

## Algoritma *Deep Learning* untuk Pengklasifikasian Penyakit Radang Paru-Paru pada Citra *Chest X-Ray* dengan *Convolutional Neural Network*

Putri Rahmadewi<sup>1)</sup>, Juli Yanti Kartika Harahap<sup>2)</sup>, Evta Indra<sup>3)\*</sup>

<sup>1)2)3)</sup> Fakultas Teknologi dan Ilmu Komputer, Universitas Prima Indonesia

\*Correspondence Author: [evtaindra@unprimdn.ac.id](mailto:evtaindra@unprimdn.ac.id), Medan, Indonesia

DOI: <https://doi.org/10.37012/jtik.v9i1.1371>

### Abstrak

*Radang paru-paru (Pneumonia)* adalah salah satu penyakit paru-paru yang menyerang saluran pernapasan yang menyebabkan penderitanya kesulitan untuk bernapas dikarenakan terjadinya pembengkakan pada paru-paru. Tindakan yang dilakukan oleh dokter untuk mengetahui seseorang mengalami radang paru-paru adalah dengan melakukan rontgen pada dada pasien dan menganalisa hasil dari foto rontgen atau gambar citra *chest x-ray* pasien secara manual. Oleh karena itu, dilakukan penelitian untuk mengklasifikasikan penyakit radang paru-paru melalui citra *x-ray* dengan menggunakan *algoritma deep learning* yaitu *Convolutional Neural Network (CNN)*. Pada metode *Convolutional Neural Network (CNN)* ini digunakan 5266 data citra *chest x-ray* sebagai data *training* dan 592 data citra *chest x-ray* sebagai data *testing*, dimana dua data tersebut sudah terbagi menjadi citra *chest x-ray* normal dan citra *chest x-ray* radang paru-paru. Pada penelitian ini diperoleh hasil berupa *Accuracy* sebesar 92.22 %, *Recall* sebesar 98.41 %, *Precision* sebesar 90.29% dan *f1-score* sebesar 94.17%. Penelitian ini juga menunjukkan *confusion matrix* bahwa model memiliki tingkat negatif palsu yang rendah, tetapi tingkat positif palsu yang tinggi. Hal tersebut lebih baik dari pada terlalu banyak citra *chest x-ray* yang tidak terdeteksi penyakit radang paru-paru.

**Kata Kunci:** Radang Paru-Paru (*Pneumonia*), Citra *chest X-ray*, *Convolutional Neural Network*

### Abstract

*Lung Inflammation (Pneumonia)* is a lung disease that attacks the respiratory tract which causes the sufferer to have difficulty breathing due to swelling of the lungs. The action taken by a doctor to find out if someone has pneumonia is to take an X-ray of the patient's chest and analyze the results of the X-ray or chest X-ray image of the patient manually. Therefore, a study was conducted to classify pneumonia through x-ray images using a deep learning algorithm, namely the *Convolutional Neural Network (CNN)*. In the *Convolutional Neural Network (CNN)* method, 5266 chest x-ray image data are used as training data and 592 chest x-ray image data as testing data, where the two data have been divided into normal chest x-ray images and chest x-ray images with pneumoniae. In this study the results obtained were *Accuracy* of 92.22%, *Recall* of 98.41%, *Precision* of 90.29% and *f1-score* of 94.17%. This study also shows that the *Confusion Matrix* model has a low false negative rate, but a high false positive rate. That's better than too many chest X-rays that don't detect pneumonia.

**Keywords:** Lung Inflammation (*Pneumonia*), Chest X-Ray Image, *Convolutional Neural Network*

## PENDAHULUAN

Pengobatan pada penyakit radang paru-paru tergantung pada jenis, penyebab dan tingkat keparahannya (Resna, 2020). Investigasi tertentu juga diperlukan sebelum memutuskan prosedur pengobatan penyakit radang paru-paru seperti rontgen dada, pemeriksaan darah, pemeriksaan dahak dan pemeriksaan kadar oksigen darah (Belinda, 2019). Meskipun status inflamasi paru pada pasien dapat dilihat dengan citra CT-Scan (*Computed Tomography-Scan*), citra MRI (*Magnetic Resonance Imaging*) dan citra USG

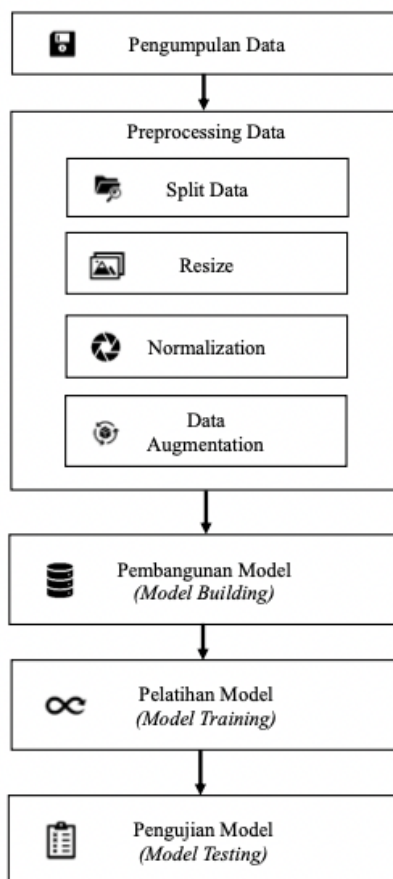
(*Ultrasonografi*) (Ekananda & Rimirasih, 2022). Gambar yang dihasilkan tidak selalu berkualitas baik, cenderung kabur, dan berbeda antar jenis penyakit paru, mungkin terdapat kesamaan yang serupa antara penyakit radang paru-paru, TBC, *pneumotoraks*, infiltrasi, bintil, dan kanker paru-paru (Maysanjaya, 2020b). Hal tersebut dikarenakan petugas khusus kesehatan secara kasat mata mengamati foto rontgen untuk mengidentifikasi penyakit seperti radang paru-paru dan TBC, sehingga menyulitkan dalam identifikasi jenis penyakit yang diderita pasien (*Klasifikasi Penyakit Paru-Paru Berbasis Pengolahan Citra X Ray Menggunakan Convolutional Neural Network (Classification Of The Lung Diseases Based On X Ray Image Processing Using Convolutional Neural Network)*, n.d.). Oleh sebab itu, kesalahan dapat terjadi karena kurangnya pengalaman dan pengetahuan khususnya dokter muda (Andhi et al., 2021). Mengingat perlunya keakuratan dan ketepatan diagnosis yang diberikan kepada pasien, ada potensi risiko positif palsu dan negatif palsu dalam diagnosis pasien (Ferdinand et al., 2021).

Penelitian yang dilakukan sebelumnya adalah melakukan penelitian untuk mengklasifikasikan pneumonia pada rontgen paru menggunakan *convolutional neural network*. Namun penelitian ini memiliki kelemahan yaitu model CNN yang dibangun memiliki kondisi *overfitting*. Dengan kata lain, model dalam keadaan non-fungsional, seperti halnya data uji (Maysanjaya, 2020a).

Metode CNN sendiri dipilih karena penggunaan *deep learning* khususnya *Convolutional Neural Network (CNN)* (Yopento & Coastera, 2022), telah berhasil melampaui *Machine Learning* tradisional dalam klasifikasi citra dalam beberapa tahun terakhir, tidak terkecuali citra medis. CNN pada dasarnya adalah susunan multi-layer yang terdiri dari *Convolution Layer*, *Pooling Layer* dan *Fully Connected Layer* (Yopento et al., 2022). Berdasarkan pemaparan tersebut, perlu sebuah penelitian untuk merancang sebuah model algoritma *deep learning* untuk klasifikasi penyakit radang paru-paru melalui citra radiografi dengan penerapan metode *Convolutional Neural Network (CNN)* yang diharapkan dapat menjadi model untuk membantu ahli medis di dunia kesehatan.

## METODE

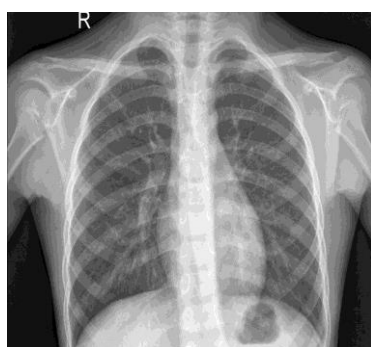
Penelitian dilakukan pada ruangan radiologi RSUD Sari Mutiara Lubuk Pakam. Notebook atau perangkat keras yang digunakan didalam penelitian ini ada Jupyter Anaconda. Agar penelitian ini berjalan dengan lancar dan selesai tepat waktu dibutuhkan sebuah prosedur kerja. Prosedur kerja dari penelitian ini dapat dilihat pada Gambar 1.



**Gambar 1.** Prosedur Kerja

### Pengumpulan Data

Datasets citra chest x-ray diperoleh dari dari RSU Sari Mutiara Lubuk Pakam dalam bentuk dataset citra chest x-ray. Dataset citra chest x-ray terdapat total gambar 5858 gambar dengan 1573 gambar normal atau tidak menunjukkan tanda-tanda radang paru-paru dan 4285 gambar menunjukkan tanda-tanda radang paru-paru. Contoh sample dari citra tersebut ditampilkan pada Gambar 2.



Normal



Radang Paru-Paru

**Gambar 2.** Perbandingan *Chest X-Ray* Normal dengan Radang Paru-Paru

Pada gambar diatas terlihat bahwa rontgen dada atau paru-paru normal di dominasi dengan gambar berwarna hitam sebagai pertanda bahwa sirkulasi udara didalam paru-paru bekerja dengan baik, sedangkan pada gambar dengan penyakit radang paruparu terlihat bercak putih didalam paru paru.

### **Image Processing (*Preprocessing*)**

Tahap *Preprocessing* data merupakan tahap untuk mempersiapkan *datasets* citra radiografi dada. Tahapan ini bertujuan untuk memproses dataset sebelum dilakukan proses selanjutnya, karena dataset yang dikumpulkan tidak bisa langsung dimengerti oleh CNN untuk itu perlu dilakukan *preprocessing*. (Muhammadiyah Jember et al., 2022)

Tahapan yang dilakukan didalam *preprocessing* diantaranya adalah *split data*, *resize*, *normalization* dan *data augmentation*. Tahapan tersebut dilakukan agar dataset dapat diproses dengan baik oleh model CNN dan mempercepat proses training maupun testing (Varshni et al., 2019) (Muhammadiyah Jember et al., 2022).

### **Pembangunan Model (*Model Building*)**

Pembangunan model CNN merupakan tahapan membangun model *deep learning* CNN. (Muhammadiyah Jember et al., 2022) Dimana hasil dari model building ini akan digunakan untuk proses training dan testing model klasifikasi CNN.

### **Pelatihan Model (*Model Training*)**

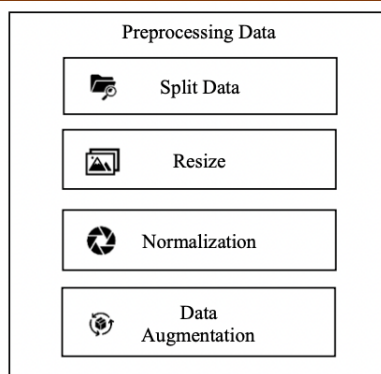
Pelatihan model adalah fase di mana model CNN mengenali karakteristik data yang diajarkan melalui pembelajaran kelas, sehingga model memiliki kemampuan klasifikasi pada data yang serupa.

### **Pengujian Model (*Model Testing*)**

Pengujian model atau model testing digunakan untuk mengetahui tingkat akurasi (*accuracy*) dan tingkat kehilangan (*loss*) rata-rata pada setiap model (Perdananto et al., n.d.). Pada proses testing ini untuk mengukur tingkat keakuratan dari hasil pengujian model menggunakan metrik kinerja (*performance metrics*). Persamaan dari metrik kinerja tersebut diturunkan dari *Confusion Matrix* (Ferdinand et al., 2021).

## **HASIL DAN PEMBAHASAN**

Dalam melakukan *image processing* atau *preprocessing* pada penelitian ini terdapat beberapa tahapan antara lain *split data*, *resize* dan *normalization*. Gambar 3 merupakan tahapan dari *preprocessing*.



**Gambar 3.** Tahapan *Preprocessing*

### Split Data

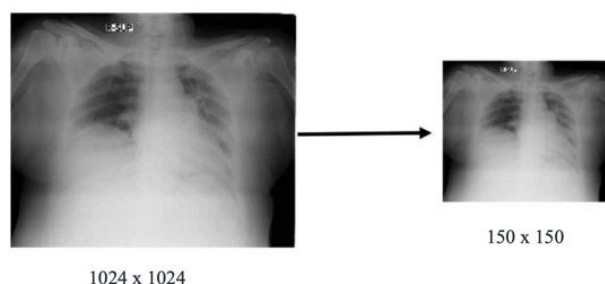
Dataset citra *chest x-ray* akan dipisah menjadi dua data yaitu data training atau data validasi. Dimana data training untuk digunakan pada proses training atau pelatihan model. Jumlah data dari masing-masing data dapat dilihat pada Tabel 1.

**Tabel 1.** Pembagian Dataset

Nama Data	Jumlah
Data Train	5266
Data Test	592
<b>Total</b>	<b>5858</b>

### Resize

Pada penelitian ini ukuran gambar citra chest x-ray asli memiliki ukuran gambar yang besar dan berbeda-beda, sehingga untuk mempermudah proses pengolahan data dilakukan *resize*, dimana ukuran semua gambar tersebut diubah menjadi 150 x 150 pixel. Kemudian dikonversi menjadi float 32. Ukuran tersebut dipilih karena mempertimbangkan jika semakin besar ukuran citra radiografi maka penggunaan memori juga semakin besar dan pemrosesan akan membutuhkan waktu yang cukup lama.



**Gambar 4.** Ilustrasi *Resize* Citra *Chest X-Ray*

## Normalization

Pixel pada citra chest x-ray terdapat diantara 0 hingga 255. Oleh sebab itu, untuk mempercepat model CNN mencapai konvergen dan generalisasi, maka perlu dilakukan proses *normalization*. Proses *normalization* pada penelitian ini adalah dengan mengubah nilai pixel pada rentang nilai 0 sampai 1, dengan cara membagi pixel dengan 255, Ketika nilai pixelnya berada pada rentang nilai 0 dan 1.

## Data Augmentation

Dibawah ini merupakan proses *data augmentation* yang dilakukan pada penelitian ini:

- *Shear Range* = 0.2 , menggeser data citra hingga 20%.
- *Zoom Range* = 0.2, memperbesar data citra dengan acak hingga 20%.
- *Horizontal Flip* = True, memutar data citra secara horizontal.

```
train_datagen = ImageDataGenerator(  
    rescale = 1./255,  
    shear_range = 0.2,  
    zoom_range = 0.2,  
    horizontal_flip = True,  
)
```

Gambar 5. Source Code Data Augmentation

## Pembangunan Model (*Model Building*)

Pembangunan model (*Model Building*) menggunakan algoritma deep learning Convolutional Neural Networ (CNN). Arsitektur model CNN pada penelitian ini menggunakan *fully connected layer* (3 layer) dan terdiri dari 5 blok dengan beberapa komponen diantaranya adalah *convolution layer* 2D (conv2D), *max-pooling layer* (MaxPool2D), *batch normalization layer* (BatchNormalization) dan *separable convolution layer* 2D (separable\_conv2d). Tabel 2 menunjukkan lapisan-lapisan arsitektur CNN yang telah terbentuk. Pada pembangunan model CNN ini menggunakan ukuran citra sebesar 150 x 150. Pada proses *convolutional* ditetapkan ukuran kernel 3 x 3, untuk *maxpooling* ditetapkan ukuran pool 2 x 2. Untuk blok 4 dan blok 5 dilakukan dropout sebesar 0.2. Sedangkan untuk *fully connected layer* dilakukan dropout 0.5. Tujuannya adalah untuk terhindar dari kondisi overfitting.

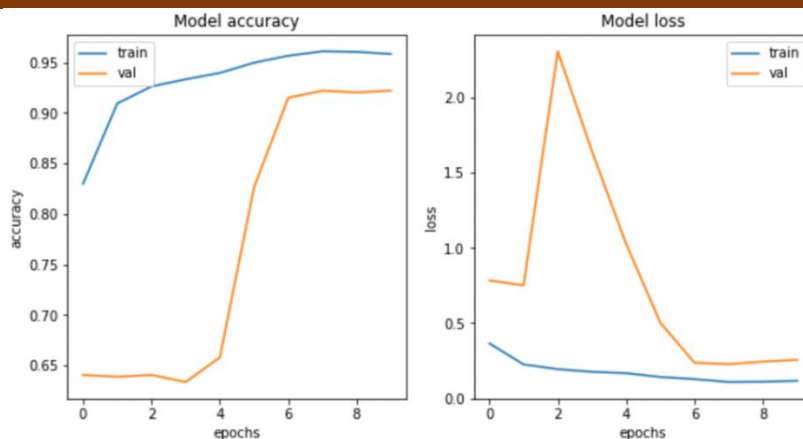
separable_conv2d_4 (SeparableConv2D)	(None, 18, 18, 128)	8896
separable_conv2d_5 (SeparableConv2D)	(None, 18, 18, 128)	17664
batch_normalization_2 (BatchNormalization)	(None, 18, 18, 128)	512
max_pooling2d_3 (MaxPooling2D)	(None, 9, 9, 128)	0
dropout (Dropout)	(None, 9, 9, 128)	0
separable_conv2d_6 (SeparableConv2D)	(None, 9, 9, 256)	34176
separable_conv2d_7 (SeparableConv2D)	(None, 9, 9, 256)	68096
batch_normalization_3 (BatchNormalization)	(None, 9, 9, 256)	1024
max_pooling2d_4 (MaxPooling2D)	(None, 4, 4, 256)	0
dropout_1 (Dropout)	(None, 4, 4, 256)	0
flatten (Flatten)	(None, 4096)	0
dense (Dense)	(None, 512)	2097664
dropout_2 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 128)	65664
dropout_3 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 64)	8256
dropout_4 (Dropout)	(None, 64)	0
dense_3 (Dense)	(None, 1)	65
=====		
Total params: 2,314,337		
Trainable params: 2,313,377		
Non-trainable params: 960		

**Gambar 6.** Summary Model CNN

### **Pelatihan Model (*Model Training*)**

Setelah *preprocessing* telah selesai dilakukan dan model *Convolutional Neural Network* untuk klasifikasi telah berhasil dibuat pada tahapan *model building*, selanjutnya adalah melakukan proses *model training*. Proses *model training* ini bertujuan untuk melatih model dari algoritma *Convolutional Neural Network* (CNN) dalam memahami pola setiap citra yang telah dipreprocessing. Dataset untuk proses model training penelitian ini menggunakan dataset yang telah dipreprocessing sebelumnya. Data digunakan sebanyak 5266 dataset citra *chest x-ray* yang telah dibagi menjadi 2 jenis yaitu citra *chest x-ray* normal dan citra *chest x-ray* radang paru-paru. Proses model training dilakukan dengan menggunakan batch size 32 dan epoch 10. Epoch bertujuan untuk mengetahui berapa kali algoritma *deep learning* bekerja melewati seluruh data training, tujuan menggunakan beberapa nilai epoch adalah untuk mendapatkan nilai akurasi yang optimal. Grafik akurasi dan grafik loss hasil proses training model dengan menggunakan epoch 10 ditunjukkan pada gambar dibawah ini.





**Gambar 7.** Grafik Akurasi Train dan Loss

### Pengujian Model (*Model Testing*)

Pada penelitian ini, tahapan pengujian (testing) model digunakan untuk mengetahui tingkat akurasi (*accuracy*) dan tingkat kehilangan (*loss*) rata-rata pada setiap model. Proses testing ini menggunakan dataset testing dengan jumlah data 592 citra *chest x-ray*.

*Akurasi*

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

\* 100%

$$Presisi = \frac{TN}{TN + FN} * 100\% \quad (2)$$

$$Recall = \frac{TN}{TN + FP} * 100\% \quad (3)$$

**Gambar 8.** Rumus Perhitungan *Testing Model*

Selanjutnya mengukur tingkat keakuratan dari hasil pengujian model menggunakan metrik kinerja (*performance metrics*). Metrik kinerja merupakan persamaan (1), (2) dan (3) yang digunakan untuk menghitung nilai dari *accuracy*, *recall* dan *precision*. Persamaan tersebut diturunkan dari *Confusion Matrix* yang terdiri dari nilai *True Normal* (TN), *True Pneumonia* (TP), *False Pneumonia* (FP) dan *False Normal* (FN). Hasil dari proses testing ditunjukkan pada Gambar 9.

```

CONFUSION MATRIX -----
[[174  40]
 [  6 372]]

TEST METRICS -----
Accuracy: 92.22972972972973%
Precision: 90.29126213592234%
Recall: 98.4126984126984%
F1-score: 94.17721518987342

TRAIN METRIC -----
Train acc: 95.82
    
```

**Gambar 9.** Hasil Perhitungan Testing



---

## KESIMPULAN DAN REKOMENDASI

Pada penelitian ini dapat disimpulkan bahwa algoritma *deep learning Convolutional Neural Network* (CNN) dapat melakukan pengklasifikasian penyakit radang paru-paru melalui citra *chest x-ray*. Dataset yang digunakan sebesar 5266 citra *chest x-ray* untuk data training dan 592 untuk data validasi dan testing. Hasil dari pengujian model menunjukkan bahwa tingkat akurasi sebesar 92.22%. Model juga memiliki nilai presisi sebesar 90.29% dan *recall* 98.41% dan *f1-score* sebesar 94.17%. Pada pengujian juga menunjukkan *confusion matrix* bahwa model memiliki tingkat negatif palsu yang rendah, tetapi tingkat positif palsu yang tinggi. Hal tersebut lebih baik dari pada terlalu banyak citra *chest x-ray* yang tidak terdeteksi penyakit radang paru-paru. Secara keseluruhan berdasarkan hasil yang diperoleh menunjukkan bahwa model algoritma CNN dengan tingkat akurasi yang tinggi lebih dari 90% dapat mengklasifikasikan penyakit radang paru-paru dengan cukup baik.

Penelitian selanjutnya disarankan dapat menambahkan citra *chest x-ray* yang digunakan untuk data train dan data testing dengan menggunakan metode optimasi lainnya, hal itu dilakukan untuk menambah tingkat akurasi dalam klasifikasi gambar. Penelitian selanjutnya juga dapat dilakukan untuk klasifikasi penyakit yang berbeda. Selain itu, model yang telah dibuat mungkin bisa diimplementasikan kedalam aplikasi website atau mobile di penelitian selanjutnya agar lebih memudahkan petugas kesehatan dalam menentukan penyakit radang paru-paru.

## REFERENSI

- Andhi, P., Pratama, H., Teguh, R., Sahay, A. S., Wilentine, V. Deteksi COVID-19 Berdasarkan Hasil Rontgen Dada (Chest Xray) Menggunakan Python. In *JOINTECOMS (Journal of Information Technology and Computer Science) p-ISSN: xxxx-xxxx* (Vol. 1, Issue 1).
- Belinda, G. (2019, February 22). *Pneumonia: Penyebab, Gejala, dan Pengobatan*. <https://www.honestdocs.id/pneumonia>.
- Ekananda, N. P., & Rimirasih, D. (2022). Identifikasi Penyakit Pneumonia Berdasarkan Citra Chest X-Ray Menggunakan Convolutional Neural Network. *Jurnal Ilmiah Informatika Komputer*, 27(1), 79–94. <https://doi.org/10.35760/ik.2022.v27i1.6487>
- Ferdinand, R., Setyawan, E., & Rere, L. M. R. (2021). Convolutional Neural Network Untuk Deteksi Covid-19 Melalui Citra Chest X-Ray. *Seminar Nasional Teknologi Informasi Dan Komunikasi STI&K (SeNTIK)*, 5(1).

---

*Klasifikasi Penyakit Paru-Paru Berbasis Pengolahan Citra X Ray Menggunakan Convolutional Neural Network (Classification Of The Lung Diseases Based On X Ray Image Processing Using Convolutional Neural Network).* (n.d.).

- Maysanjaya, I. Md. D. (2020a). Klasifikasi Pneumonia pada Citra X-rays Paru-paru dengan Convolutional Neural Network. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 9(2), 190–195. <https://doi.org/10.22146/jnteti.v9i2.66>
- Maysanjaya, I. Md. D. (2020b). Klasifikasi Pneumonia pada Citra X-rays Paru-paru dengan Convolutional Neural Network. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 9(2), 190–195. <https://doi.org/10.22146/jnteti.v9i2.66>
- Muhammadiyah Jember, U., Trisetya Indrawan, G., Nilogiri, A., & Azizah Al Faruq, H. (2022). Diagnosis COVID-19 Berdasarkan Citra X-ray Paru-Paru Menggunakan Metode Convolutional Neural Network Diagnose Of COVID-19 Based On X-ray Image Of The Lungs Using Convolutional Neural Network. In *Jurnal Smart Teknologi* (Vol. 3, Issue 4). <http://jurnal.unmuhjember.ac.id/index.php/JST>
- Perdananto, A., Udin Zailani, A., Kencana No, J., & Tangerang Selatan, P. (n.d.). Penerapan Deep Learning Pada Aplikasi Prediksi Penyakit Pneumonia Berbasis Convolutional Neural Networks. *DES 2019 Journal of Informatics and Communications Technology*, 1(2), 1–010.
- Resna, N. (2020, November 23). *Berbagai Pengobatan Pneumonia Berdasarkan Penyebab dan Tingkat Keparahannya*. <https://www.sehatq.com/artikel/jenis-antibiotik-untuk-pneumonia-dan-pilihan-pengobatan-lainnya>.
- Varshni, D., Nijhawan, R., Thakral, K., Mittal, A., & Agarwal, L. (2019). Pneumonia Detection Using CNN based Feature Extraction. In *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*.
- Yopento, J., & Coastera, F. (2022). Identifikasi Pneumonia Pada Citra X-Ray Paru-Paru Menggunakan Metode Convolutional Neural Network (CNN) Berdasarkan Ekstraksi Fitur Sobel. In *Jurnal Rekursif* (Vol. 10, Issue 1). <http://ejournal.unib.ac.id/index.php/rekursif/40>
- Yopento, J., Ernawati, E., & Coastera, F. F. (2022). Identifikasi Pneumonia Pada Citra X-Ray Paru-Paru Menggunakan Metode Convolutional Neural Network (CNN) Berdasarkan Ekstraksi Fitur Sobel. *Rekursif: Jurnal Informatika*, 10(1), 40–47. <https://doi.org/10.33369/rekursif.v10i1.17247>