



ANALISIS PERBANDINGAN PENYAKIT PARU-PARU DARI CITRA X-Ray DADA MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK DAN SUPPORT VECTOR MACHINE

Ningsiah¹, Suhendro Yusup², Joko Triloka³

^{1,2,3}Fakultas Komputer, Institut Informatika & Bisnis Darmajaya

Email: caesarintanputri@gmail.com¹, suhendro@darmajaya.ac.id², triloka@darmajaya.ac.id³

Abstrak

Paru-paru merupakan organ tubuh manusia yang penting dalam tubuh manusia terutama dalam sistem pernafasan. Fungsi lain paru-paru menjaga suhu tubuh agar tetap stabil melindungi tubuh dari zat berbahaya, hidung sebagai indra penciuman, namun terkadang paru-paru akan mengalami kondisi dimana tidak berfungsi secara normal. Gambar rontgen dada adalah metode klinis yang paling terkenal untuk diagnosis penyakit paru-paru. Namun, mendiagnosis penyakit paru-paru dari gambar rontgen dada merupakan tugas yang menantang bahkan untuk ahli radiologi. Penelitian ini mengusulkan suatu sistem yang dapat digunakan untuk Analisis Perbandingan penyakit paru-paru dengan menerapkan metode *Convolutional Neural Network* dan *Support Vector Machine*. CNN adalah metode di bidang pengenalan objek yang memiliki lapisan khusus, yaitu lapisan konvolusi dan lapisan *pooling* yang memungkinkan proses pembelajaran fitur dengan baik. SVM adalah metode Analisis perbandingan yang mengandalkan hasil dari teori pembelajaran statistik untuk menjamin kinerja generalisasi. Pada penelitian ini terdapat 2 proses utama yaitu *preprocessing*, dan Analisis perbandingan. Terdapat 3 kelas penyakit yang analisis perbandingan yaitu penyakit Covid-19, penyakit Tuberculosis, penyakit Pneumonia, dan Normal. Pada penelitian ini juga dilakukan perbandingan antara klasifikasi yang dilakukan oleh CNN dan SVM. Data penelitian menggunakan dataset citra X-ray dada. Penelitian ini menghasilkan sebuah algoritma terbaik yang diimplementasikan untuk mengklasifikasi penyakit paru-paru dari citra rontgen dada.

Kata Kunci: Penelitian Ini Menghasilkan Sebuah Algoritma Terbaik Yang Diimplementasikan Untuk Mengklasifikasi Penyakit Paru-Paru Dari Citra Rontgen Dada.

Abstract

The lungs are an important human organ in the human body, especially in the respiratory system. Another function of the lungs is to maintain stable body temperature, protect the body from dangerous substances, the nose is the sense of smell, but sometimes the lungs will experience conditions where they do not function normally. Chest x-ray images are the most well-known clinical method for the diagnosis of lung diseases. However, diagnosing lung diseases from chest x-ray images is a challenging task even for radiologists. This research proposes a system that can be used for comparative analysis of lung disease by applying the Convolutional Neural Network and Support Vector Machine methods. CNN is a method in the field of object recognition that has special layers, namely convolution layers and pooling layers which enable a good feature learning process. SVM is a comparative analysis method that relies on results from statistical learning theory to guarantee generalization performance. In this research there are 2 main processes, namely preprocessing and comparative analysis.



There are 3 classes of disease for comparative analysis, namely Covid-19 disease, Tuberculosis disease, Pneumonia disease, and Normal disease. In this study, a comparison was also carried out between the classification carried out by CNN and SVM. The research data uses a chest X-ray image dataset. This research produces the best algorithm that is implemented to classify lung diseases from chest x-ray images.

Keywords: *This Research Produces The Best Algorithm That Is Implemented To Classify Lung Diseases From Chest X-Ray Images.*

1. PENDAHULUAN

Paru-paru merupakan organ tubuh manusia yang penting dalam tubuh manusia terutama dalam sistem pernafasan. Fungsi lain paru-paru menjaga suhu tubuh agar tetap stabil melindungi tubuh dari zat berbahaya, hidung sebagai indra penciuman, namun terkadang paru-paru akan mengalami kondisi dimana tidak berfungsi secara normal (Mujahid et al., 2022). Penyakit pada organ tubuh manusia khususnya paru-paru biasanya disebabkan oleh merokok, polusi udara, atau infeksi bakteri, yang akan menyerang sistem pernapasan hingga gangguan serius. Penyakit paru-paru tercatat pada tahun 2019 mencapai 212,3 juta kasus dilaporkan secara global dengan jumlah kematian mencapai 74,4 juta (Safiri et al., 2022). Penyakit paru-paru menyebabkan kelemahan alami dalam sistem pernapasan, yang mengharuskan pasien untuk berhati-hati dan perhatian untuk meringankan masalah ini (Hussein et al., 2022). Gambar rontgen dada diketahui memiliki potensi dalam pemantauan dan pemeriksaan penyakit paru-paru seperti pneumonia, tuberculosis dan COVID-19 (Ismael & Şengür, 2021).

Coronavirus Disease 2019 (COVID-19) adalah penyakit menular yang disebabkan oleh virus yang belum pernah terjadi sebelumnya (Alom et al., 2020). Penyakit coronavirus 2019 (COVID-19) mengikuti berbagai epidemi di masa lalu yang disebabkan oleh infeksi virus pada pernapasan yang sangat menular. Kasus COVID-19 dilaporkan pertama kali terjadi di Wuhan, China pada Desember 2019 dan menjadi perhatian masyarakat di seluruh dunia khusus nya pada bidang kesehatan (Giri et al., 2021; World Health Organization, 2020). Pneumonia merupakan penyakit menular yang disebabkan oleh virus, kuman, jamur, dan bakteri. Ini mengorbankan paru-paru yang bisa sangat berbahaya dan mengancam jiwa jika tidak diidentifikasi pada tahap awal (Yi et al., 2021). Penyakit pneumonia menyebabkan lebih dari 1,1 juta orang rawat inap dan 50.000 kematian pada tahun 2010, Sebagian besar kematian



karena pneumonia terjadi pada pasien di atas 65 tahun. Tuberculosis merupakan penyakit menular infeksi bakteri dengan penyebaran infeksi melalui kontak langsung atau udara (Fati et al., 2022). TBC umumnya menyerang paru-paru, tetapi juga dapat menyerang organ tubuh lain, seperti ginjal, tulang, belakang, dan otak. Pada tahun 2020 diperkirakan hampir 4,3 juta orang jatuh sakit dengan Tuberculosis dan diperkirakan 700.000 dengan kasus kematian (World Health Organization, 2020).

Gambar rontgen dada adalah metode klinis yang paling terkenal untuk diagnosis penyakit paru-paru. Namun, mendiagnosis penyakit paru-paru dari gambar rontgen dada merupakan tugas yang menantang bahkan untuk ahli radiologi. Hasil rontgen penderita COVID-19, Pneumonia, dan Tuberculosis sering tidak jelas, inkonsistensi ini menyebabkan banyak keputusan subjektif di antara ahli radiologi untuk penyakit paru-paru (Ayan & Ünver, 2019). Oleh karena itu, perlu ada nya sistem pendukung untuk membantu ahli radiologi untuk mendiagnosis penyakit paru-paru dari gambar rontgen dada. Di bidang pemrosesan gambar dalam pembelajaran mendalam, CNN yang merupakan teknologi utama telah menerima perhatian yang luar biasa. Mengikuti kesuksesan besar AlexNet, model peningkatan kinerja seperti VGG, ResNet, dan InceptionNet terus berkembang, dan studi baru masih dilakukan untuk menempati State-of-the-Art (Hong et al., 2021).

Dalam literatur, ada berbagai pendekatan berbasis pembelajaran mendalam yang menggunakan gambar rontgen dada untuk mendukung diagnose penyakit paru-paru (Kesim et al., 2019) mengusulkan Convolutional Neural Networks untuk klasifikasi pada gambar rontgen dada. (Kesim et al., 2019) mengembangkan arsitektur CNN berukuran kecil karena model CNN yang dilatih sebelumnya mendapatkan kesulitan dalam pengaplikasian. 12 kelas dataset gambar rontgen dada digunakan oleh penulis, dengan skor akurasi 86% yang dilaporkan dalam pengujian mereka. (Horry et al., 2020) mengusulkan pendekatan berbasis pembelajaran yang mendalam untuk mendeteksi COVID-19. Penulis menunjukkan bahwa dengan kumpulan data COVID-19 yang terbatas, model VGG-19 dapat digunakan untuk mengembangkan alat berbasis pembelajaran mendalam yang sesuai untuk deteksi COVID-19. Model VGG19 mampu mengklasifikasikan kondisi Pneumonia vs Normal dan COVID-19 vs Pneumonia untuk beberapa mode pencitraan termasuk X-Ray, Ultrasound, dan CT scan. Berdasarkan uraian diatas maka penulis tertarik untuk mengangkat permasalahan dengan mengambil judul

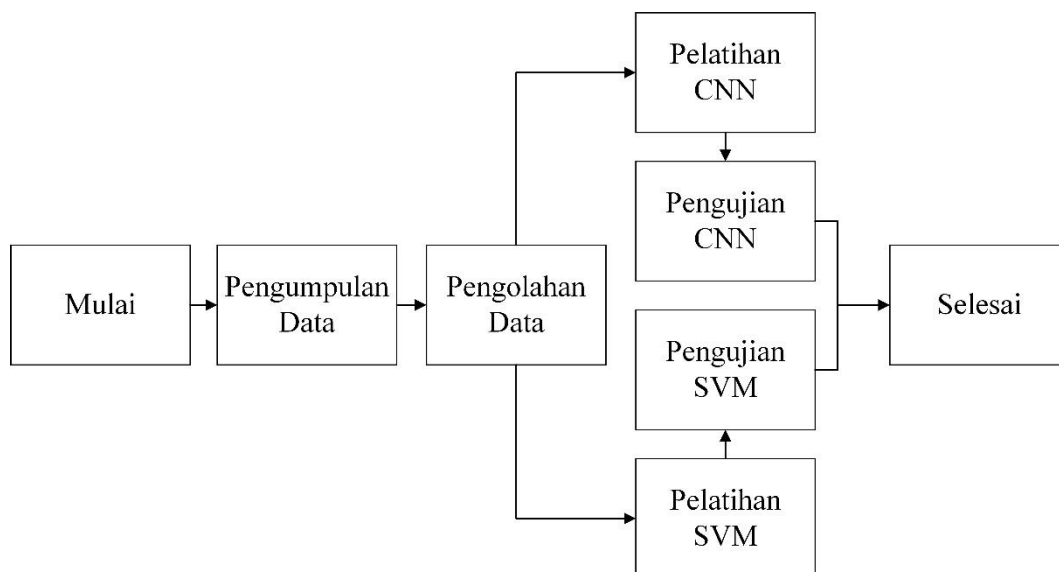


Analisis Perbandingan Penyakit Paru-Paru Dari Citra X-Ray Dada Menggunakan Metode Convolutional Neural Network dan Support-vector machine.

2. METODE PENELITIAN

Tahapan Penelitian

Tahapan dalam melakukan penelitian dimulai dari pengumpulan data, pengolahan data awal, pelatihan model dengan menggunakan CNN dan SVM, dan melakukan pengujian model. Pada pengumpulan data, kegiatan yang dilakukan yaitu mengumpulkan data dari website Kaggle dan akuisisi data citra rontgen dada. Tahap pengolahan data awal atau preprocessing untuk mempersiapkan data citra menjadi citra masukkan dalam CNN dan SVM. Tahap pelatihan model untuk proses pembelajaran dan mendapatkan model terbaik dengan mencari parameter CNN dan SVM yang dapat menghasilkan akurasi terbaik. Tahap terakhir yaitu tahap pengujian untuk melakukan evaluasi dari model terbaik yang dihasilkan dari tahap penelitian.



Gambar 3.1 Tahapan Penelitian

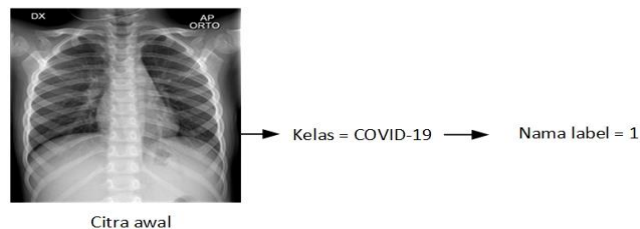
1) Pengumpulan Data

Dataset citra rontgen dada dikumpulkan dengan cara mengunduh pada website: <https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset>. Tahap pengambilan citra rontgen dada dilakukan dengan mengunduh dataset dari website. Pada gambar 3.2 menunjukkan contoh dataset citra rontgen dada.



Gambar 3.2 Contoh Citra Rontgen Dada Penyakit COVID-19

Kemudian masing-masing citra rontgen dada diberi label sesuai dengan penyakit. Setelah masing-masing citra diberi label data dibagi menjadi 4 jenis yaitu COVID-19, Pneumonia, Tuberculosis, dan Normal. Alur pelabelan data citra rontgen dada ditunjukkan pada Gambar 3.3 dan hasil pelabelan data citra rontgen dada ditunjukkan pada Gambar 3.4.



Gambar 3.3 Pelabelan Data Citra Rontgen Dada

Dari hasil pelabelan maka total data yang digunakan dalam penelitian ini dipaparkan pada Tabel 3.1.

Tabel 3.1 Total Data Citra Rontgen

No	Data	Total Data
1	Covid-19	3616
2	Normal	9573
3	Tuberculosis	700
4	Pneumonia	1995

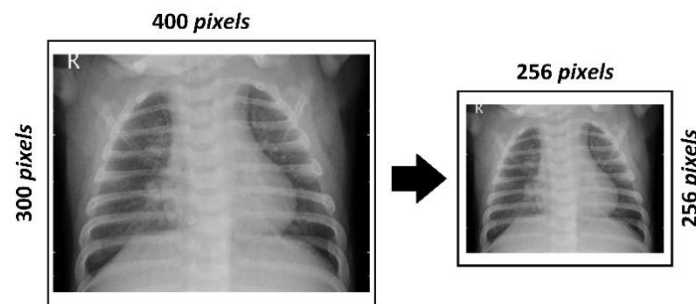


2) Preprocessing

Preprocessing adalah proses yang dilakukan terhadap data, bertujuan agar data siap dijadikan input proses pelatihan. Pada bagian pelatihan dan pengujian, preprocessing dilakukan dengan 2 tahap yaitu resize, dan augmentasi data. Pada tahap Pada penelitian ini preproccressing yang akan dilakukan adalah sebagai berikut:

1. *Resize*

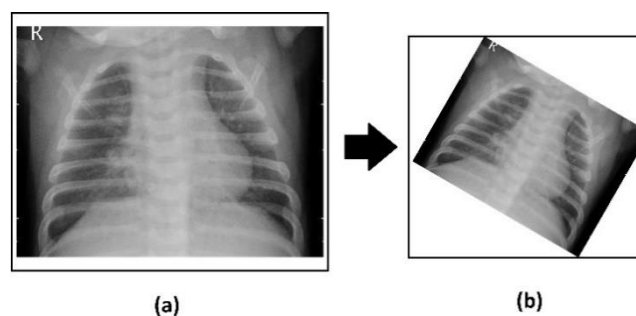
Resize merupakan tahap untuk menyamakan citra dari setiap data yang dimiliki, karena data yang dimiliki masih beraneka ragam ukuran. Pada penelitian ini data citra akan diresize menjadi 256x256 pixel, ukuran tersebut dipilih karena dianggap cukup kecil meminimalisir penggunaan memori tetapi tidak menghilangkan informasi atau feature yang terdapat pada citra. Contoh resize ditunjukkan pada Gambar 3.5.



Gambar 3.4. *Resize* Data Rontgen Dada

2. *Augmentasi Data*

Preprocessing selanjutnya yaitu augmentasi data hal ini dilakukan karena data yang didapatkan pada saat proses pengumpulan data sangat terbatas untuk setiap kelas, agar klasifikasi mendapatkan hasil yang baik dibutuhkan data yang banyak. Proses augmentasi data meliputi *rescale*, *shear range*, *zoom range*, *rotation range*, *horizontal flip*, dan *fill mode*. Contoh data *augmentation* ditunjukkan pada Gambar 3.6.



Gambar 3.5 *Augmentasi* data Rontgen Dada

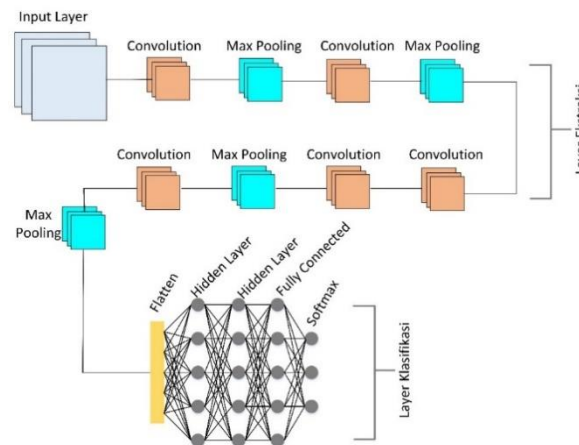


3) Perancangan Model

Pada penelitian ini, setelah data terkumpul maka selanjutnya dilakukan proses perancangan model untuk melakukan proses pembelajaran terhadap data *training* dari setiap kelas, model yang dirancang sangat berpengaruh pada hasil klasifikasi

1. Convolutional Neural Network

Pada model CNN sangat banyak arsitektur yang dapat dibuat, setiap arsitektur dengan data tertentu harus melewati *tuning* untuk mendapatkan model yang optimal. Arsitektur CNN pada penelitian ini menggunakan layer ekstraksi arsitektur AlexNet (Alex Krizhevsky, Ilya Sutskever, 2007). Rancangan arsitektur model CNN ditunjukkan Gambar 3.7.



Gambar 3.6 Rancangan arsitektur CNN

Pada arsitektur CNN, terdapat 2 bagian yaitu *layer ekstraksi* dan *layer klasifikasi*. Pada *layer ekstraksi* terdapat *input layer*, 5 *convolution layer*, 4 *max pooling layer*. Pada *layer klasifikasi* terdapat *flatten layer*, 2 *hidden layer*, dan *fully connected layer* dengan aktivasi *softmax*. Tabel 3.2 merupakan *detail* perancangan model arsitektur CNN yang digunakan pada penelitian ini.

Tabel 3.2 Rancangan model arsitektur CNN

Layers	CNN	Output
<i>Input Layer</i>	256x256x3	256x256x3
<i>Convolution Layer</i>	11x11x128	123x123x128
<i>Pooling Layer</i>	3x3 <i>max pooling</i>	61x61x128
<i>Convolution Layer</i>	5x5x128	61x61x128
<i>Pooling Layer</i>	3x3 <i>max pooling</i>	30x30x128
<i>Convolution Layer</i>	3x3x256	30x30x256



<i>Convolution Layer</i>	3x3x384	30x30x384
<i>Pooling Layer</i>	3x3 <i>max pooling</i>	14x14x384
<i>Convolution Layer</i>	3x3x256	14x14x256
<i>Pooling Layer</i>	3x3 <i>max pooling</i>	6x6x256
<i>Flatten Layer</i>	-	9216
<i>Hidden Layer</i>	1024, ReLU	1024
<i>Hidden Layer</i>	512, ReLU	512
<i>Output Layer</i>	<i>num class, softmax</i>	5

a. *Layer Ekstraksi*

Pada layer ekstraksi dilakukan proses ekstraksi feature dari data, proses ekstraksi pada CNN dilakukan dengan operasi matematis yang direpresentasikan sebagai layer, yaitu convolution layer, pooling layer, dan fungsi aktivasi ReLU. Pada arsitektur ini penentuan jumlah dari setiap layer tidak memiliki aturan yang tetap, sehingga penyusunan layer pada arsitektur ini disesuaikan dengan data.

Pada penelitian ini, perancangan arsitektur CNN dapat dilihat pada Gambar 3.7 dan detail perancangan dapat dilihat pada Tabel 3.2. Layer input menggunakan matriks berukuran 256x256x3 ukuran tersebut digunakan berdasarkan ukuran data citra pelatihan, yang telah melalui proses preprocessing, angka 3 yang dimaksud merupakan 3 channel pada matriks data citra yaitu Red, Green, Blue (RGB).

Layer ekstraksi pertama yaitu *convolution layer* dengan ukuran 11x11 dengan jumlah kernel 128. Matriks *input* akan melalui proses perkalian dengan matriks kernel, ilustrasi dapat dilihat pada Gambar 2.3. Proses lapisan ini akan diaktivasi dengan fungsi ReLU, dimana fungsi tersebut merubah semua nilai negative pada matriks menjadi 0, ilustrasi dapat dilihat pada Gambar 2.6.

Pada penelitian ini terdapat beberapa convolution layer yang digunakan. Proses dan ukuran *convolution* pertama sudah dijelaskan sebelumnya hanya berbeda ukuran dan jumlah kernel nya. *Convolution layer* kedua dengan ukuran 5x5 dengan jumlah kernel 128, *convolution* ketiga dengan ukuran 3x3 dengan jumlah kernel 256, *convolution* keempat dengan ukuran 3x3 dengan jumlah kernel 384, *convolution* kelima dengan ukuran 3x3 dengan jumlah kernel 256.

Layer ekstraksi kedua yaitu *pooling layer*, pada penelitian ini menggunakan *max pooling layer*. *Max pooling* memiliki operasi sederhana untuk mengambil nilai maksimal dari



range tertentu yang ditentukan dari kernel dan mengurangi dimensi *feature map*. Pada penelitian ini menggunakan empat layer *max pooling* dimana memiliki ukuran berbeda. Pada layer pertama sampai keempat *max pooling* memiliki ukuran 3x3, ilustrasi *max pooling* dapat dilihat pada Gambar 2.5.

Layer ekstraksi ketiga yaitu *flatten layer* yang bertujuan membentuk ulang fitur menjadi sebuah *vector* agar dapat digunakan sebagai input pada *fully connected layer*, setelah melakukan proses pada *convolution layer* dan *pooling layer* menghasilkan *vector* berukuran 1x9216 yang dijadikan sebagai input untuk layer klasifikasi.

b. Layer Klasifikasi

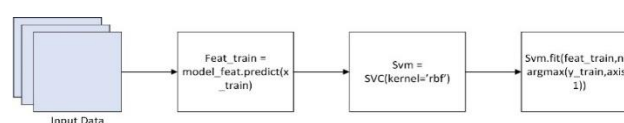
Pada layer klasifikasi ini menggunakan *fully connected layer* yang berbasis *backpropagation*. Vektor dari layer klasifikasi akan menjadi input untuk melatih bobot pada layer klasifikasi. Pada penelitian ini layer klasifikasi terdiri dari 2 hidden layer dengan jumlah node 1024 dan 512 dengan activation ReLU.

Vektor input akan melalui proses *forwardpropagation* yang dilakukan dengan mengkalikan matriks input dengan matriks bobot yang menghubungkan dengan *hidden layer*, hasil dari perkalian tersebut akan diaktivasi dengan fungsi ReLU, outputnya akan dikalikan dengan matriks bobot yang menghubungkan *hidden layer* dengan layer output dan fungsi aktivasi dengan softmax.

Selanjutnya adalah proses *backpropagation*, yang bertujuan untuk menyesuaikan kembali tiap *weight* berdasarkan *error* yang didapat pada proses *forwardpropagation*. Proses *forwardpropagation* dan *backpropagation* akan dilakukan secara *iterative* sehingga mendapatkan bobot yang dianggap optimal, menghasilkan *accuracy* yang tinggi dan *error* yang kecil.

2. Support Vector Machine

Metode setelah CNN adalah penggunaan SVM untuk klasifikasi data citra rontgen dada. Rancangan SVM ditunjukkan pada Gambar 3.8.



Gambar 3.7 Rancangan SVM



Pada Gambar 3.8 menunjukkan rancangan dari SVM yang akan digunakan. Input data untuk program SVM berasal dari data citra rontgen dada. Langkah selanjutnya adalah dilakukan prediksi terhadap data x_{train} . Inisialisasi kernel SVM yang akan digunakan untuk pengenalan karakternya, dan selanjutnya dilakukan *fitting* terhadap model SVM yang telah dibuat.

4) *Tuning Model*

Pada proses *tuning* bertujuan agar mendapatkan model yang lebih optimal, model awal akan melalui proses *tuning* untuk menyesuaikan model terhadap data yang ada. Pada rancangan model CNN dan SVM yang akan detuning sebagai berikut.

1. *Batch Normalization*

Penambahan *batch normalization* pada proses *tuning* yang bertujuan membuat jaringan saraf lebih cepat dan lebih stabil dalam proses pelatihan.

2. *Hidden Layer*

Pada proses *tuning* model tidak selalu menambahkan *hidden layer*, karena model awal itu sudah ada *hidden layer*, tetapi jika hasil pada saat pelatihan belum memuaskan maka *hidden layer* ditambahkan pada layer klasifikasi. Jadi *hidden layer* disini bersifat opsional.

3. *Node Hidden Layer*

Meningkatkan jumlah *node* pada *hidden layer* pada saat *tuning* memungkinkan untuk mengurangi kesalahan pada saat pelatihan, tetapi juga mengurangi jumlah generalisasi. Pada penelitian ini jumlah *node* yang akan digunakan sebesar 1024 dan 512.

4. *Dropout*

Pada proses *tuning dropout* sangat penting dilakukan, karena mencegah terjadi *overfitting* dalam jaringan saraf tiruan. Penelitian ini menggunakan *dropout* sebesar 0.2 dan 0.2.

5. *Loss function*

Loss function adalah fungsi yang menghitung jarak antara *output* algoritma saat ini dan *output* yang diharapkan. Fungsi ini merupakan metode untuk mengevaluasi bagaimana algoritma kita memodelkan data. Penelitian ini menggunakan ‘*categorical_crossentropy*’ sebagai *loss function*.



6. *Learning function*

Learning rate adalah *hyper parameter* yang mengontrol seberapa banyak kita menyesuaikan bobot jaringan kita dengan memperhatikan kerugian pada *gradient*.

7. *Decay*

Fungsi *decay* adalah dimana setiap setelah pembaruan bobot dikalikat dengan *factor* yang sedikit yaitu kurang dari 1, dengan tujuan mencegah bobot tumbuh terlalu besar dan dapat dilihat sebagai penurunan *gradient*.

8. *Momentum*

Fungsi *momentum* adalah mempercepat penurunan *gradient* ke arah yang relevan dan meredam osilasi.

9. *Nesterov*

Fungsi *nesterov* yaitu menghitung rata-rata *momentum* yang bergerak menurun ke dalam *gradient*, secara *default* nilai *nesterov* adalah *false*.

10. *Kernel*

Fungsi utama kernel adalah mengambil ruang *input* berdimensi rendah dan mengubahnya menjadi ruang berdimensi lebih tinggi. Kernel sangat berguna dalam masalah non-linear. Pada penelitian ini kernel yang digunakan adalah 'rbf'.

11. *C (Regularization)*

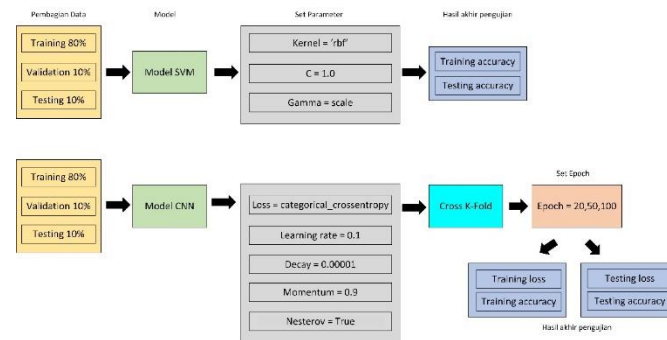
C adalah parameter *penalty* yang mewakili kesalahan klasifikasi. Kesalahan klasifikasi memberi tahu pengoptimalan SVM berapa banyak kesalahan yang dapat ditanggung. Kita dapat mengontrol *trade-off* antara batas keputusan dan istilah kesalahan klasifikasi. Pada penelitian ini nilai C adalah 1.0.

12. *Gamma*

Gamma bertujuan menentukan seberapa jauh pengaruh perhitungan garis pemisah yang dapat diterima. Pada penelitian ini *gamma* yang digunakan adalah 'scale'.

5) Perancangan Pengujian

Setelah proses perancangan model selesai, maka didapatkan 2 model yaitu CNN dan SVM. Selanjutnya akan dilakukan beberapa skenario pengujian. Rancangan pengujian ditunjukkan pada Gambar 3.9.



Gambar 3.8 Rancangan pengujian

Pada perancangan pengujian langkah pertama adalah pembagian data. Data yang digunakan untuk model CNN dan SVM dibagi menjadi 3 dengan persentase yang berbeda, yaitu *training* 80%, *validation* 10%, dan *testing* 10%. Langkah kedua adalah model CNN dan SVM yang telah dibuat dilakukan proses pembelajaran.

Langkah ketiga yaitu mengatur beberapa parameter. Pada model SVM parameter yang diatur yaitu *Kernel*, *C*, *Gamma* dan untuk model CNN parameter yang akan diatur yaitu *Loss*, *Learning rate*, *Decay*, *Momentum*, dan *Nesterov*. Pada model CNN juga *K-fold* dan *epoch* akan diatur, untuk *K-fold* bernilai 5 *fold*. Jika kita mengatur *epoch* 20 namun hasilnya akhir sesuai dengan harapan maka pengujian berhasil, namun jika tidak berhasil maka *epoch* ditingkatkan menjadi 50 atau 100. Setelah semuanya sudah diatur maka hasil akhir pada model SVM adalah *training accuracy*, *testing accuracy* dan pada model CNN adalah *training loss*, *training accuracy*, *testing loss*, dan *testing accuracy*.

6) Evaluasi Accuracy, Precision, Recall, dan F1-score

Proses terakhir adalah evaluasi untuk mengetahui kinerja dari model yang telah dibuat. Pada tahap ini dilakukan dengan melakukan klasifikasi terhadap semua data *testing* yaitu memiliki 3 kelas. Kemudian akan dilakukan proses perhitungan *accuracy*, *precision*, *recall*, dan *f1-score* dengan memanfaatkan *confusion matrix*.

1. Accuracy

$$Akurasi = \sum_{i=1}^l \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} * 100\%$$

2. Precision

$$Presisi_{class} = \frac{TP_{class}}{TP_{class} + FP_{class}} * 100\%$$



3. *Recall*

$$Recall_{class} = \frac{TP_{class}}{TP_{class} + FN_{class}} * 100\%$$

4. *F1-score*

$$F1-score_{class} = 2 * \frac{Recall_{class} * Presisi_{class}}{Recall_{class} + Presisi_{class}}$$

3. HASIL DAN PEMBAHASAN

a) Metodologi Pelatihan Model

Dalam model CNN, setiap lapisan untuk peningkatan kedalaman dilatih menggunakan optimisasi Stochastic Gradient Descent dengan parameter kinerja standar (learning_rate=0.001, momentum=0.9, dan nesterov=True) dengan menggunakan ukuran batch sebesar 16, dilakukan melalui validasi silang 5 lipatan dengan 10 epoch di seluruh dataset. Selain itu, teknik penjadwalan tingkat pembelajaran yang efektif yang disebut "ReduceLROnPlateau" dimasukkan untuk mengatur relevansi kerugian. Ini dimulai dengan tingkat pembelajaran sebesar 0.0001 sesuai kebutuhan.

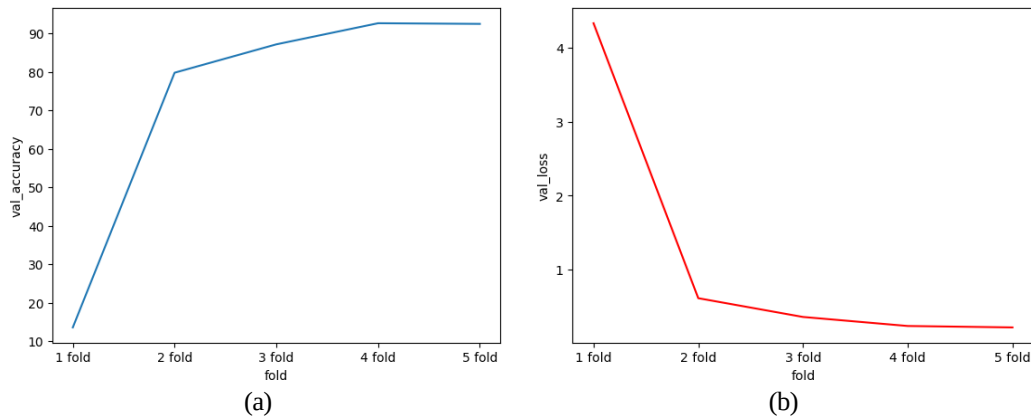
Ekstraksi fitur PCA mengurangi dimensi gambar menjadi n_components=153, menghasilkan rasio variasi sebesar 0.920. Hasil n_components digunakan sebagai data pelatihan untuk SVM. Dalam model SVM, berbagai parameter termasuk kernel='rbf', C=5, class_weight='balanced', dan gamma='scale' digunakan, dipilih berdasarkan estimasi optimal dari GridSearch.

b) Hasil Model CNN dan SVM

Eksperimen penelitian menggunakan model CNN dan SVM. Tabel 3 dan Gambar 9 menunjukkan hasil dari model yang dilatih dan diuji.

Tabel 3
Akurasi Pelatihan dan Pengujian Model CNN

Model	Training Accuracy	Testing Accuracy
CNN	0.92	0.23



Gambar 9. Kinerja Umum Klasifikasi CNN

Tabel 4 menggambarkan bahwa model CNN mencapai skor akurasi tertinggi sebesar 0,92 dan 0,23 untuk pelatihan dan pengujian, secara berturut-turut. Kurva pembelajaran untuk model CNN digambarkan dalam Figur 9. Hasil ini diantisipasi, mengingat variasi yang cukup besar dalam dataset gambar sinar-X. Eksperimen berikutnya melibatkan kombinasi klasifikasi PCA dengan SVM, seperti yang terindikasi dalam Tabel tersebut.

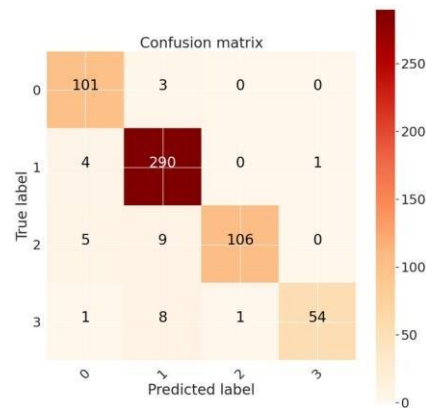
c) Hasil dari Metrik Evaluasi

Model yang diusulkan mengalami evaluasi melalui matriks kebingungan, meliputi metrik seperti akurasi, presisi, recall, dan skor F1. Model arsitektur CNN dinilai menggunakan gambar sinar-X dada (CXR), dengan total 583 gambar. Di antaranya, dataset uji terdiri dari 295 kasus COVID-19, 120 kasus Normal, 64 kasus Tuberkulosis, dan 104 kasus Pneumonia Viral. Tabel 5 menggambarkan hasil evaluasi dari model CNN, sementara matriks kebingungan diilustrasikan dalam Gambar 10.

Tabel 5

Hasil Evaluasi Kinerja Model CNN

Model	Accuracy	Precision	Recall	F1-score
CNN	0.95	0.95	0.92	0.93



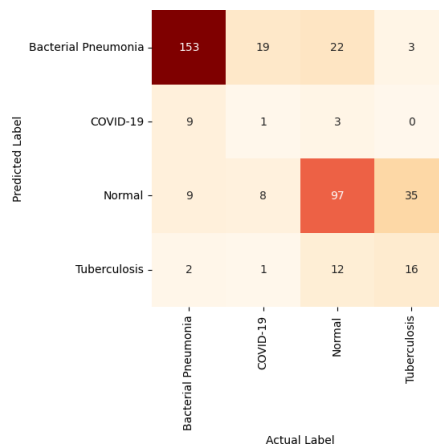
Gambar 10. Hasil Matriks Konfusi Kinerja Model CNN

Tabel 6 menunjukkan hasil evaluasi dari kombinasi klasifikasi PCA dan SVM, sementara matriks kebingungan untuk model PCA-SVM disajikan dalam Gambar 11.

Tabel 6

Hasil Evaluasi Kinerja Model SVM

Model	Accuracy	Precision	Recall	F1-score
SVM	0.68	0.51	0.48	0.48



Gambar 11. Hasil Matriks Kebingungan Kinerja Model SVM

KESIMPULAN DAN SARAN

Studi ini memperkenalkan sebuah sistem yang dirancang untuk analisis perbandingan penyakit paru-paru menggunakan Metode Convolutional Neural Network (CNN) dan Metode Support Vector Machine (SVM). CNN, sebuah teknik yang terkenal dalam pengenalan objek, menggabungkan lapisan khusus - lapisan konvolusi dan lapisan penggabungan - yang



memfasilitasi pembelajaran fitur yang efektif. SVM, di sisi lain, adalah metode analisis perbandingan yang berakar dalam teori pembelajaran statistik, yang menjamin kinerja generalisasi yang tangguh. Penelitian ini mencakup dua proses utama: pra-pemrosesan dan analisis perbandingan. Empat kelas penyakit - COVID-19, Tuberkulosis, Pneumonia, dan Normal - dipertimbangkan untuk analisis perbandingan. Selain itu, dilakukan perbandingan antara metode klasifikasi CNN dan SVM. Dataset penelitian terdiri dari gambar sinar-X dada. Studi ini mencapai puncaknya dalam pengembangan algoritma yang dioptimalkan untuk mengklasifikasikan penyakit paru-paru dari gambar sinar-X dada. Hasil menunjukkan bahwa model CNN mengungguli model SVM, mencapai akurasi sebesar 0,95 dibandingkan dengan 0,65 SVM.

DAFTAR PUSTAKA

- Ahmad, B., Usama, M., Huang, C. M., Hwang, K., Hossain, M. S., & Muhammad, G. (2020). Discriminative Feature Learning for Skin Disease Classification Using Deep Convolutional Neural Network. *IEEE Access*, 8, 39025–39033. <https://doi.org/10.1109/ACCESS.2020.2975198>
- Aldhyani, T. H. H., Verma, A., Al-Adhaileh, M. H., & Koundal, D. (2022). Multi-Class Skin Lesion Classification Using a Lightweight Dynamic Kernel Deep-Learning-Based Convolutional Neural Network. *Diagnostics*, 12(9). <https://doi.org/10.3390/diagnostics12092048>
- Alex Krizhevsky, Ilya Sutskever, G. E. H. (2007). ImageNet Classification with Deep Convolutional Neural Networks Alex. *Handbook of Approximation Algorithms and Metaheuristics*, 1–1432. <https://doi.org/10.1201/9781420010749>
- Alom, M. Z., Rahman, M. M. S., Nasrin, M. S., Taha, T. M., & Asari, V. K. (2020). *COVID_MTNet: COVID-19 Detection with Multi-Task Deep Learning Approaches*. <http://arxiv.org/abs/2004.03747>
- Arrofiqoh, E. N., & Harintaka, H. (2018). Implementasi Metode Convolutional Neural Network Untuk Klasifikasi Tanaman Pada Citra Resolusi Tinggi. *Geomatika*, 24(2), 61. <https://doi.org/10.24895/jig.2018.24-2.810>
- Athena Anwar, I. D. (2014). Pneumonia among Children Under Five Years of Age in Indonesia. *Urology*, 56(6), 956–961. [https://doi.org/10.1016/S0090-4295\(00\)00847-5](https://doi.org/10.1016/S0090-4295(00)00847-5)



-
- Ayan, E., & Ünver, H. M. (2019). Diagnosis of pneumonia from chest X-ray images using deep learning. *2019 Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science, EBBT 2019*, 2–6. <https://doi.org/10.1109/EBBT.2019.8741582>
- Candemir, S., & Antani, S. (2019). A review on lung boundary detection in chest X-rays. *International Journal of Computer Assisted Radiology and Surgery*, 14(4), 563–576. <https://doi.org/10.1007/s11548-019-01917-1>
- Demir, F., Sengur, A., & Bajaj, V. (2020). Convolutional neural networks based efficient approach for classification of lung diseases. *Health Information Science and Systems*, 8(1), 1–8. <https://doi.org/10.1007/s13755-019-0091-3>
- Fati, S. M., Senan, E. M., & ElHakim, N. (2022). Deep and Hybrid Learning Technique for Early Detection of Tuberculosis Based on X-ray Images Using Feature Fusion. *Applied Sciences*, 12(14), 7092. <https://doi.org/10.3390/app12147092>
- Giri, B., Pandey, S., Shrestha, R., Pokharel, K., Ligler, F. S., & Neupane, B. B. (2021). Review of analytical performance of COVID-19 detection methods. *Analytical and Bioanalytical Chemistry*, 413(1), 35–48. <https://doi.org/10.1007/s00216-020-02889-x>
- Hidayat, A., Darusalam, U., & Irmawati, I. (2019). Detection of Disease on Corn Plants Using Convolutional Neural Network Methods. *Jurnal Ilmu Komputer Dan Informasi*, 12(1), 51. <https://doi.org/10.21609/jiki.v12i1.695>
- Hong, M., Rim, B., Lee, H. C., Jang, H. U., Oh, J., & Choi, S. (2021). Multi-class classification of lung diseases using cnn models. *Applied Sciences (Switzerland)*, 11(19), 1–17. <https://doi.org/10.3390/app11199289>
- Horry, M. J., Chakraborty, S., Paul, M., Ulhaq, A., Pradhan, B., Saha, M., & Shukla, N. (2020). COVID-19 Detection through Transfer Learning Using Multimodal Imaging Data. *IEEE Access*, 8, 149808–149824. <https://doi.org/10.1109/ACCESS.2020.3016780>
- Hou, J., & Gao, T. (2021). Explainable DCNN based chest X-ray image analysis and classification for COVID-19 pneumonia detection. *Scientific Reports*, 11(1), 1–15. <https://doi.org/10.1038/s41598-021-95680-6>
- Hussein, F., Mughaid, A., AlZu'bi, S., El-Salhi, S. M., Abuhaija, B., Abualigah, L., & Gandomi, A. H. (2022). Hybrid CLAHE-CNN Deep Neural Networks for Classifying Lung



-
- Diseases from X-ray Acquisitions. *Electronics*, 11(19), 3075.
<https://doi.org/10.3390/electronics11193075>
- Irfan, A., Adivishnu, A. L., Sze-To, A., Dehkharghanian, T., Rahnamayan, S., & Tizhoosh, H. R. (2020). Classifying Pneumonia among Chest X-Rays Using Transfer Learning. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2020-July*, 2186–2189.
<https://doi.org/10.1109/EMBC44109.2020.9175594>
- Ismael, A. M., & Şengür, A. (2021). Deep learning approaches for COVID-19 detection based on chest X-ray images. *Expert Systems with Applications*, 164(March 2020).
<https://doi.org/10.1016/j.eswa.2020.114054>
- Kesim, E., Dokur, Z., & Olmez, T. (2019). X-ray chest image classification by a small-sized convolutional neural network. *2019 Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science, EBBT 2019*, 5–9.
<https://doi.org/10.1109/EBBT.2019.8742050>
- Khan, A., Sohail, A., Zahoor, U., & Qureshi, A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, 1–67.
<https://doi.org/10.1007/s10462-020-09825-6>
- Khan, S., Siddique, R., Shereen, M. A., Ali, A., Liu, J., Bai, Q., Bashir, N., & Xue, M. (2020). Emergence of a Novel Coronavirus, Severe Acute Respiratory Syndrome Coronavirus 2: Biology and Therapeutic Options. *Journal of Clinical Microbiology*, 58(8), 1–9.
<https://doi.org/10.1128/JCM.01297-20>
- Koushik, J. (2016). *Understanding Convolutional Neural Networks*. 3, 1–6.
<http://arxiv.org/abs/1605.09081>
- Krause, R., & Smolle, J. (2022). Covid-19 mortality and local burden of infectious diseases: A worldwide country-by-country analysis. *Journal of Infection and Public Health*, 15(12), 1370–1375. <https://doi.org/10.1016/j.jiph.2022.10.018>
- Krüger, F. (2018). Activity, Context, and Plan Recognition with Computational Causal Behaviour Models. *ResearchGate*, August.
https://www.researchgate.net/figure/Confusion-matrix-for-multi-class-classification-The-confusion-matrix-of-a_fig7_314116591



-
- Kumar, S., Negi, A., Singh, J. N., & Gaurav, A. (2018). Brain Tumor Segmentation and Classification Using MRI Images via Fully Convolution Neural Networks. *Proceedings - IEEE 2018 International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2018*, 1178–1181. <https://doi.org/10.1109/ICACCCN.2018.8748614>
- Lee, N. Y., Li, C. W., Tsai, H. P., Chen, P. L., Syue, L. S., Li, M. C., Tsai, C. S., Lo, C. L., Hsueh, P. R., & Ko, W. C. (2020). A case of COVID-19 and pneumonia returning from Macau in Taiwan: Clinical course and anti-SARS-CoV-2 IgG dynamic. *Journal of Microbiology, Immunology and Infection*, 53(3), 485–487. <https://doi.org/10.1016/j.jmii.2020.03.003>
- Link, A. N., & Scott, J. T. (1998). Performance Evaluation Metrics. *Public Accountability*, 1995, 17–21. https://doi.org/10.1007/978-1-4615-5639-8_4
- Masías, V. H., Valle, M., Morselli, C., Crespo, F., Vargas, A., & Laengle, S. (2016). Modeling verdict outcomes using social network measures: The watergate and caviar network cases. *PLoS ONE*, 11(1), 1–24. <https://doi.org/10.1371/journal.pone.0147248>
- Masud, M., Eldin Rashed, A. E., & Hossain, M. S. (2022). Convolutional neural network-based models for diagnosis of breast cancer. *Neural Computing and Applications*, 34(14), 11383–11394. <https://doi.org/10.1007/s00521-020-05394-5>
- Mujahid, M., Rustam, F., Álvarez, R., Luis Vidal Mazón, J., Díez, I. de la T., & Ashraf, I. (2022). Pneumonia Classification from X-ray Images with Inception-V3 and Convolutional Neural Network. *Diagnostics*, 12(5), 1–16. <https://doi.org/10.3390/diagnostics12051280>
- O'Shea, K., & Nash, R. (2015). *An Introduction to Convolutional Neural Networks*. 1–11. <http://arxiv.org/abs/1511.08458>
- Phan, T. (2020). Genetic diversity and evolution of SARS-CoV-2. *Infection, Genetics and Evolution*, 81(February), 104260. <https://doi.org/10.1016/j.meegid.2020.104260>
- Prasetya, T. A., Harjanto, C. T., & Setiyawan, A. (2020). Analysis of student satisfaction of e-learning using the end-user computing satisfaction method during the Covid-19 pandemic. *Journal of Physics: Conference Series*, 1700(1). <https://doi.org/10.1088/1742-6596/1700/1/012012>



-
- Ragab, D. A., Sharkas, M., Marshall, S., & Ren, J. (2019). Breast cancer detection using deep convolutional neural networks and support vector machines. *PeerJ*, 2019(1), 1–23. <https://doi.org/10.7717/peerj.6201>
- Rouan, M. C., Lounis, N., Gevers, T., Dillen, L., Gilissen, R., Raoof, A., & Andries, K. (2012). Pharmacokinetics and pharmacodynamics of TMC207 and its N-desmethyl metabolite in a murine model of tuberculosis. *Antimicrobial Agents and Chemotherapy*, 56(3), 1444–1451. <https://doi.org/10.1128/AAC.00720-11>
- Safiri, S., Carson-Chahhoud, K., Noori, M., Nejadghaderi, S. A., Sullman, M. J. M., Ahmadian Heris, J., Ansarin, K., Mansournia, M. A., Collins, G. S., Kolahi, A. A., & Kaufman, J. S. (2022). Burden of chronic obstructive pulmonary disease and its attributable risk factors in 204 countries and territories, 1990-2019: Results from the Global Burden of Disease Study 2019. *The BMJ*. <https://doi.org/10.1136/bmj-2021-069679>
- Salehi, M., Ghasemian, A., Shokouhi Mostafavi, S. K., Najafi, S., & Rajabi Vardanjani, H. (2017). Sero-prevalence of. *Iran J Pathol*, 12(2), 183–188. <https://www.ncbi.nlm.nih.gov/pubmed/29515642>
- Sharma, H., Jain, J. S., Bansal, P., & Gupta, S. (2020). Feature extraction and classification of chest X-ray images using CNN to detect pneumonia. *Proceedings of the Confluence 2020 - 10th International Conference on Cloud Computing, Data Science and Engineering*, 227–231. <https://doi.org/10.1109/Confluence47617.2020.9057809>
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem*, 2818–2826. <https://doi.org/10.1109/CVPR.2016.308>
- Tomashefski, J. F. (2008). *Dail and Hammar's Pulmonary Pathology*.
- Turkoglu, M. (2021). COVIDetectionNet: COVID-19 diagnosis system based on X-ray images using features selected from pre-learned deep features ensemble. *Applied Intelligence*, 51(3), 1213–1226. <https://doi.org/10.1007/s10489-020-01888-w>
- Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F., & Pinheiro, P. R. (2020). CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-



-
- 19 Detection. *IEEE Access*, 8, 91916–91923.
<https://doi.org/10.1109/ACCESS.2020.2994762>
- World Health Organization. (2020). *Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases*. 2019(March).
- Xie, Y., Zaccagna, F., Rundo, L., Testa, C., Agati, R., Lodi, R., Manners, D. N., & Tonon, C. (2022). Convolutional Neural Network Techniques for Brain Tumor Classification (from 2015 to 2022): Review, Challenges, and Future Perspectives. *Diagnostics*, 12(8).
<https://doi.org/10.3390/diagnostics12081850>
- Yi, R., Tang, L., Tian, Y., Liu, J., & Wu, Z. (2021). Identification and classification of pneumonia disease using a deep learning-based intelligent computational framework. *Neural Computing and Applications*, 7. <https://doi.org/10.1007/s00521-021-06102-7>