

## DEEP LEARNING IMAGE CLASSIFICATION RONTGEN DADA PADA KASUS COVID-19 MENGGUNAKAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK

Leni Anggraini Susanti<sup>\*1</sup>, Agus M Soleh<sup>2</sup>, Bagus Sartono<sup>3</sup>

<sup>1,2,3</sup>Institut Pertanian Bogor, Bogor

Email: <sup>1</sup>lenianggraini@apps.ipb.ac.id, <sup>2</sup>agusms@apps.ipb.ac.id, <sup>3</sup>bagusco@apps.ipb.ac.id

<sup>\*</sup>Penulis Korespondensi

(Naskah masuk: 21 Maret 2023, diterima untuk diterbitkan: 26 September 2023)

### Abstrak

Penelitian ini mengusulkan penggunaan *Convolutional Neural Network* (CNN) dengan arsitektur VGGNet-19 dan ResNet-50 untuk diagnosis COVID-19 melalui analisis citra rontgen dada. Modifikasi dilakukan dengan membandingkan nilai regularisasi *dropout* 50% dan 80% untuk kedua arsitektur dan mengubah jumlah lapisan klasifikasi menjadi 4 kelas. Selanjutnya, kinerja model dibandingkan berdasarkan ukuran dataset. Dataset terdiri dari 21165 citra, dengan pembagian 10% sebagai data uji dan 90% data dibagi menjadi data latih (80%) dan data validasi (20%). Kinerja model dievaluasi menggunakan metode validasi silang berulang 5 kali lipat. Proses pelatihan menggunakan *learning rate* 0.0001, optimasi *stochastic gradient descent* (SGD), dan sepuluh iterasi. Hasil penelitian menunjukkan bahwa penambahan lapisan *dropout* dengan peluang 50% untuk kedua arsitektur secara efektif mengatasi *overfitting* dan meningkatkan performa model. Ditemukan bahwa kinerja yang lebih baik dicapai pada ukuran kumpulan data lebih besar dan memberikan peningkatan signifikan pada kinerja model. Hasil klasifikasi menunjukkan arsitektur ResNet-50 mencapai akurasi rata-rata 94.4%, *recall* rata-rata 94.1%, presisi rata-rata 95.5%, spesifisitas rata-rata 97% dan *F1-score* rata-rata 94.8%. Sedangkan arsitektur VGGNet-19 mencapai akurasi rata-rata 91%, *recall* rata-rata 89%, presisi rata-rata 95.0%, spesifisitas rata-rata 96.8% dan *F1-score* rata-rata 92.7%. Pemanfaatan model ini dapat membantu mengidentifikasi penyebab kematian pasien dan memberikan informasi yang berharga bagi pengambilan keputusan medis dan epidemiologi.

**Kata kunci:** Citra, CNN, Klasifikasi, ResNet-50, VGGNet-19

## DEEP LEARNING IMAGE CLASSIFICATION CHEST X-RAY IN COVID-19 CASES USING THE CONVOLUTIONAL NEURAL NETWORK ALGORITHM.

### Abstract

*This research proposes using a Convolutional Neural Network (CNN) with VGGNet-19 and ResNet-50 architectures for COVID-19 diagnosis through chest X-ray image analysis. Modifications were made by comparing the dropout regularization values of 50% and 80% for both architectures and altering the number of classification layers to 4 classes. Furthermore, the model's performance was compared based on dataset size. The dataset comprised 21,165 images, with a division of 10% for testing and 90% divided into training data (80%) and validation data (20%). The model's performance was evaluated using the 5-fold repeat cross-validation method. The training process employed a learning rate of 0.0001, stochastic gradient descent (SGD) optimization, and ten iterations. The study's results indicate that adding dropout layers with a 50% probability for both architectures effectively addressed overfitting and improved the model's performance. It was found that better performance was achieved with larger dataset sizes. The classification results indicate the ResNet-50 architecture achieved an average accuracy of 94.4%, average recall of 94.1%, average precision of 95.5%, average specificity of 97%, and average F1-score of 94.8%. Meanwhile, the VGGNet-19 architecture achieved an average accuracy of 91%, an average recall of 89%, average precision of 95.0%, average specificity of 96.8%, and an average F1-score of 92.7%. Utilizing these models can assist in identifying the causes of patient mortality and offer valuable information for medical and epidemiological decision-making.*

**Keywords:** Image, Classification, CNN, ResNet-50, VGGNet-19

## 1. PENDAHULUAN

Penyakit COVID-19 disebabkan oleh virus baru yang pertama kali ditemukan di Wuhan, China pada Desember 2019. Penyebarannya sangat cepat mengancam serius kesehatan masyarakat secara global. Salah satu metode yang diandalkan dalam mendiagnosis pasien terinfeksi adalah metode RT-PCR. Namun, metode ini mahal dan memakan waktu. Oleh karena itu, penting untuk mengembangkan sistem diagnostik yang cepat dan akurat untuk mengidentifikasi pasien yang terinfeksi COVID-19. Salah satu pendekatan yang dapat dilakukan adalah analisis citra rontgen dada. *Computerised Tomography* (CT) untuk tingkat infeksi COVID-19 sekitar 98% dibandingkan dengan sensitivitas RT-PCR 71% (Y. Fang, H. Zhang, J. Xie, M. Lin *et al.*, 2020). Penelitian dilakukan oleh Long *et al* (2020) di China, ditemukan bahwa dari 36 kasus COVID-19 dengan pneumonia CT-Chest, terdapat 35 pasien yang menunjukkan hasil CT abnormal. Sementara itu, melalui pengujian RT-PCR, sebanyak 30 pasien dinyatakan positif. Sensitivitas CT-Chest dilaporkan mencapai 97.2%, sedangkan sensitivitas RT-PCR mencapai 84.6%.

Data citra *chest radiograph* merupakan data gambar dua dimensi dan dapat diolah menggunakan metode *Convolutional Neural Network* (CNN) dalam *deep learning*. Dengan menggunakan arsitektur yang khusus dikembangkan untuk memproses data gambar, CNN mampu mengidentifikasi pola dan fitur-fitur penting dalam gambar dengan cara yang mirip dengan cara kerja visual manusia. Dalam CNN, terdapat lapisan konvolusi yang dapat mempelajari fitur-fitur yang mendeteksi pola-pola spesifik dalam gambar, serta lapisan *pooling* yang membantu mereduksi dimensi dan mengambil fitur-fitur yang paling signifikan. Hasil analisis citra rontgen dada menggunakan CNN dapat memberikan informasi apakah kematian tersebut disebabkan oleh COVID-19 atau faktor lain yang tidak terkait. Hal ini dapat membantu dalam pemahaman lebih lanjut tentang dampak dan karakteristik COVID-19 serta memberikan informasi yang berharga untuk pengambilan keputusan medis dan epidemiologi.

Penelitian oleh Yi Zhong (2020) mengenai *image classification rontgen* dada kedalam 2 kelas yaitu positif COVID-19 dan normal menggunakan arsitektur VGG16. Hasil yang diperoleh adalah model mencapai akurasi 87.3%, presisi 89.7% dan recall 84.446%. Penelitian yang dilakukan oleh Swastika (2020) menggunakan arsitektur VGGNet-16 dengan membandingkan optimizer SGD dan Adam untuk mengklasifikasikan infeksi COVID-19 dan normal. Hasilnya menunjukkan akurasi sebesar 92.85% dalam mengklasifikasikan infeksi COVID-19 dan normal. Selain itu, untuk pelatihan dengan optimizer SGD, nilai spesifisitas dan sensitivitas mencapai 100% dan 85.71% secara berturut-turut.

Victor Ikechukwu *et al* (2021) membandingkan VGGNet-19 dan ResNet-50 untuk mengklasifikasikan normal dan viral pneumonia. Penelitian lainnya dilakukan oleh Apostolopoulos *et al* (2020) menggunakan CNN yang dilatih untuk mengklasifikasikan citra *rontgen* dada kedalam tiga kelas. Penelitian mengklasifikasikan citra *rontgen* dada ke dalam tiga kelas juga dilakukan oleh Chowdhury *et al* (2020) dan Sakib *et al* (2021).

Literatur pada penelitian-penelitian tersebut menggunakan kumpulan data citra yang berbeda beda, dengan jumlah kasus COVID-19 bervariasi yaitu dari 25 hingga 224 di antara jumlah total kasus dari 50 hingga 2905 dan jumlah kelas 2 hingga 3 kelas. Diketahui sensitivitas yang dilaporkan untuk mendeteksi kasus COVID-19 berkisar antara 79.0%-97.95%. Terlepas dari hasil yang menjanjikan yang dilaporkan dalam penelitian sebelumnya, terdapat masalah yang belum diselidiki dengan baik terkait bagaimana melatih model *deep learning* secara optimal. Misalnya, data yang digunakan terbilang kecil, hal ini menyulitkan untuk menggeneralisasikan hasil yang dilaporkan dan tidak dapat menjamin bahwa performa yang dilaporkan akan dipertahankan saat model akan diuji pada kumpulan data yang lebih besar. Selain itu adalah apakah pada model dilakukan metode penanganan *overfitting*. Model juga rentan terhadap terjadinya *overfitting*, yaitu model mampu bekerja dengan baik pada data latih namun tidak untuk data uji ataupun data lainnya yang belum dikenal model.

Oleh karena itu, pada penelitian ini dikembangkan sebuah sistem untuk mendeteksi kemungkinan COVID-19 pada suatu citra *rontgen* dada dengan skala data lebih besar yaitu 21165 citra, dengan 4 kelas (COVID-19, *lung opacity*, normal, *viral pneumonia*) dan melakukan perbandingan performa model menggunakan jumlah dataset 300 citra untuk setiap kelas dengan total dataset 1200 citra. Adapun model yang digunakan adalah model CNN dengan metode *transfer learning* untuk identifikasi positif COVID-19.

## 2. METODE PENELITIAN

Terdapat beberapa tahapan untuk mencapai tujuan penelitian. Berikut adalah langkah-langkah tahapan analisis data secara menyeluruh yang dilakukan dalam penelitian ini.

### 2.1 Tahap Pengumpulan Data

Mengumpulkan dataset citra rontgen dada yang tersedia pada laman *kaggle.com* oleh Tawisfurrahman. Terdapat 21165 citra *rontgen* dada, terdiri dari 3616 citra *rontgen* dada COVID-19, 6012 citra *rontgen* dada *lung opacity*, 10192 citra *rontgen* dada normal dan 1345 citra *viral pneumonia*. Setiap citra memiliki format *portable network graphics* (PNG) dengan ukuran 299x299 piksel.

## 2.2 Tahap Pembagian Data

Data dibagi menjadi tiga bagian, yaitu data uji, data latih dan data validasi. Data uji menggunakan 10% dari total keseluruhan data dan 90% dibagi menjadi data latih (80%) dan data validasi (20%) untuk mengevaluasi kinerja model pada setiap tahap pelatihan. Penelitian ini menggunakan validasi silang berulang 5 kali. Hal ini berarti proses pelatihan evaluasi model dilakukan sebanyak lima kali, dengan membagi data secara acak menjadi lima kelompok (*fold*) yang setara. Setiap kelompok bergantian menjadi data validasi sementara yang lainnya digunakan sebagai data latih. Dengan demikian, kinerja model dinilai berdasarkan rata-rata hasil dari kelima percobaan tersebut.

## 2.3 Tahap Praproses Data

Sebelum pelatihan, dilakukan *preprocessing* citra untuk mengurangi beban kerja mesin, mempercepat waktu dan meningkatkan akurasi proses pelatihan. Tahap awal adalah melakukan augmentasi untuk meningkatkan variasi data. Beberapa proses augmentasi yang dilakukan adalah rotasi citra, memperbesar gambar dan menerapkan pergeseran sudut citra. Kemudian mengubah ukuran citra yang disesuaikan dimensinya dengan bentuk masukan citra pada model VGGNet-19 dan ResNet-50 menjadi  $224 \times 224$  piksel. Selanjutnya menggunakan *channel* warna *grayscale*. Tahap praproses (augmentasi data, mengubah ukuran citra dan *channel* warna) dilakukan sekaligus dengan pembangkitan dataset menggunakan fitur *ImageDataGenerator* pada package *Keras*.

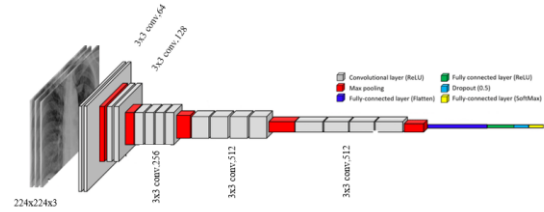
## 2.4 Pembentukan model CNN

Pembentukan model CNN menggunakan arsitektur VGGNet-19 dan ResNet-50 sebagai model dasar yang dimodifikasi.

### 2.4.1 Visual Geometry Group Network 19 (VGGNet-19)

VGGNet-19 menggunakan konsep blok untuk membentuk lapisan konvolusi yang masing masing memiliki besaran *filter*  $3 \times 3$ . Pada akhir blok konvolusi dilakukan operasi *max pooling* dengan besaran *pool*  $2 \times 2$  dan *stride* 2. Blok pertama memiliki 2 lapisan konvolusi 64 *channel* dengan besaran *filter*  $3 \times 3$  dan *max pooling*. Blok kedua terdapat 2 lapisan konvolusi 128 *channel* dan operasi *max pooling*. Pada blok ketiga terdapat 4 lapisan konvolusi 256 *channel* dan pada akhir blok sama seperti blok sebelumnya yaitu *max pooling*. Blok keempat dan kelima terdapat 4 lapisan konvolusi 512 *channel* dan operasi *max pooling*. Total terdapat 16 lapisan konvolusi dengan *filter*  $3 \times 3$  dan 3 lapisan *pooling* dengan operasi yang sama pada semua blok. *Feature map* hasil dari semua blok konvolusi dengan dimensi (7,7,512)

tersebut akan disederhanakan menjadi bentuk vektor (1, 25088).



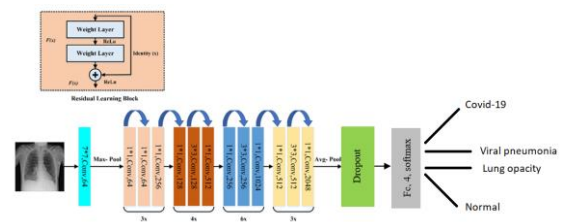
Gambar 1. Modifikasi arsitektur VGGNet-19

Modifikasi yang dilakukan pada arsitektur VGGNet-19 dalam penelitian ini menghasilkan perbedaan dengan arsitektur aslinya. Berikut adalah modifikasi yang dilakukan.

1. Menambah lapisan *dropout* dengan peluang 50% dan 80% secara acak untuk mengurangi *overfitting*, diikuti oleh lapisan aktivasi *ReLU* sebelum lapisan *fully connected*.
2. Mengatur jumlah *node* pada lapisan *fully connected* dari 1000 kelas menjadi 4 kelas

### 2.4.2 Residual Network 50 (ResNet-50)

Arsitektur kedua pada penelitian ini adalah ResNet-50.



Gambar 2. Modifikasi arsitektur ResNet-50

Pada tahap pertama, citra berukuran  $224 \times 224$  piksel diproses melalui lapisan konvolusi dengan *filter*  $7 \times 7$  dan *stride* 2. Hasilnya dinormalisasi menggunakan *batch normalization* dan diaktivasi dengan fungsi *ReLU* untuk mendapatkan ekstraksi fitur non-linear. Selanjutnya, hasilnya direduksi melalui lapisan *max pooling* sebelum dilanjutkan ke tahap konvolusi kedua. Proses ekstraksi fitur pada tahap kedua hingga kelima dilakukan menggunakan kombinasi *convolutional block* dan *identity block*. Setelah ekstraksi fitur, *feature map* diproses melalui lapisan *fully connected* untuk prediksi. Pada penelitian ini, terdapat beberapa modifikasi pada arsitektur aslinya. Salah satunya adalah penambahan lapisan *dropout* dengan peluang 50% dan 80%. Selain itu, jumlah *output* pada lapisan *fully connected* diubah menjadi 4 kelas sesuai dengan kebutuhan klasifikasi yang dilakukan.

## 2.5 Tahap Evaluasi

Hasil prediksi menggunakan CNN dievaluasi dengan menghitung nilai presisi, *recall* (sensitivitas), spesifisitas, *F1-score* dan akurasi. Proses tersebut

melibatkan pembagian dataset menjadi beberapa *subset*, dimana setiap *subset* akan digunakan secara bergantian sebagai data validasi sementara *subset* lainnya digunakan sebagai data latih. Proses ini diulang sebanyak 5 kali untuk memastikan bahwa model dievaluasi secara menyeluruh dengan menggunakan berbagai kombinasi data latih dan validasi yang berbeda. Nilai-nilai tersebut dari setiap model akan dibandingkan untuk memilih model terbaik, yang kemudian akan digunakan untuk menguji set data baru untuk melakukan deteksi citra.

Tabel 1. Contoh confusion matrix 4x4 (Garillos-Manliguez, C.A., 2016)

Aktual	Prediksi			
	Kelas 1	Kelas 2	Kelas 3	Kelas 4
Kelas 1	$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$
Kelas 2	$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$
Kelas 3	$x_{31}$	$x_{32}$	$x_{33}$	$x_{34}$
Kelas 4	$x_{41}$	$x_{42}$	$x_{43}$	$x_{44}$

Total nilai *false negative* (TFN), *false positive* (TFP) dan *true negative* (TN) untuk tiap kelas  $i$  akan dihitung berdasarkan persamaan (1), (2) dan (3). *Total true positive* (TTP) dihitung menggunakan persamaan (4).

$$TFN_i = \sum_{j=1}^4 x_{ij}, j \neq i \quad (1)$$

$$TFP_i = \sum_{j=1}^4 x_{ji}, j \neq i \quad (2)$$

$$TTN_i = \sum_{j=1}^4 \sum_{k=1}^4 x_{jk}, j \neq i, k \neq i \quad (3)$$

$$TTP_{all} = \sum_{j=1}^4 x_{jj} \quad (4)$$

Nilai presisi, *recall*, spesifisitas, *F-1 score* dan akurasi untuk setiap kelas  $i$  digunakan persamaan 5, 6, 7, 8 dan 9.

$$presisi_i = \frac{TTP_{all}}{TTP_{all} + TFP_i} \quad (5)$$

$$recall_i = \frac{TTP_{all}}{TTP_{all} + TFN_i} \quad (6)$$

$$spesifisitas_i = \frac{TTN_{all}}{TTN_{all} + TFP_i} \quad (7)$$

$$F1 - Score_i = 2 \times \frac{presisi_i \times recall_i}{presisi_i + recall_i} \quad (8)$$

$$Akurasi = \frac{TP_{all}}{TP_{all} + \sum_i \sum_j x_{ij}} \times 100\% \quad (9)$$

$$i \neq j, i, j = 1, 2, 3, 4$$

Nilai-nilai tersebut kemudian divisualisasikan menggunakan *boxplot*. Pada *boxplot*, nilai presisi, *recall*, spesifisitas dan *F1-score* yang lebih tinggi akan tercermin dalam posisi median yang lebih tinggi, serta letak kotak (*box*) dan batas atas (*upper*

*whisker*) yang lebih tinggi. Semakin besar kotak dan rentang *whisker* semakin besar variabilitas dalam data tersebut.

## 2.6 Tahap Deteksi

Tahap selanjutnya adalah menerapkan pendeteksian pada data baru berupa citra *rontgen* dada menggunakan model terbaik yang diperoleh. Data baru akan melalui tahap *preprocessing* sebelum masuk proses klasifikasi. Setelah klasifikasi dilakukan, akan ditampilkan citra *rontgen* dada beserta label hasil klasifikasi dan persentase akurasi masing-masing label.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Pembagian Data

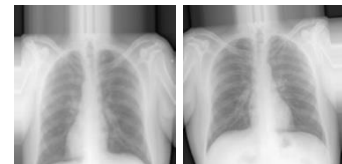
Penelitian ini menerapkan validasi silang berulang 5 kali. Citra dibagi menjadi tiga bagian yaitu 10% dari keseluruhan dataset sebagai data uji dan 90% sisanya dibagi menjadi dua bagian yaitu 80% data latih 20% data validasi. Hasil ditunjukkan pada Tabel 2.

Tabel 2. Distribusi pembagian data pada masing-masing kelas

Kelas	Data Latih	Data Validasi	Data Uji	Total
COVID-19	2637	613	366	3616
Lung opacity	4278	1137	597	6012
Viral pneumonia	955	239	151	1345
Normal	7368	1821	1003	10192
Total	15238	3810	2117	21165

### 3.2 Preprocessing Data

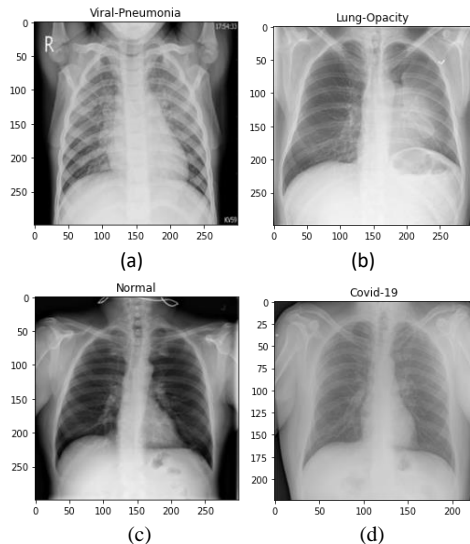
Pada penelitian ini, *preprocessing* yang dilakukan diawal adalah augmentasi yang bertujuan untuk generalisasi model. Proses augmentasi dilakukan dengan bantuan fungsi *ImageDataGenerator* dari package *Keras* dengan ketentuan yaitu secara acak memutar gambar sebesar  $30^\circ$ , memperbesar gambar secara acak dengan kisaran 0.85 sampai 1.15 dan melakukan *shear* dengan skala 0.15. Berikut hasil augmentasi ditunjukkan pada Gambar 3.



Gambar 3. Contoh hasil augmentasi citra

Selanjutnya, resolusi citra perlu diseragamkan denganmenubah ukuran sampel data menjadi 224x224 piksel, sesuai dengan standar input yang digunakan oleh arsitektur VGGNet-19 dan arsitektur ResNet-50 pada CNN. Kemudian, citra

hasil konversi ukuran diubah *channel* warnanya menjadi *grayscale* agar citra dari keseluruhan data seragam. Berikut adalah contoh hasil *preprocessing* dari masing-masing kelas citra.



Gambar 4. Konversi citra (a) COVID-19 (b) lung opacity (c) viral pneumonia (d) normal pada praproses menjadi 224x224 piksel

Proses konversi citra yang dilakukan membuat ukuran citra tersebut menjadi lebih kecil. Data masukan dalam jaringan CNN merupakan sebuah *array* berdimensi  $N \times \text{Width} \times \text{Height} \times 1$  yang terdiri dari nilai-nilai piksel dari dataset citra *rontgen* dada masing-masing kelas yang sudah diekstrak dalam proses *preprocessing*.  $N$  merupakan jumlah data, *width* dan *height* adalah dimensi citra yaitu 224x224 piksel, dan 1 merupakan dimensi untuk *channel grayscale*. *Range* nilai pada citra *channel grayscale* adalah nilai 0 sampai 255. Hal ini ditunjukkan seperti Gambar 5.



Gambar 5. Citra *Rontgen* Dada COVID-19 dengan nilai pada tiap piksel

### 3.3 Implementasi Model

Penelitian ini menggunakan dua model CNN yaitu VGGNet-19 dan ResNet-50. Masing-masing model tersebut menggunakan lapisan fitur pada arsitektur aslinya, sedangkan lapisan klasifikasi dimodifikasi dengan penambahan lapisan *dropout*. Penambahan lapisan *dropout* berarti beberapa *neuron* akan dinonaktifkan secara acak pada setiap iterasi pelatihan, sehingga memaksa model untuk tidak bergantung terlalu banyak pada fitur-fitur tertentu. Selain itu, jumlah label atau kelas yang

digunakan pada arsitektur VGGNet-19 dan ResNet-50 sebanyak 1000, sedangkan dalam penelitian ini hanya terdapat empat kelas. Oleh karena itu, lapisan klasifikasi pada kedua arsitektur tersebut juga perlu disesuaikan dengan jumlah kelas yang ada. Kedua arsitektur tersebut juga diterapkan untuk membandingkan performa model ketika digunakan jumlah data yang lebih kecil.

### 3.4 Pembuatan Model CNN

Metode yang digunakan dalam pembuatan model pada penelitian ini adalah membandingkan arsitektur VGGNet-19 dan arsitektur ResNet-50 pada CNN.

#### 3.4.1 Parameter Pelatihan Model

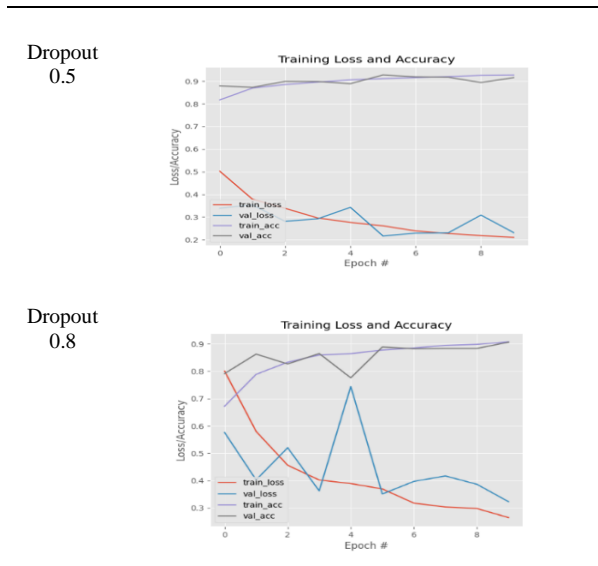
Proses pelatihan dilakukan untuk mengoptimalkan parameter guna membentuk sebuah model yang akan digunakan sebagai sistem klasifikasi. Kedua model tersebut menggunakan fungsi *optimizer* yang sama, yaitu *stochastic gradient descent* (SGD). Fungsi *optimizer* berfungsi untuk mengurangi *loss* dan meningkatkan akurasi dengan cara mengatur bobot dan *learning rate* pada model. Pada proses *training*, ada beberapa parameter yang harus di atur agar mendapatkan hasil model yang baik. Parameter tersebut diantaranya adalah *batch size*, *epoch*, iterasi dan *learning rate*. *Batch size* adalah istilah yang digunakan untuk menggambarkan jumlah sampel data yang diproses secara bersamaan dalam satu iterasi pada *neural network*. Karena jumlah data latih adalah 15238 dan menggunakan *batch size* 8, algoritma ini akan menggunakan 8 sampel data pertama dari total 15238 data untuk dilatih oleh jaringan saraf (*neural network*) hingga selesai. Kemudian, dilakukan pengambilan 8 sampel data berikutnya dari 15238 data dan proses ini dilanjutkan hingga mencapai delapan sampai data ke-1904 (15238/8). Nilai *epoch* yang digunakan adalah 10, artinya terdapat 10 tahapan atau iterasi proses *training* setiap *dataset* yang telah melalui proses *training* sampai di kembalikan ke awal untuk sekali putaran. Dan akan di perbaharui sebanyak 1904 iterasi. *Learning rate* digunakan untuk menghitung perubahan bobot yang diterapkan selama proses pelatihan pada *neural network*. Dalam kasus ini, digunakan *learning rate* sebesar 0.0001.

#### 3.4.2 Pelatihan Model Penanganan *Overfitting*

Setelah membuat model selanjutnya adalah melakukan proses pelatihan model. Untuk menangani dugaan adanya *overfitting* pada model, pada masing-masing model diterapkan lapisan *dropout*. Pada penelitian ini terdapat dua skema perbandingan nilai regularisasi *dropout* yang dilakukan terhadap model yaitu menggunakan *dropout* dengan nilai peluang 50% dan 80%.

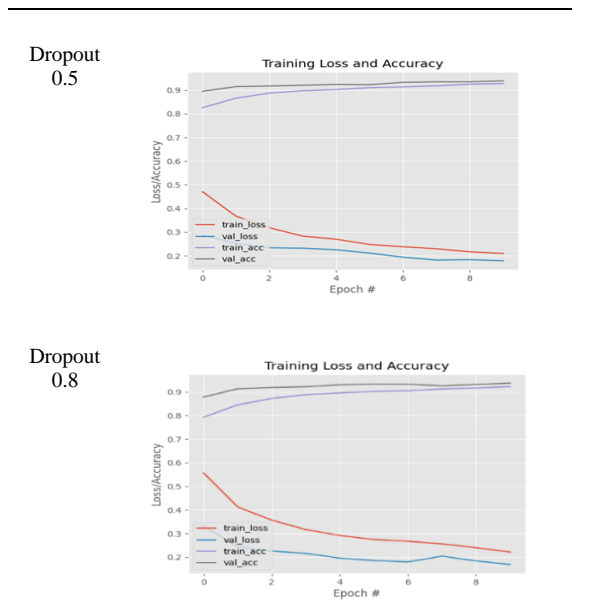
Artinya, sebanyak 50% dan 80% *neuron* akan dipilih secara acak dan tidak digunakan selama proses pelatihan.

Tabel 3. Perbandingan Penggunaan Regularisasi *Dropout* 50% dan 80% pada Model VGGNet-19



Tabel 3, menunjukkan perbandingan antara penggunaan *dropout* 0.5 dan 0.8 pada model VGGNet-19. Penggunaan *dropout* 0.5 menunjukkan peningkatan nilai akurasi yang cukup stabil. Setiap kenaikan dan penurunan nilai *val\_acc* dan *train\_acc* berjalan cukup wajar. Namun, pada penggunaan *dropout* 0.8, terdapat dugaan adanya *overfitting*, karena *loss* validasi meningkat sementara dan *loss* pelatihan menurun.

Tabel 4. Perbandingan Penggunaan Regularisasi *Dropout* 50% dan 80% pada Model ResNet-50



Tabel 4 menunjukkan, pada arsitektur ResNet-50 baik skema dengan menggunakan *dropout* 0.5 maupun dengan *dropout* 0.8 menunjukkan nilai *val\_acc* dan *train\_acc* berbanding lurus. Hal ini berarti diasumsikan tidak adanya *overfitting* pada model yang dihasilkan dari arsitektur ResNet-50. Namun, penggunaan *dropout* 0.5 menghasilkan model dengan *gap* atau jarak antara validasi dan train lebih kecil dibandingkan *dropout* 0.8. Hasil perbandingan ini membantu dalam memilih parameter *dropout* yang optimal untuk mencegah *overfitting* dan meningkatkan performa model. Dengan menggunakan *dropout* 0.5, model dapat mencapai hasil yang lebih baik dalam mengklasifikasi citra rontgen dada COVID-19, *viral pneumonia*, *lung opacity* dan normal dengan mengurangi perbedaan antara citra hasil validasi dan pelatihan.

### 3.4.3 Perbandingan Performa Model terhadap Jumlah Data

Pada jumlah dataset lebih kecil, terdapat 864 citra digunakan sebagai data latih, 216 citra sebagai data validasi dan 120 citra sebagai data uji. Sedangkan untuk jumlah dataset lebih besar, terdapat 15238 citra sebagai data latih, 3810 citra sebagai data validasi dan 2117 citra sebagai data uji.

Tabel 5 menunjukkan perbandingan performa masing-masing model terhadap jumlah data yang digunakan dari hasil model terbaik yang diperoleh kemudian diterapkan pada data uji. Nilai presisi, *recall*, spesifisitas, *F1-score* serta akurasi pada jumlah data yang lebih besar menunjukkan hasil yang lebih tinggi untuk setiap kelasnya dibandingkan dengan jumlah data lebih kecil. Hal ini mengindikasikan bahwa semakin banyak data yang digunakan dalam pelatihan model, semakin baik pula performa model tersebut. Dengan jumlah data yang lebih besar, model memiliki lebih banyak informasi yang digunakan untuk mempelajari pola dan fitur yang relevan. Akibatnya, model memiliki kemampuan yang lebih baik dalam mengklasifikasi citra rontgen dada COVID-19, *lung opacity*, *viral pneumonia* dan normal dengan presisi, *recall*, spesifisitas, *F1-score* dan akurasi yang lebih tinggi. Hasil ini konsisten dengan penelitian sebelumnya, seperti yang disebutkan oleh Chowdurry (2020), bahwa penggunaan jumlah data lebih kecil dapat menyulitkan dalam menggeneralisasikan hasil yang diperoleh dan tidak dapat menjamin bahwa performa yang dilaporkan akan dipertahankan ketika model diuji pada kumpulan data yang lebih besar.

Selain itu, terlihat bahwa model ResNet-50 pada jumlah data lebih besar memiliki akurasi sebesar 94.4%, sedangkan model VGGNet-19 memiliki akurasi sebesar 91.6%. Artinya, pada model ResNet-50 dengan jumlah dataset lebih

besar, 94.4% citra berhasil diklasifikasikan dengan benar.

Tabel 5. Performa Model terhadap Keseluruhan Dataset

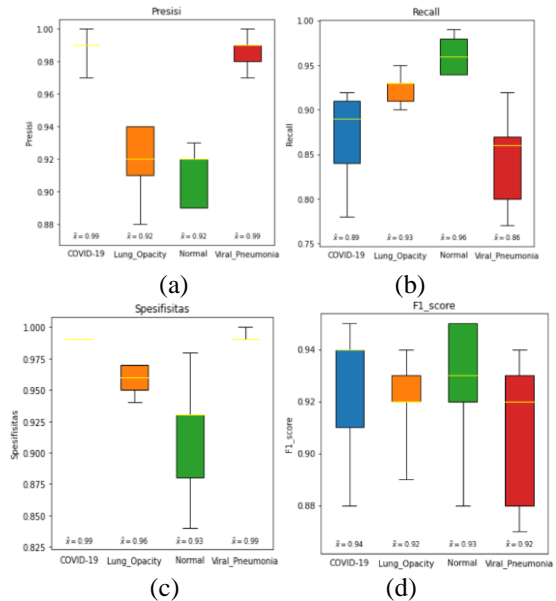
Total Dataset: 1200 Model: VGGNet-19 Akurasi: 84%				
Kelas	Presisi	Recall	Spesifisitas	F1-score
COVID-19	0.88	0.77	0.94	0.82
Lung opacity	0.76	0.87	0.91	0.80
Normal	0.80	0.91	0.90	0.85
Viral pneumonia	0.97	0.82	0.99	0.88
Total Dataset: 1200 Model: ResNet-50 Akurasi: 85%				
Kelas	Presisi	Recall	Spesifisitas	F1-score
COVID-19	0.90	0.92	0.95	0.91
Lung Opacity	0.84	0.85	0.94	0.85
Normal	0.85	0.87	0.94	0.86
Viral pneumonia	0.98	0.89	0.99	0.93
Total Dataset: 21165 Model: VGGNet-19 Akurasi: 91.6%				
Kelas	Presisi	Recall	Spesifisitas	F1-score
COVID-19	0.99	0.87	0.99	0.94
Lung Opacity	0.92	0.92	0.96	0.93
Normal	0.91	0.96	0.92	0.93
Viral pneumonia	0.99	0.84	0.99	0.92
Total Dataset: 21165 Model: ResNet-50 Akurasi: 94.4%				
Kelas	Presisi	Recall	Spesifisitas	F1-score
COVID-19	0.97	0.97	0.99	0.96
Lung Opacity	0.94	0.90	0.97	0.92
Normal	0.93	0.96	0.92	0.95
Viral pneumonia	0.98	0.93	0.99	0.95

Hal ini menunjukkan kemampuan model dalam mengenali pola dan fitur yang relevan dalam citra rontgen dada untuk mengklasifikasikan dengan akurasi yang tinggi. Sedangkan pada model VGGNet-19 dengan jumlah dataset lebih besar, 91.6% citra berhasil diklasifikasi dengan benar. Meskipun akurasi ini sedikit lebih rendah dibandingkan model ResNet-50, namun tetap menunjukkan performa yang baik dalam memprediksi kelas citra rontgen dada dengan tingkat keakuratan yang tinggi. Dengan demikian, peningkatan jumlah dataset yang digunakan dalam pelatihan model memberikan dampak positif terhadap akurasi dan performa model, terutama

pada model ResNet-50 yang mencapai tingkat akurasi yang lebih tinggi

### 3.5 Evaluasi Model

Kedua model dievaluasi menggunakan 2117 citra *rontgen* dada data uji dengan metrik presisi, *recall*, spesifisitas dan *F1-score* dan divisualisasikan menggunakan *boxplot*. Pada setiap percobaan, diperoleh 5 nilai metrik untuk setiap kelas yang dievaluasi.

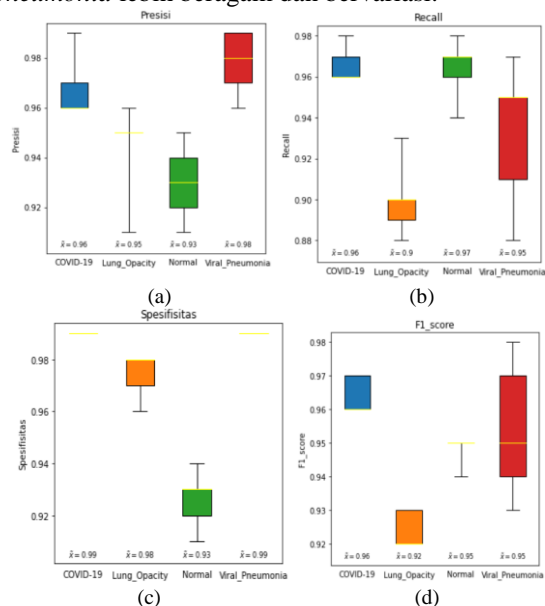


Gambar 6. *Boxplot* (a) presisi (b) *recall* (c) spesifisitas dan (d) *F1-score* model VGGNet-19

Gambar 6 menampilkan *boxplot* sebaran metrik presisi, *recall*, spesifisitas dan *F1-score* pada model arsitektur VGGNet-19. Pada Gambar 6(a), terlihat bahwa median presisi kelas COVID-19 dan *viral pneumonia* memiliki nilai yang sama dan lebih tinggi dibandingkan dengan kedua kelas lainnya, yaitu *lung opacity* dan normal. Namun pada kelas *viral pneumonia* memiliki kotak lebih besar dibandingkan COVID-19. Hal ini mengindikasikan bahwa hasil *tuning* pada kelas *viral pneumonia* cenderung memiliki keragaman dibandingkan dengan kelas COVID-19.

Pada Gambar 6(b) menunjukkan nilai median *recall* pada kelas normal lebih tinggi daripada kelas lainnya. Nilai *recall* lebih tinggi mengindikasikan bahwa model memiliki kemampuan yang lebih baik dalam mendeteksi dengan akurat kasus-kasus normal. Pada kelas COVID-19 kotak (*interquartile range*) memiliki lebar yang hampir sama dengan *viral pneumonia*, namun nilai kotak pada kelas COVID-19 lebih tinggi. Hal ini menunjukkan bahwa distribusi *recall* pada kelas COVID-19 memiliki konsentrasi yang lebih tinggi pada nilai-nilai yang lebih tinggi dibandingkan kelas *viral pneumonia*. Nilai spesifisitas masing-masing kelas pada Gambar 6(c) menunjukkan bahwa kelas COVID-19 dan *viral pneumonia* memiliki median yang sama dan hasil konsisten.

Berdasarkan Gambar 6(d), terlihat bahwa median *F1-score* untuk kelas COVID-19 lebih tinggi dibandingkan kelas lainnya. Hal ini menunjukkan bahwa model memiliki keseimbangan yang lebih baik antara presisi dan *recall* ketika memprediksi kasus COVID-19. Sedangkan kelas *viral pneumonia* memiliki median *F1-score* sedikit lebih rendah dibandingkan kelas lainnya. Kotak yang ditunjukkan pada kelas viral pneumonia lebih panjang dibandingkan kelas lainnya. Hal ini menunjukkan bahwa nilai *F1-score* pada kelas *viral pneumonia* lebih beragam dan bervariasi.



Gambar 7. Boxplot (a) presisi (b) *recall* (c) spesifisitas dan (d) *F1-score* model ResNet-50

Pada Gambar 7(a), menunjukkan median presisi tertinggi secara berurutan adalah kelas *viral pneumonia*, COVID-19, normal dan *lung opacity*. Hal ini menunjukkan bahwa model memiliki kinerja yang lebih baik dalam mendeteksi kasus *viral pneumonia* dengan tingkat presisi yang lebih tinggi dibandingkan kelas lainnya.

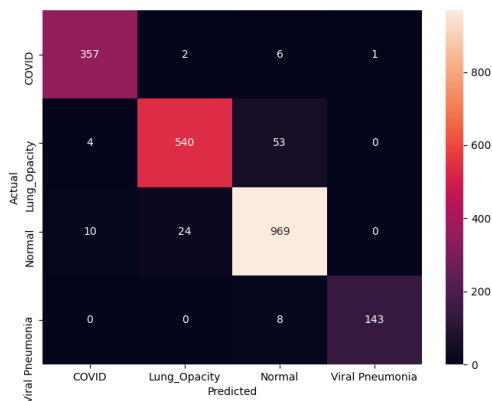
Pada Gambar 7(b), terlihat bahwa median *recall* pada kelas normal lebih tinggi dibandingkan kelas lainnya. Namun pada kelas COVID-19 terlihat bahwa terdapat rentang nilai *recall* antara kuartil tiga (Q3) hingga batas atas. Sedangkan pada kelas normal, terdapat rentang nilai *recall* di atas Q3 hingga batas atas dan rentang nilai *recall* antara kuartil pertama (Q1) dan batas bawah. Hal ini mengindikasikan terdapat variasi dalam hasil *tuning* pada kelas normal, di mana beberapa hasil *tuning* menghasilkan *recall* yang lebih rendah. Sementara itu, kelas COVID-19 menunjukkan konsistensi dalam mendeteksi kasus dengan tingkat *recall* yang lebih tinggi.

Nilai spesifisitas masing-masing kelas pada Gambar 7(c) menunjukkan bahwa kelas COVID-19 dan viral pneumonia memiliki median yang sama dan hasil konsisten.

Pada Gambar 7(d) menunjukkan median *F1-*

*score* pada kelas COVID-19 secara signifikan lebih tinggi dibandingkan kelas lainnya. Hal ini menunjukkan bahwa model memiliki keseimbangan yang lebih baik antara presisi dan *recall* ketika memprediksi kasus COVID-19.

Berdasarkan Gambar 6 dan Gambar 7, disimpulkan bahwa model ResNet-50 dengan modifikasi pada lapisan asli, yaitu penambahan lapisan *dropout* dengan peluang 0.5, menunjukkan kinerja yang lebih baik dibandingkan dengan model VGGNet-19. Hasil tersebut terlihat dari nilai median yang lebih tinggi pada setiap kelas dalam model ResNet-50 jika dibandingkan dengan model VGGNet-19. Selanjutnya, model terbaik ResNet-50 digunakan untuk melakukan analisis lebih lanjut terhadap kelas yang terprediksi dengan benar dan salah. Hal ini dapat dilakukan dengan melihat hasil pada *confusion matrix*.



Gambar 8 *Confusion Matrix* Model Terbaik (ResNet-50)

Gambar 8 menunjukkan bahwa masing-masing kelas terdapat misklasifikasi. Pada kelas COVID-19 terdapat 357 citra (98%) dari 366 keseluruhan citra yang diuji mampu di prediksi dengan tepat. Namun, terdapat 9 citra mengalami misklasifikasi, yaitu 2 citra salah diprediksi sebagai kelas *lung opacity*, 6 citra salah diprediksi masuk kelas normal dan 1 citra salah diprediksi masuk kelas *viral pneumonia*. Kelas *lung opacity* diperoleh 59 citra miskalsifikasi. Dari citra tersebut, 4 citra salah diprediksi sebagai kelas COVID-19 dan 53 citra salah diprediksi sebagai kelas normal. Total citra yang diuji adalah 597 citra dan 540 citra lainnya diprediksi tepat. Selanjutnya pada kelas normal juga mengalami misklasifikasi sebanyak 34 citra. Dari citra-citra tersebut, 10 citra salah diprediksi sebagai kelas COVID-19 dan 24 citra salah diprediksi sebagai kelas *lung opacity*. Sementara itu, pada kelas *viral pneumonia* didapatkan 143 citra (95%) diprediksi tepat dari total 151 citra yang diuji dan sisanya 8 citra diprediksi normal.





Misklasifikasi dapat terjadi karena beberapa alasan, termasuk ketidakseimbangan jumlah dataset antara kelas-kelas yang ada. Ketika jumlah dataset pada kelas normal lebih banyak, hal ini dapat mempengaruhi kemampuan model dalam membedakan dengan tepat antara kelas normal

dengan kelas lainnya. Model cenderung lebih familiar dan terlatih dengan fitur-fitur yang mewakili kelas normal, sementara fitur-fitur yang mewakili kelas lain mungkin tidak memiliki representasi yang kuat dalam model. Inilah salah satu yang dapat menyebabkan terjadinya misklasifikasi pada kelas normal.

### 3.6 Deteksi

Tujuan dari deteksi adalah untuk mengenali dan mengidentifikasi kelas data yang belum diketahui berdasarkan pola citra, dengan menampilkan persentase akurasi pada setiap kelas. Model yang digunakan adalah model ResNet-50 karena memiliki performa lebih baik dari VGGNet-19 pada kasus ini.

Tabel 6. Hasil Prediksi ResNet-50

<p>Normal</p>  <p>Covid-19: 0.08% Lung_Opacity: 2.48% Normal: 97.46% Viral_Pneumonia: 0.00%</p>	<p>Lung opacity</p>  <p>Covid-19: 0.00% Lung_Opacity: 87.12% Normal: 2.50% Viral_Pneumonia: 0.00%</p>
<p>COVID-19</p>  <p>Covid-19: 96.32% Lung_Opacity: 0.00% Normal: 1.10% Viral_Pneumonia: 0.00%</p>	<p>Viral pneumonia</p>  <p>Covid-19: 0.00% Lung_Opacity: 0.00% Normal: 1.10% Viral_Pneumonia: 94.20%</p>

Citra – citra tersebut dimasukkan pada sistem dan menerapkan model *ResNet-50* untuk melakukan prediksi. Berdasarkan Tabel 6 menghasilkan prediksi seluruh label dengan benar. Citra yang memiliki kelas normal diprediksi normal dengan akurasi sebesar 97.46%. Kemudian, menggunakan model ResNet-50 juga berhasil memprediksi citra *lung opacity* sebagai *lung opacity* dengan akurasi 87.12%. Citra yang diklasifikasikan sebagai kelas COVID-19 berhasil diprediksi sebagai COVID-19 dengan akurasi 96.32% dan citra *viral pneumonia* berhasil diprediksi sebagai *viral pneumonia* dengan akurasi yang diperoleh sebesar 94.20%.

### 4. KESIMPULAN DAN SARAN

Hasil modifikasi dengan penambahan lapisan teknik regularisasi *dropout* dengan peluang 50% pada arsitektur VGGNet-19 dan ResNet-50 mampu mengatasi *overfitting* dan meningkatkan ukuran kebaikan model. Performa model masing-masing arsitektur memberikan hasil lebih baik pada jumlah

data yang lebih besar. Performa terbaik diperoleh pada model arsitektur ResNet-50 dengan nilai rata-rata presisi, *recall*, spesifistas dan *F-1 score* berturut-turut sebesar 95.5%, 94.1%, 97% dan 94.8%. Hasil penelitian menunjukkan tingkat akurasi sebesar 94.4%. Hasil ini dapat dipertimbangkan untuk penerapannya pada kecerdasan buatan (*artificial intelligent*), diharapkan dapat digunakan sebagai pendukung dalam proses diagnosa klinis penyakit COVID-19 atau *viral pneumonia* dan *lung opacity* pada populasi yang lebih besar.

Terdapat beberapa hal yang dapat diperhatikan dalam pengembangan penelitian selanjutnya. Pertama, direkomendasikan untuk menerapkan teknik-teknik penanganan ketidakseimbangan dataset seperti *oversampling* atau *undersampling*. Kedua, penerapan metode LIME (*Local Interpretable Model-Agnostic Explanations*) yang dapat digunakan untuk menginterpretasikan keputusan yang diambil oleh model *deep learning* dengan melihat ciri-ciri yang signifikan pada citra. Ketiga, mempertimbangkan penggunaan metode optimasi *hyperparameter* otomatis, seperti menggunakan algoritma pencarian *grid* atau teknik optimasi berbasis algoritma genetika. Keempat, menggabungkan beberapa metode *deep learning*, seperti penggunaan *ensemble learning*. Pengembangan selanjutnya dapat mengoptimalkan proses komputasi dengan menggunakan perangkat keras yang lebih canggih dan perangkat lunak yang efisien dapat mempercepat waktu komputasi dan memungkinkan eksperimen lebih luas.

### DAFTAR PUSTAKA

- APOSTOLOPOULOS, I.D., MPESIANA, T.A. 2020. Covid-19: Automatic Detection from X-Ray Images Utilizing Transfer Learning with Convolutional Neural Network. *Phys Eng Sci Med* 43, 635-640. <https://doi.org/10.1007/s13246-020-00865-4>
- BEJIGA MB, ZEGGADA A, NOUFFIDJ A, MELGANI F. 2017. A Convolutional Neural Network Approach for Assisting Avalanche Search and Rescue Operations with UAV Imagery. *Remote Sensing*. 9(2):100. <https://doi.org/10.3390/rs9020100>
- CHOWDHURY. M E H, TAWSIFUR R, AMITH K, dkk. 2020. Can AI Help in Screening Viral and COVID-19 Pneumonia? *IEEE Access*, Vol. 8, 2020, pp. 132665 – 132676. <https://doi.org/10.48550/arXiv.2003.13145>
- FANG, Y., ZHANG, H., XIE, J., LIN, M., YING, L., PANG, P., & JI, W. 2020. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology*, 296(2), E115–E117. <https://doi.org/10.1148/radiol.2020200432>
- GARILLOS MANLIGUEZ, C.A. 2016. Generalized Confusion Matrix for Multiple Classes. [online] Tersedia di:

- <https://www.researchgate.net/publication/310799885>. [Diakses 12 Februari 2023]
- IKECHUKWU, A. VICTOR, S. MURALI, R. DEEPU, R.C. SHIVAMURTHY. 2021. Resnet-50 Vs VGG-19 Vs Training from Scratch: A Comparative Analysis of The Segmentation and Classification of Pneumonia from Chest X-Ray Images. *Global Transitions Proceedings*. Volume 2 Issue 2, 375-381. <https://doi.org/10.1016/j.gltp.2021.08.027>
- LONG, C., XU, H., SHEN, Q., ZHANG, X., FAN, B., WANG, C., ZENG, B., LI, Z., LI, X., & LI, H. 2020. Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT?. *European Journal of Radiology*, 126, 108961. <https://doi.org/10.1016/j.ejrad.2020.108961>
- RAHMAN T., KHANDAKAR A., QIBLAWEY Y., TAHIR A., KIRANYAZ S., KASHEM SBA., ISLAM MT., MADEED SA., ZUGHAIER SM., KHAN MS., dan CHOWDHURY ME. 2020. Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images. <https://doi.org/10.48550/arXiv.2012.02238>
- RAHMAN T, 2021. COVID-19 Radiography Database, [online] Tersedia di: <https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database>. [Diakses 19 Juli 2021]
- RISNANDAR. 2022. DeSa COVID-19: Deep salient COVID-19 image-based quality assessment. *Journal of King Saud University - Computer and Information Sciences*. Volume 34, Issue 10, Part B, 9501-9512. <https://doi.org/10.1016/j.jksuci.2021.11.013>
- S. ALBAWI, T. A. MOHAMMED AND S. AL-ZAWI. 2017. Understanding of a convolutional neural network. *International Conference on Engineering and Technology (ICET)*, Antalya, Turkey, 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.
- SAKIB et al. 2020. Detection of COVID-19 Disease from Chest X-Ray Images: A Deep Transfer Learning Framework. *medRxiv*. <https://doi.org/10.1101/2020.11.08.20227819>
- SCHMIDHUBER, J. 2015. Deep learning in neural networks: An overview. *Neural Network*. 61, 85-117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- SHIBLY K, HASSAN S.K., MD TAHZIB-UL-ISLAM D., MD MAHBUBUR RAHMAN. 2020. COVID Faster R-CNN: A Novel Framework to Diagnose Novel Coronavirus Disease (COVID-19) in X-Ray Images. *Informatics in Medicine Unlocked*. Volume 20 100405. <https://doi.org/10.1016/j.imu.2020.100405>
- SINGH, D., KUMAR, V., VAISHALI, & KAUR, M. 2020. Classification Of COVID-19 Patients from Chest CT Images using Multi-Objective Differential Evolution-Based Convolutional Neural Networks. *European Journal of clinical microbiology & infectious diseases: official publication of the European Society of Clinical Microbiology*, 39(7), 1379–1389. <https://doi.org/10.1007/s10096-020-03901-z>
- SWASTIKA, WINDRA. 2020. Studi Awal Deteksi COVID-19 menggunakan Citra CT Berbasis Deep Learning. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)* Vol. 7, No. 3, hlm. 629-634. <http://dx.doi.org/10.25126/jtiik.2020733399>
- WORLD HEALTH ORGANIZATION. 2020b. Laboratory Testing for Coronavirus Disease (COVID-19) In Suspected Human Cases. WHO – Interim Guidance, 19 March, 1-7
- WORLDMETERS. World Population Clock. <https://www.worldmeters.info>. 2023. [Diakses 14 Maret 2023]
- YING X. 2019. An Overview of Overfitting and its Solutions. *J Phys Conf*. 1168(2):1- 6. <https://dx.doi.org/10.1088/1742-6596/1168/2/022022>
- ZHANG. H, LIN HAN, KE CHEN, Y. PENG, J. LIN. 2020. Diagnostic Efficiency of the Breast Ultrasound Computer-Aided Prediction Model Based on Convolutional Neural Network in Breast Cancer. *Journal of Digital Imaging*. 1-6. <https://doi.org/10.1007/s10278-020-00357-7>
- ZHONG. YI. 2020. Using Deep Convolutional Neural Networks to Diagnose COVID-19 From Chest X-Ray Images. <https://doi.org/10.48550/arXiv.2007.09695>