

Application of Shapley Additive Explanations (SHAP) in Deep Learning for Lung Disease Detection Using X-ray Images

Sarifah Muliani¹, Benny Sukma Negara², Muhammad Irsyad³, Jasril⁴, Iwan Iskandar⁵

^{1,2,3,4,5} Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Kota Pekanbaru, 28293, Indonesia.

Informasi Artikel

Diterima : 30 Mei 2025
Revisi : 9 Juni 2025
Publikasi : 20 Juni 2025

Kata Kunci:

COVID-19
Pneumonia
Citra X-ray
Explainable Artificial Intelligence
Shapley Additive Explanations

ABSTRAK

Pemeriksaan menggunakan citra *x-ray* merupakan metode yang efektif dalam membantu deteksi penyakit paru-paru, seperti COVID-19, dan pneumonia. Seiring dengan perkembangan teknologi yang meningkat, proses diagnosis kini dapat dilakukan secara lebih akurat dengan memanfaatkan sistem berbasis kecerdasan buatan. Salah satu metode yang banyak digunakan adalah *deep learning* namun metode ini bersifat *black-box*, sehingga hasil prediksi sulit dipahami dengan alasan dibalik keputusan model. Tujuan penelitian ini adalah untuk membangun sistem klasifikasi citra *x-ray* menggunakan model *deep learning* berbasis *Convolutional Neural Network* (CNN) dengan arsitektur VGG-16, serta menerapkan metode *Shapley Additive Explanations* (SHAP) untuk memberikan penjelasan mengenai visual terkait area citra yang mempengaruhi hasil prediksi. Model dilatih menggunakan beberapa konfigurasi, dan hasil terbaik diperoleh pada rasio data 80% : 20%, *learning rate* 0.001, *batch size* 32, dan 50 *epoch* dengan mencapai akurasi 96,00%. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*. Penerapan metode *Shapley Additive Explanations* (SHAP) pada penelitian ini berhasil meningkatkan pemahaman terhadap hasil prediksi model dengan menampilkan area-area penting pada citra yang berkontribusi terhadap keputusan model.

ABSTRACT

X-ray imaging is an effective method to help detect lung diseases, such as COVID-19 and pneumonia. Along with the increasing development of technology, the diagnosis process can now be carried out more accurately by utilizing artificial intelligence-based systems. One of the methods widely used is deep learning, but this method is black-box, so the prediction results are difficult to understand with the reasons behind the model's decisions. The purpose of this study is to build an x-ray image classification system using a deep learning model based on Convolutional Neural Network (CNN) with VGG-16 architecture, and apply the Shapley Additive Explanations (SHAP) method to provide an explanation of the visuals related to the image area that affects the prediction results. The model was trained using several configurations, and the best results were obtained at a data ratio of 80%: 20%, a learning rate of 0.001, a batch size of 32, and 50 epochs. The results showed that the model was able to achieve an accuracy of 95.75% on the training data and 96.00% on the validation data. The SHAP method is used to improve understanding of the prediction results. The results show that the combination of deep learning and SHAP is able to provide a visual explanation of the model's prediction results.

This is an open-access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



*Penulis Koresponden

Email: bsnegara@uin-suska.ac.id

S. Muliani, B.S. Negara, M. Irsyad, Jasril, & I. Iskandar "Application of Shapley Additive Explanations (SHAP) in Deep Learning for Lung Disease Detection Using X-ray Images" *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, vol. 5, no. 2, pp. 709-719, Juni 2025. doi: 10.30811/jaise.v5i2.7044

1. PENDAHULUAN

Mengingat penyakit seperti COVID-19 dan pneumonia dengan dampak yang luas serta tingkat kematian yang tinggi, menjadi ancaman yang serius bagi kesehatan global [1]. Tahun 2020 sendiri bahkan dikenang sebagai tahun COVID-19, penyakit ini disebabkan oleh virus corona [2]. Virus penyebab pandemi ini yaitu *severe acute respiratory syndrome coronavirus 2* (SARS-CoV-2), pertama kali muncul di Tiongkok pada bulan Januari 2020 [3]. Demikian pula, pneumonia juga merupakan penyakit infeksi saluran pernapasan yang dapat menimbulkan komplikasi serius hingga kematian, terutama pada kelompok rentan seperti anak-anak, lansia, dan individu dengan penyakit penyerta [4]. COVID-19 atau pneumonia memiliki gejala klinis yang mirip, seperti demam, batuk, dan sesak napas, sehingga seringkali sulit dibedakan tanpa pemeriksaan penunjang berupa pencitraan medis, salah satunya citra *x-ray* dada [5].

Meskipun pemeriksaan seperti MRI (*Magnetic Resonance Imaging*), dan CT scan (*Computed Tomography*) mampu memberikan hasil gambar yang lebih detail, kedua metode tersebut relatif mahal dan terbatasnya ketersediaan terutama di wilayah berpenghasilan rendah [6]. Pemanfaatan citra *x-ray* menjadi pilihan utama yang lebih efektif dan biaya yang relatif terjangkau dalam mendeteksi berbagai penyakit paru-paru [7]. Seiring berkembangnya teknologi, khususnya dibidang kecerdasan buatan (AI), metode pembelajaran mendalam seperti *deep learning* telah terbukti andal dalam berbagai tugas pencitraan medis, termasuk analisis citra *x-ray* [8]. Teknologi ini mampu mengenali pola visual secara otomatis dan menghasilkan prediksi yang akurat, sehingga berpotensi besar membantu proses diagnosis penyakit paru-paru secara lebih cepat dan efisien [7].

Deep learning merupakan salah satu cabang dari *machine learning* yang dirancang untuk meningkatkan efisiensi dalam pengolahan data kompleks, serta menghasilkan model prediksi dan klasifikasi yang lebih akurat [9]. Salah satu arsitektur *deep learning* adalah *artificial neural network* (ANN) yang meniru cara kerja jaringan saraf biologis dalam memproses informasi. ANN terdiri dari lapisan neuron yang saling terhubung dan mampu mempelajari pola data melalui pelatihan [10]. Salah satu pengembangan ANN adalah *convolutional neural network* (CNN) yang dirancang untuk pengolahan data citra. Pada penelitian ini, CNN digunakan untuk mengenali pola pada citra rontgen dada guna mendeteksi kelainan dengan sensitivitas dan spesifisitas tinggi [11]. Pada penelitian ini arsitektur CNN yang digunakan adalah *visual geometry group 16* (VGG-16) yang memiliki kemampuan dalam mendeteksi fitur visual melalui susunan layer konvolusi yang sederhana namun efektif [3]. Penelitian sebelumnya menunjukkan bahwa VGG-16 memiliki performa yang sangat baik, tidak hanya mencapai akurasi tinggi, tetapi juga memiliki potensi besar untuk diterapkan dalam berbagai domain klasifikasi citra. Salah satunya, penelitian [12] menggunakan arsitektur VGG-16 untuk klasifikasi ikan dan berhasil mencapai akurasi sebesar 99,5%. Selain itu, penelitian [13] mengenai klasifikasi penyakit pneumonia dan COVID-19 berbasis citra *x-ray* dengan algoritma *Deep Residual Network* memperoleh akurasi 91,2%. Penelitian lain [14] membandingkan kemampuan klasifikasi citra *x-ray* paru-paru menggunakan metode *transfer learning* dengan ResNet-50 dan VGG-16, yaitu hasilnya menunjukkan bahwa VGG-16 lebih unggul dari segi akurasi.

Meskipun model *deep learning* mampu memberikan hasil prediksi yang akurat, namun model ini bersifat *black-box* atau proses pengambilan keputusannya sulit dipahami oleh pengguna akhir, khususnya di bidang kesehatan [15]. Sehingga membutuhkan metode *explainable AI* (XAI) untuk memberikan gambaran yang lebih transparan mengenai alasan dibalik prediksi yang dihasilkan [16]. Penerapan XAI memungkinkan peningkatan pemahaman terhadap model, membangun kepercayaan pengguna, serta membantu mengidentifikasi potensi dan kesalahan dalam sistem AI secara lebih efektif [17]. Pada penelitian ini metode XAI yang digunakan adalah *Shapley Additive Explanations* (SHAP) yang dapat membantu proses interpretasi citra medis [11].

Beberapa penelitian sebelumnya membandingkan metode *transfer learning* seperti EfficientNetB0, InceptionV3, dan LetNet, dengan hasil akurasi mencapai 99,20% penelitian tersebut menyatakan bahwa metode SHAP dan LIME mampu memberikan wawasan penting terhadap proses pengambilan keputusan model, sementara Grad-CAM dan Grad-CAM++ menyoroti area citra yang berkontribusi terhadap hasil klasifikasi [18]. Selain itu, penelitian lain menerapkan metode *transfer learning* menggunakan model Xception dan berhasil memperoleh akurasi sebesar 98,2% dengan pemanfaatan teknik Grad-CAM dan SHAP untuk memberikan penjelasan visual terhadap proses prediksi AI, sehingga dapat meningkatkan pemahaman serta kepercayaan di kalangan tenaga medis [7].

Meskipun penelitian sebelumnya telah mengimplementasikan metode XAI seperti SHAP, Grad-CAM, dan LIME pada model *deep learning* untuk klasifikasi citra *x-ray*, sebagian penelitian tersebut lebih

berfokus pada model *transfer learning* modern seperti ResNet, Inception, Xception, dan EfficientNet. Sementara itu, pada penelitian ini penerapan metode SHAP dilakukan menggunakan arsitektur VGG-16 yang masih jarang digunakan dalam kombinasi dengan SHAP untuk interpretasi citra *x-ray*. Penelitian ini juga melakukan pengujian menggunakan beberapa konfigurasi parameter untuk mendapatkan performa terbaik, sehingga diharapkan mampu memberikan hasil klasifikasi yang akurat sekaligus interpretasi visual yang jelas terhadap area citra yang memengaruhi keputusan model.

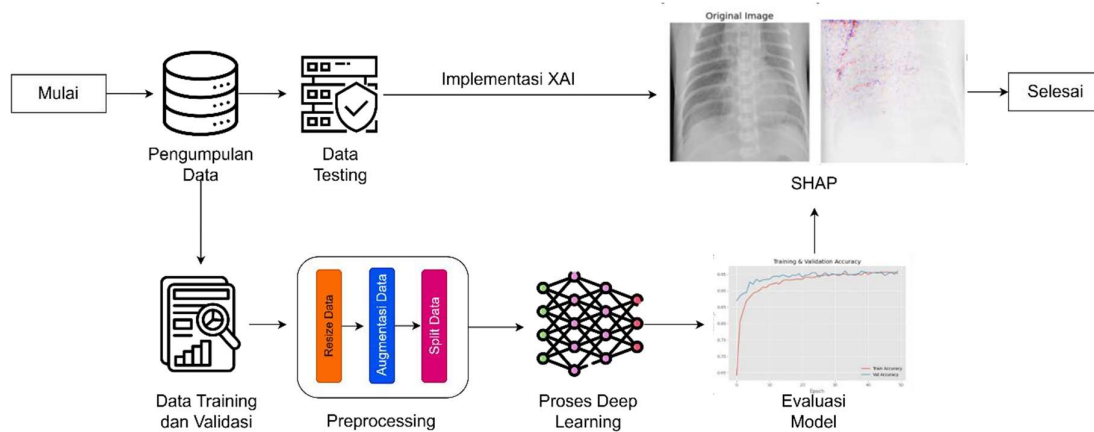
Berdasarkan penelitian sebelumnya, metode SHAP terbukti mampu memberikan interpretasi lokal terhadap prediksi model dengan menunjukkan kontribusi tiap fitur terhadap hasil klasifikasi [19]. Keunggulan SHAP yang bersifat model-agnostic mampu mengaplikasikan berbagai arsitektur *deep learning* menjadikannya fleksibel untuk interpretasi citra medis. Penelitian ini menerapkan SHAP untuk meningkatkan transparansi model dan mendukung proses pengambilan keputusan [20].

Batasan masalah dalam penelitian ini berkaitan dengan proses interpretasi hasil prediksi, yaitu metode yang digunakan adalah *Shapley Additive Explanations* (SHAP) untuk memberikan penjelasan visual terhadap area citra yang mempengaruhi keputusan model.

1. Penelitian ini memanfaatkan citra *x-ray* dada sebagai sumber data utama untuk melakukan klasifikasi terhadap tiga kategori kondisi paru-paru, yaitu COVID-19, pneumonia, dan normal. Dalam proses pengembangan sistem klasifikasi, model *deep learning* yang digunakan dibatasi pada arsitektur *Convolutional Neural Network* (CNN) VGG-16 tanpa melibatkan arsitektur lain sebagai perbandingan.
2. Proses interpretasi hasil prediksi dalam penelitian ini dilakukan dengan menerapkan metode *Shapley Additive Explanations* (SHAP), yang berfungsi untuk menampilkan visualisasi area pada citra yang memiliki kontribusi terhadap keputusan klasifikasi model.

2. METODE

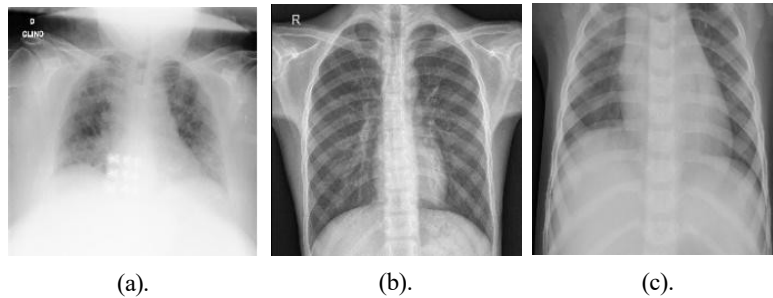
Tahapan metode penelitian pada penelitian ini terdiri dari beberapa proses, yang dimulai dari tahapan pengumpulan data, dilanjutkan dengan *preprocessing* data, proses *deep learning*, evaluasi performa model, hingga tahapan akhir yaitu implementasi *explainable AI* (XAI) menggunakan *shapley additive explanations* (SHAP) untuk memberikan interpretasi terhadap hasil prediksi.



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Data

Pada penelitian ini dataset yang dikumpulkan berasal dari dataset publik yang terdiri dari tiga kelas, yaitu COVID-19, normal, dan pneumonia. Secara keseluruhan, terdapat 3.030 citra *x-ray* dada yang dibagi menjadi data training dan data testing, yang mana data training terdiri dari 3.000 citra *x-ray* yang diperoleh dari situs mendeley [21]. Sementara itu data testing terdiri dari 10 citra *x-ray* perkelas yang diambil dari situs yang sama, namun tidak pernah digunakan atau dilihat oleh model selama proses pelatihan. Masing-masing kelas di data testing terdiri dari 10 citra *x-ray*, sehingga dapat digunakan untuk menguji kemampuan model dalam melakukan prediksi terhadap data baru.



Gambar 2. (a) COVID-19, (b) Normal, (c) Pneumonia

2.2 Preprocessing Data

Pada tahapan ini dilakukan serangkaian proses preprocessing yang bertujuan untuk menyiapkan citra *x-ray* sebelum digunakan dalam proses pelatihan model *deep learning*. Proses ini penting dilakukan agar data yang digunakan berada dalam kondisi standar dan sesuai dengan format yang dibutuhkan oleh arsitektur model. Beberapa tahapan preprocessing yang dilakukan yaitu pemuatan citra beserta labelnya, pengubahan ukuran citra, augmentasi data, serta pembagian dataset menjadi data training dan data validasi.

2.2.1. Resize

Setelah seluruh citra berhasil dimuat, langkah selanjutnya yaitu melakukan *resize* dengan menyesuaikan ukuran gambar menjadi 224 x 224 piksel. Ukuran ini dipilih karena sesuai dengan format input yang dibutuhkan oleh arsitektur VGG-16. Penyesuaian ukuran ini dilakukan agar memastikan semua citra memiliki dimensi yang sama, sehingga dapat meminimalisir error saat proses pelatihan berlangsung.

2.2.2 Augmentasi Data

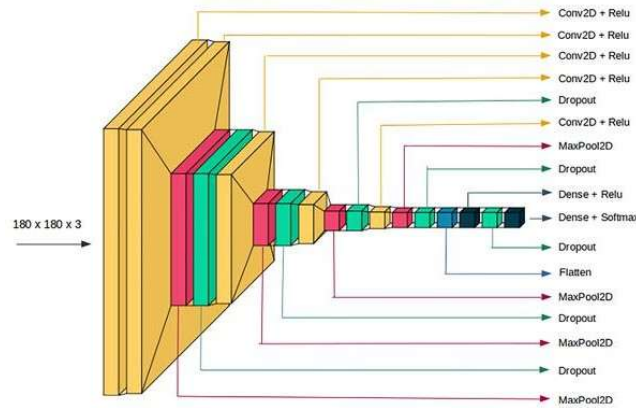
Berikutnya, dilakukan proses augmentasi data untuk menambah variasi citra. Teknik augmentasi yang digunakan adalah *rotation range* sebesar 10 derajat dengan metode interpolasi *nearest*. Proses ini bertujuan agar model dapat beradaptasi terhadap perbedaan orientasi citra, selain itu dilakukan proses *rescaling* pada citra ke dalam skala 1/255 untuk menyesuaikan rentang nilai piksel menjadi 0 hingga 1, yang menjadi standar input dari model *deep learning* [18].

2.2.3 Split Data

Selanjutnya dilakukan proses split data dengan membagi data training dan data validasi. Pembagian ini dilakukan untuk memastikan distribusi data pada masing-masing kategori seimbang. Data training dan validasi dipisah berdasarkan tahapan *deep learning* dengan beberapa rasio, yaitu 70% data training 30% data validasi, 80% data training 20% data validasi, dan 90% data training 10% data validasi. Beberapa pengujian telah dilakukan untuk menentukan rasio pembagian data yang memberikan performa terbaik bagi model. Hasil pengujian dapat dilihat pada tabel 2.

2.3 Proses Deep Learning

Proses *deep learning* diterapkan menggunakan algoritma *convolutional neural network* (CNN) dengan menggunakan arsitektur *visual geometry group-16* (VGG-16) [3]. Arsitektur ini dipilih karena banyak digunakan pada berbagai penelitian klasifikasi citra dan terbukti memiliki performa yang baik. VGG-16 memiliki karakteristik berupa susunan lapisan konvolusi yang berjumlah 13 layer, diikuti dengan 5 *pooling layer* dan 3 *fully connected layer* pada bagian akhir [22]. Seluruh *convolution layer* dalam VGG-16 menggunakan kernel berukuran 3x3 dengan *padding* 1 dan *stride* 1 sehingga mampu menangkap fitur visual secara detail [23]. Pada penelitian ini, proses *deep learning* terdiri dari dua tahap yaitu pembangunan model dan pelatihan model.



Gambar 3. Arsitektur VGG-16

2.3.1 Pembangunan Model

Pada tahapan ini dilakukan proses pembangunan model *deep learning* dengan menggunakan arsitektur VGG-16 sebagai *base model*. VGG-16 di-load tanpa *fully connected layer* dan ditambahkan beberapa *layer classifier* di atasnya, yaitu *layer AveragePooling2D* untuk mengubah *output* dari matrix menjadi vektor satu dimensi. Selanjutnya, ditambahkan *Danse* dengan jumlah 64 neuron kemudian diaktifkan menggunakan fungsi aktivasi ReLU. Disisipkan *dropout layer* sebanyak 0.5 sebagai *output layer* untuk mengurangi risiko *overfitting*, kemudian digunakan *danse layer* sebanyak 3 unit sesuai dengan jumlah kelas COVID-19, normal, dan pneumonia, serta diaktifkan menggunakan fungsi aktivasi *softmax* agar menghasilkan probabilitas untuk masing-masing kelas. Selanjutnya, dilakukan proses *freeze* seluruh layer pada *basemodel* agar bobot pretrained dari ImageNet tidak ikut ter-*update* saat proses pelatihan. Model kemudian di-*compile* menggunakan *optimizer Adam*, *loss function*, dan matrix akurasi.

2.3.2 Pelatihan Model

Setelah model selesai dibangun, dilakukan proses pelatihan menggunakan *optimizer Adam* dengan berbagai variasi *learning rate* yaitu 0.1, 0.01, 0.001, 0.0001, 0.00001. Selain itu, dilakukan pengujian terhadap kombinasi parameter *epoch* sebanyak 20, 30, dan 50, serta *batch size* sebesar 8, 16, dan 32. Tujuan dari variasi parameter ini agar menemukan kombinasi terbaik yang dapat menghasilkan performa yang optimal dalam proses klasifikasi.

2.4 Evaluasi Model

Setelah melakukan pelatihan model, model dievaluasi menggunakan *classification report* yang menampilkan *precision*, *recall*, *f1-score*, dan *accuracy* untuk masing-masing kelas. Selain itu, performa model divisualisasikan dalam bentuk *confusion matrix* untuk menggambarkan lebih rinci seberapa akurat model dalam mengklasifikasikan citra *x-ray* dada ke dalam tiga kelas [24]. Nilai-nilai matrix evaluasi diperoleh menggunakan rumus sebagai berikut :

		Actual Values	
		1 (Positive)	0 (Negative)
Predicted Values	1 (Positive)	<p>TP (True Positive)</p>	<p>FP (False Positive) Type I Error</p>
	0 (Negative)	<p>FN (False Negative) Type II Error</p>	<p>TN (True Negative)</p>

Gambar 4. Confusion Matrix

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

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Keterangan :

TP = Jumlah data positif yang diklasifikasi benar.

TN = Jumlah data negatif yang diklasifikasi benar.

FP = Jumlah data positif yang diklasifikasi salah.

FN = Jumlah data negatif yang diklasifikasi salah

2.5 Implementasi XAI

Pada penelitian ini, diterapkan metode *Explainable AI (XAI)* untuk memberikan penjelasan terhadap hasil prediksi model *deep learning* yang bersifat *black-box* atau tidak dapat dijelaskan [16]. Implementasi XAI bertujuan untuk meningkatkan transparansi dan interpretabilitas hasil prediksi, sehingga pengguna seperti dokter atau tenaga medis dapat memahami alasan dibalik keputusan yang dihasilkan oleh model [18]. Salah satunya menggunakan metode *Shapley Additive Explanations (SHAP)* yang bisa menunjukkan seberapa besar kontribusi tiap fitur terhadap keputusan model.

2.5.1 Shapley Additive explanations (SHAP)

Shapley Additive Explanations (SHAP) adalah teknik *explainability* yang diperkenalkan oleh Lundberg and Lee [20]. Konsep dasar metode ini berasal dari nilai *Shapley*, yang pertama kali diperkenalkan oleh Lloyd Shapley pada tahun 1953 dengan teori kooperatif [19]. SHAP termasuk ke dalam metode *explainable AI (XAI)* yang dapat menjelaskan kontribusi masing-masing fitur terhadap prediksi model [18]. Nilai *Shapley* dari teori permainan yang menjadi dasar perhitungan nilai kontribusi dalam metode SHAP, sehingga dapat menjelaskan seberapa besar pengaruh setiap fitur dalam proses pengambilan keputusan [25]. Rumus perhitungan nilai *shapley* adalah sebagai berikut:

$$\varphi_j(v) = \varphi_j = \sum_{S \subseteq M \setminus \{j\}} \frac{|S|!(M - |S| - 1)!}{M!} (v(S \cup \{j\}) - v(S)) \quad (5)$$

Keterangan:

M = himpunan semua fitur dalam model.

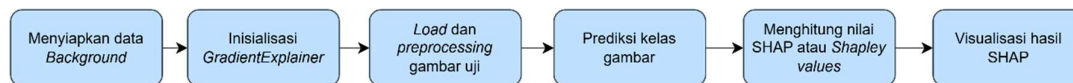
S = himpunan subset fitur tanpa fitur j.

$f(S)$ = nilai prediksi model ketika hanya fitur dalam S yang digunakan.

$f(S \cup \{j\})$ = nilai prediksi model ketika fitur j ditambahkan ke subset S.

$\frac{|S|!(M - |S| - 1)!}{M!}$ = adalah bobot probabilitas bahwa suatu subset S dipilih dalam perhitungan

Proses implementasi SHAP untuk menjelaskan hasil prediksi model *deep learning* berbasis citra *x-ray* dada, dilakukan melalui beberapa tahapan sebagai berikut :



Gambar 5. Proses Implementasi SHAP

SHAP dilakukan melalui beberapa tahapan sebagai berikut :

1. Menyiapkan data *background*

Data *background* merupakan sekumpulan citra dari dataset validasi yang digunakan sebagai *baseline* dalam proses perhitungan nilai SHAP. Data ini di muat menggunakan *ImageDataGenerator* dengan proses normalisasi piksel ke skala 0-1, kemudian diambil beberapa batch gambar untuk dijadikan *background*.

2. Inisialisasi *GradientExplainer*
Setelah data *background* siap, tahap berikutnya adalah menginisialisasi *explainer* dengan menggunakan *GradientExplainer* dari library SHAP. *Explainer* ini memerlukan model *deep learning* yang telah dilatih serta data *background* sebagai acuan dalam menghitung *baseline* prediksi [18].
3. *Load* dan *preprocessing* gambar uji
Citra uji yang dianalisis dimuat dari direktori penyimpanan, kemudian diubah ke dalam *array*, di-*resize* ke ukuran 224x224 piksel, dan dinormalisasi ke skala 0-1 agar sesuai dengan format input model *deep learning*.
4. Prediksi kelas gambar
Citra uji yang telah diproses kemudian diprediksi menggunakan model. Hasil prediksi berupa nilai probabilitas untuk masing-masing kelas, lalu dipilih kelas dengan probabilitas tertinggi sebagai hasil akhir prediksi.
5. Menghitung nilai SHAP (*SHAP values*)
Nilai SHAP dihitung menggunakan fungsi `explainer.shap.values()`. Nilai ini menunjukkan seberapa besar kontribusi setiap piksel terhadap prediksi yang dihasilkan oleh model untuk masing-masing kelas.
6. Visualisasi hasil SHAP
Nilai SHAP yang telah dihitung kemudian divisualisasikan dalam bentuk heatmap menggunakan *Matplotlib*. Heatmap ini menampilkan citra asli beserta peta kontribusi piksel terhadap masing-masing kelas. Warna merah menunjukkan area dengan kontribusi positif terhadap piksel, sedangkan warna biru menunjukkan kontribusi negatif [18].

3. HASIL DAN PEMBAHASAN

Pada penelitian ini, dilakukan serangkaian percobaan dengan beberapa skenario pengujian yang digunakan untuk mengevaluasi kinerja model. Setiap skenario menggunakan parameter pelatihan yang berbeda-beda, pembagian dataset, jenis *optimizer*, *activation function*, nilai *learning rate*, *batch size*, dan jumlah *epoch*. Tujuan dari percobaan ini adalah untuk mencari konfigurasi terbaik yang mampu memberikan performa yang optimal dalam mengklasifikasikan citra *x-ray* dada ke dalam tiga kategori COVID-19, normal, dan pneumonia.

Tabel 1 menunjukkan daftar parameter pelatihan yang digunakan dalam penelitian ini. Parameter dengan yang meliputi rasio pembagian data, jumlah *epoch*, *learning rate*, dan *batch size*. *Optimizer* yang digunakan adalah Adam dengan *activation function* ReLU pada seluruh percobaan.

Table 1. Parameter pelatihan model

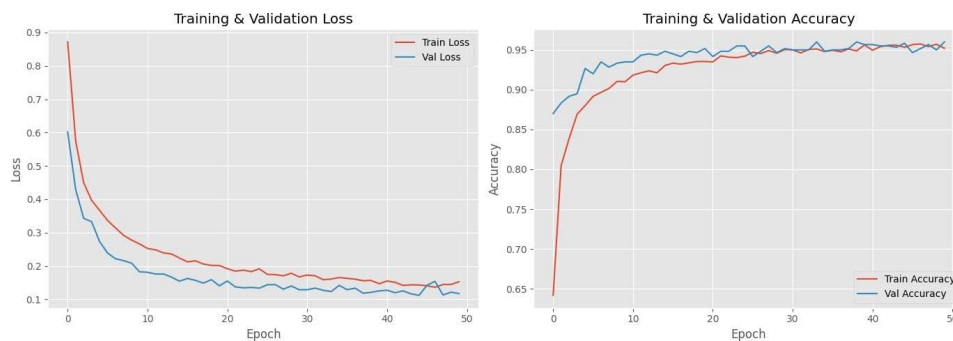
Parameter	Nilai
Pembagian data	70% : 30%, 80% : 20%, 90% : 10%
<i>Optimizer</i>	Adam
EPOCH	20, 30, dan 50
<i>Activation Function</i>	ReLU
<i>Batch Size</i>	8, 16, dan 32
<i>Learning Rate</i>	0.1, 0.01, 0.001, 0.0001, 0.00001

Setelah model dilatih menggunakan kombinasi parameter yang berbeda, diperoleh hasil akurasi, nilai loss, serta akurasi validasi dan nilai loss validasi pada setiap skenario pengujian. Hasil setiap kombinasi dapat dilihat pada tabel 2 berikut

Table 2. Seluruh hasil pengujian

No.	Rasio	Learning Rate	EPOCH	Batch Size	Accuracy	Val Accuracy	Loss	Val Loss
1	70%:30%	0.1	20	8	34.88	33.11	1.1176	1.1034
2	70%:30%	0.01	30	16	82.13	93.33	0.4134	0.1934
3	70%:30%	0.001	50	32	94.04	93.12	0.1935	0.1813
4	70%:30%	0.0001	30	32	85.3	87.71	0.4326	0.3694
5	70%:30%	0.00001	30	32	59.23	79.32	0.2994	0.8738
6	80%:20%	0.1	20	8	34.18	33.33	1.1067	1.1094
7	80%:20%	0.01	30	16	82.97	91.67	0.4084	0.2885
8	80%:20%	0.001	50	32	95.75	96.00	0.1408	0.1175
9	80%:20%	0.0001	30	32	86.18	87	0.4296	0.3673
10	80%:20%	0.00001	50	32	68.11	80.83	0.8335	0.7657
11	90%:10%	0.1	20	8	32.41	33.33	1.1072	1.1027
12	90%:10%	0.01	30	16	85.99	93.11	0.3573	0.2341
13	90%:10%	0.001	50	32	93.27	94.52	0.2089	0.1663
14	90%:10%	0.0001	30	32	86.38	89.93	0.4162	0.3279
15	90%:10%	0.00001	30	32	57.26	77.67	0.9185	0.8494

Berdasarkan tabel 2, terlihat bahwa model dengan rasio 80%:20% dengan *learning rate* 0.001, *batch size* 32, dan jumlah *epoch* 50. Nilai akurasi validasi tertinggi diperoleh pada konfigurasi dengan rasio 80% : 20% yaitu 96.00% dengan nilai loss validasi 0.1175%. Sebaliknya, model dengan *learning rate* 0.1 menghasilkan akurasi yang sangat rendah pada seluruh skenario, dengan kisaran angka 33-34%, baik pada data training ataupun validasi. Hal ini menunjukkan bahwa nilai *learning rate* yang terlalu besar dapat menyebabkan proses pembelajaran model menjadi tidak stabil. Selain menggunakan tabel, hasil training dan validasi model divisualisasikan ke dalam bentuk grafik untuk mempermudah analisis. Gambar 7 menunjukkan grafik perubahan nilai akurasi dan loss selama proses pelatihan.



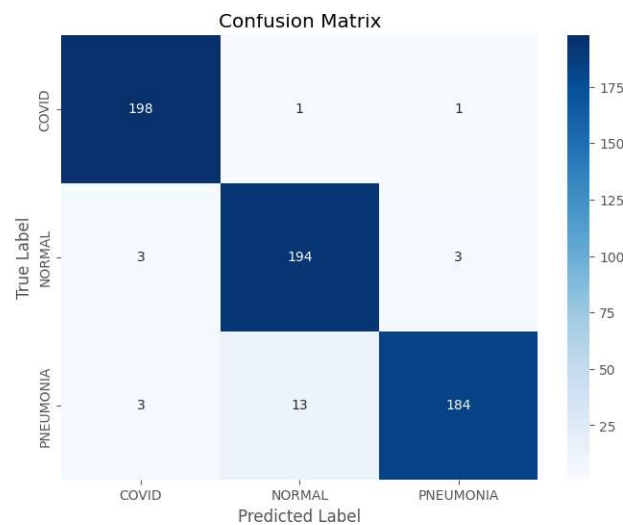
Gambar 6. Grafik akurasi

Gambar diatas menunjukkan bahwa akurasi pada data training dan validasi sama sama mengalami peningkatan hingga 90%. Nilai *loss* juga terus turun dan stabil di bawah 0.2 setelah melewati *epoch* ke-20. Hasil menunjukkan model belajar dengan baik tanpa *overfitting*, karena performa training dan validasi seimbang sampai akhir epoch ke-50. Selain grafik akurasi, dilakukan pengujian menggunakan matrix evaluasi numerik seperti *precision*, *recall*, dan *F1-score* untuk mengetahui performa model dalam mengklasifikasikan masing-masing kelas. Hasilnya dapat dilihat pada tabel 3.

Table 3. Classification report

	Precision	Recall	F1-Score	Support
COVID-19	97	99	98	200
Normal	93	97	95	200
Pneumonia	98	92	95	200
Accuracy			96	600

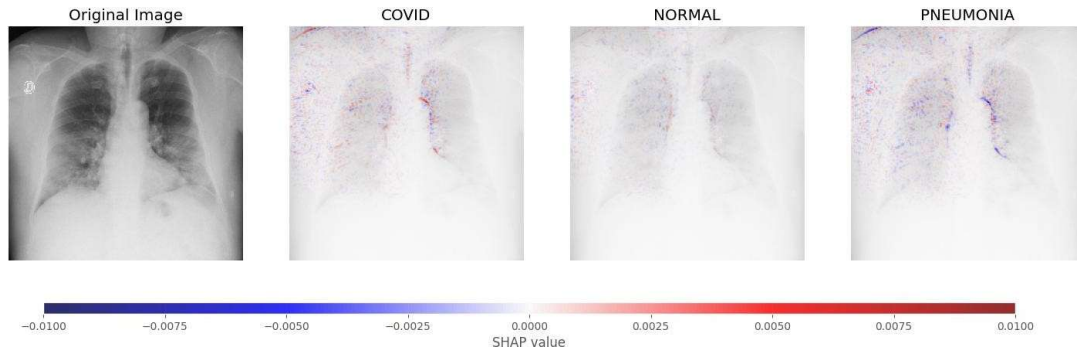
Selain menggunakan matrix evaluasi numerik seperti *precision*, *recall*, dan *F1-score*, dilakukan juga analisis menggunakan *confusion matrix* untuk melihat jumlah prediksi benar dan salah pada masing-masing kelas. Gambar berikut merupakan *confusion matrix* dari hasil prediksi model terbaik :



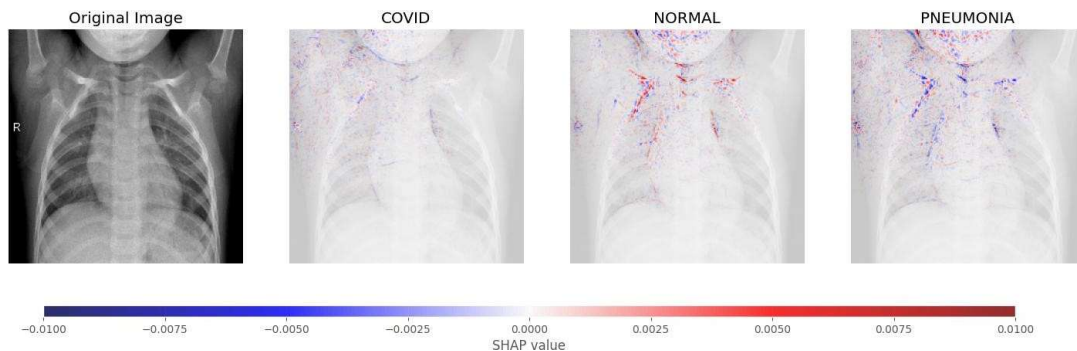
Gambar 7. Confusion matrix

Berdasarkan *confusion matrix* diatas, dapat dilihat bahwa untuk kelas COVID-19 berhasil prediksi 198 benar, sementara 1 data salah prediksi sebagai normal, dan 1 sebagai pneumonia. Pada kelas normal terdapat 194 data yang diprediksi benar, 3 data salah prediksi sebagai COVID-19, dan 3 data salah prediksi sebagai pneumonia. Sedangkan kelas pneumonia berhasil prediksi 184 data prediksi benar, 3 data salah prediksi sebagai COVID-19, dan 13 data salah prediksi sebagai normal.

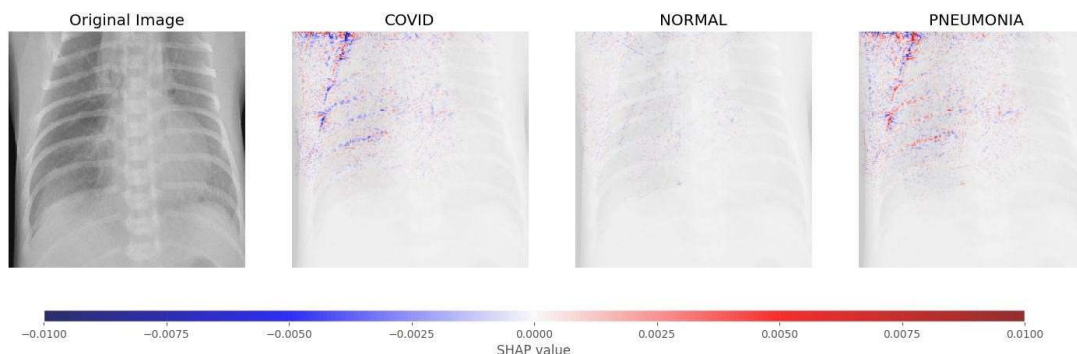
Selain mengevaluasi performa model menggunakan matrix numerik dan *confusion matrix*, penelitian ini juga melakukan interpretasi terhadap hasil prediksi menggunakan metode *Shapley Additive Explanations* (SHAP) untuk memberikan gambaran visual terkait kontribusi masing-masing piksel terhadap hasil prediksi yang dihasilkan oleh model. Proses interpretasi dilakukan dengan menerapkan SHAP dengan model terbaik, yaitu model dengan rasio data 80% : 20%, *learning rate* 0.001, *batch size* 32, dan *epoch* 50. Hasil visualisasi SHAP ditampilkan dalam bentuk heatmap, di mana warna merah menunjukkan area piksel yang memberikan kontribusi positif terhadap prediksi kelas, sedangkan warna biru menunjukkan kontribusi negatif atau tidak terlalu berpengaruh. *Colorbar* pada bagian bawah gambar berfungsi untuk menunjukkan rentang nilai SHAP yang dihasilkan, yaitu -0.01 hingga 0.01 nilai positif di sisi kanan ditandai dengan gradiasi dari putih ke merah, sedangkan nilai negatif di sisi kiri ditandai dengan gradiasi dari putih ke biru. Semakin mendekati angka 0, warna yang ditampilkan akan semakin putih, yang berarti piksel di area tersebut bersifat netral dan tidak memberikan pengaruh signifikan terhadap keputusan model.



Gambar 8. Berdasarkan nilai SHAP tertinggi pada kolom kedua, citra x-ray terdiagnosis COVID-19



Gambar 9. Berdasarkan nilai SHAP tertinggi pada kolom ketiga, citra x-ray terdiagnosis NORMAL



Gambar 10. Berdasarkan nilai SHAP tertinggi pada kolom keempat, citra x-ray terdiagnosis PNEUMONIA

Hasil eksperimen menunjukkan bahwa model VGG-16 dengan konfigurasi terbaik, yaitu rasio 80% : 20%, *Learning rate* 0.001, *batch size* 32, dan 50 *epoch*, mampu mencapai akurasi tinggi yaitu 96,00%. Capaian ini membuktikan bahwa model mampu mengenali pola-pola penting pada citra *x-ray* dada untuk membedakan kondisi paru-paru COVID-19, normal, dan pneumonia secara akurat, meskipun gejala klinis ketiganya seringkali serupa. Penerapan metode SHAP berhasil memberikan visualisasi heatmap yang informatif, menyoroti area citra yang berkontribusi terhadap keputusan model dengan menampilkan fitur-fitur relevan terkait penyakit paru. Interpretasi ini tidak hanya meningkatkan kepercayaan terhadap hasil klasifikasi, tetapi juga menjadi alat pendukung penting dalam pengembangan sistem keputusan berbasis *deep learning* di bidang medis, di mana transparansi alasan prediksi sangat dibutuhkan oleh tenaga kesehatan.

4. KESIMPULAN

Hasil penelitian menunjukkan bahwa model *deep learning* menggunakan arsitektur VGG-16 efektif dalam mengklasifikasikan citra *x-ray* dada untuk mendeteksi penyakit COVID-19, kondisi normal, dan pneumonia. Model menunjukkan performa terbaik pada rasio data 80% : 20%, menggunakan *optimizer* Adam, *learning rate* 0.001, *epoch* sebanyak 50, dan *batch size* 32. Konfigurasi ini, model berhasil mencapai akurasi pelatihan sebesar 95.75% dan akurasi validasi 96.00% tanpa indikasi *overfitting*. Selain itu, penerapan metode

SHAP berhasil memberikan visualisasi *heatmap* yang informatif, dengan area penting pada citra yang sejalan dengan hasil prediksi model. Pada penelitian selanjutnya, disarankan untuk melakukan uji validasi interpretasi hasil metode SHAP secara langsung kepada tenaga medis, khususnya dokter radiologi, untuk dinilai manfaatnya dalam mendukung proses diagnosis.

REFERENSI

- [1] A. A. D. Halim and S. Anraeni, "Analisis Klasifikasi Dataset Citra Penyakit Pneumonia menggunakan Metode K-Nearest Neighbor (KNN)," *Indones. J. Data Sci.*, vol. 2, no. 1, pp. 01–12, 2021, doi: 10.33096/ijodas.v2i1.23.
- [2] A. Kumar *et al.*, "Wuhan to World: The COVID-19 Pandemic," *Front. Cell. Infect. Microbiol.*, vol. 11, no. March, pp. 1–21, 2021, doi: 10.3389/fcimb.2021.596201.
- [3] S. Sharma and K. Guleria, "A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks," *Procedia Comput. Sci.*, vol. 218, pp. 357–366, 2022, doi: 10.1016/j.procs.2023.01.018.
- [4] M. Bhandari, T. B. Shahi, B. Siku, and A. Neupane, "Explanatory classification of CXR images into COVID-19, Pneumonia and Tuberculosis using deep learning and XAI," *Comput. Biol. Med.*, vol. 150, no. September, 2022, doi: 10.1016/j.compbiomed.2022.106156.
- [5] N. Umzubongile, S. Mandindi, H. Vadapalli, and A. Mathematics, "Detecting and Understanding COVID-19 Misclassifications : A Deep Learning and Explainable AI Approach," no. August, 2023.
- [6] M. Bhandari, T. Bahadur, B. Siku, and A. Neupane, "Explanatory classification of CXR images into COVID-19 , Pneumonia and Tuberculosis using deep learning and XAI," *Comput. Biol. Med.*, vol. 150, no. October, p. 106156, 2022, doi: 10.1016/j.compbiomed.2022.106156.
- [7] A. V. Ikechukwu, "Leveraging Transfer Learning for Efficient Diagnosis of COPD Using CXR Images and Explainable AI Techniques," *Intel. Artif.*, vol. 27, no. 74, pp. 133–151, 2024, doi: 10.4114/intartif.vol27iss74pp133-151.
- [8] J. Yopento, E. Emawati, and F. F. Coastera, "Identifikasi Pneumonia Pada Citra X-Ray Paru-Paru Menggunakan Metode Convolutional Neural Network (CNN) Berdasarkan Ekstraksi Fitur Sobel," *Rekursif J. Inform.*, vol. 10, no. 1, pp. 40–47, 2022, doi: 10.33369/rekursif.v10i1.17247.
- [9] R. A. Tilasefana and R. E. Putra, "Penerapan Metode Deep Learning Menggunakan Algoritma CNN Dengan Arsitektur VGG NET Untuk Pengenalan Cuaca," *J. Informatics Comput. Sci.*, vol. 05, no. 1, pp. 48–57, 2023.
- [10] M. Adil, R. Ullah, S. Noor, and N. Gohar, "Effect of number of neurons and layers in an artificial neural network for generalized concrete mix design," *Neural Comput. Appl.*, vol. 8, pp. 1–9, 2020, doi: 10.1007/s00521-020-05305-8.
- [11] M. A. Sufian *et al.*, "AI-Driven Thoracic X-ray Diagnostics: Transformative Transfer Learning for Clinical Validation in Pulmonary Radiography," *J. Pers. Med.*, vol. 14, no. 8, 2024, doi: 10.3390/jpm14080856.
- [12] I. M. R and T. M. Johan, "Klasifikasi Citra Ikan Menggunakan Algoritma Convolutional Neural Network dengan Arsitektur VGG-16," vol. 4, no. 2, pp. 978–985, 2023, doi: 10.30865/klik.v4i2.1209.
- [13] P. P. Illahi, H. Fauzi, and T. S. Siadari, "Klasifikasi Penyakit Pneumonia Dan Covid-19 Berbasis Citra X-Ray Menggunakan Arsitektur Deep Residual Network Classification Of Pneumonia And Covid-19 Based On X-Ray Images Using The Deep Residual Network Architecture," vol. 9, no. 4, pp. 1837–1843, 2022, [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/18232>
- [14] T. Berliani *et al.*, "Perbandingan Kemampuan Klasifikasi Citra X-ray Paru-paru menggunakan Transfer Learning ResNet-50 dan VGG-16 VGG-16 Transfer Learning," vol. 5, no. 2, pp. 123–135, 2023.
- [15] R. Rajpoot, M. Gour, S. Jain, and V. B. Semwal, "Integrated ensemble CNN and explainable AI for COVID-19 diagnosis from CT scan and X-ray images," *Sci. Rep.*, vol. 14, no. 1, pp. 1–21, 2024, doi: 10.1038/s41598-024-75915-y.
- [16] P.-D. Arsenaute, S. Wang, and J.-M. Patenande, "A Survey of Explainable Artificial Intelligence (XAI) in Financial Time Series Forecasting," vol. 1, no. 1, pp. 1–35, 2024, [Online]. Available: <http://arxiv.org/abs/2407.15909>
- [17] S. Walia, K. Kumar, S. Agarwal, and H. Kim, "Using XAI for Deep Learning-Based Image Manipulation Detection with Shapley Additive Explanation," *Symmetry (Basel)*, vol. 14, no. 8, 2022, doi: 10.3390/sym14081611.
- [18] E. Mahamud, N. Fahad, M. Assaduzzaman, S. M. Zain, K. O. M. Goh, and M. K. Morol, "An explainable artificial intelligence model for multiple lung diseases classification from chest X-ray images using fine-tuned transfer learning," *Decis. Anal. J.*, vol. 12, no. July, p. 100499, 2024, doi: 10.1016/j.dajour.2024.100499.
- [19] H. Sahlaoui, E. A. A. Alaoui, A. Nayyar, S. Agoujil, and M. M. Jaber, "Predicting and Interpreting Student Performance Using Ensemble Models and Shapley Additive Explanations," *IEEE Access*, vol. 9, pp. 152688–152703, 2021, doi: 10.1109/ACCESS.2021.3124270.
- [20] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," *Adv. Neural Inf. Process. Syst.*, vol. 2017-Decem, no. Section 2, pp. 4766–4775, 2017.
- [21] doi: 10.17632/dvntn9yhd2. . Kumar, Sachin (2022), "Covid19-Pneumonia-Normal Chest X-Ray Images", Mendeley Data, V1, "Covid19-Pneumonia-Gambar Rontgen Dada Normal."
- [22] M. Yaseliani, A. Z. Hamadani, A. I. Maghsoodi, and A. Mosavi, "Pneumonia Detection Proposing a Hybrid Deep Convolutional Neural Network Based on Two Parallel Visual Geometry Group Architectures and Machine Learning Classifiers," *IEEE Access*, vol. 10, pp. 62110–62128, 2022, doi: 10.1109/ACCESS.2022.3182498.
- [23] D. K. Jain *et al.*, "Deep Learning-Aided Automated Pneumonia Detection and Classification Using CXR Scans," *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/7474304.
- [24] M. Vakili, M. Ghamsari, and M. Rezaei, "Performance Analysis and Comparison of Machine and Deep Learning Algorithms for IoT Data Classification," 2020, [Online]. Available: <http://arxiv.org/abs/2001.09636>
- [25] T. Chrysoula and I. M. Systems, "Explainable Artificial Intelligence for Deep Learning Methods in Chest X-Ray Classification," no. February, 2023.