yadav, S. S., & Jadhav, S. M. (2019). Deep convolutional neural network based medical image classification for disease diagnosis. *Journal of Big Data*, 6(1). <https://doi.org/10.1186/s40537-019-0276-2>
- Yasin, R., & Gouda, W. (2020). Chest X-ray findings monitoring COVID-19 disease course and severity. *Egyptian Journal of Radiology and Nuclear Medicine*, 51(1). <https://doi.org/10.1186/s43055-020-00296-x>
- Zhang, H., Zhang, L., & Jiang, Y. (2019). Overfitting and Underfitting Analysis for Deep Learning Based End-to-end Communication Systems. *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*. <https://doi.org/10.1109/WCSP.2019.8927876>