

ANALISIS PERBANDINGAN ALGORITMA *MACHINE LEARNING* DAN *DEEP LEARNING* UNTUK KLASIFIKASI CITRA SISTEM ISYARAT BAHASA INDONESIA (SIBI)

Mohammad Farid Naufal^{*1}, Selvia Ferdiana Kusuma²

¹Universitas Surabaya, Surabaya, ²Politeknik Elektronika Negeri Surabaya

Email: ¹faridnaufal@staff.ubaya.ac.id, ²selvia@pens.ac.id

*Penulis Korespondensi

(Naskah masuk: 19 Desember 2022, diterima untuk diterbitkan: 26 Juli 2023)

Abstrak

Terdapat orang yang tidak mampu berkomunikasi secara verbal yang menyebabkan kesulitan dalam berkomunikasi. Orang tersebut mengalami gangguan seperti tuli atau bisu. Mereka hanya dapat berkomunikasi melalui bahasa isyarat salah satunya adalah Sistem Isyarat Bahasa Indonesia (SIBI). Pengenalan Bahasa Isyarat adalah permasalahan klasifikasi yang kompleks untuk dipecahkan. Setiap bahasa isyarat memiliki sintaks dan tata bahasanya sendiri. Computer vision adalah sebuah Teknik yang digunakan komputer untuk melakukan klasifikasi citra. Computer vision membantu pengenalan citra SIBI secara otomatis sehingga memudahkan orang normal berkomunikasi dengan orang tuli atau bisu. Pada penelitian sebelumnya belum ada yang melakukan perbandingan algoritma klasifikasi *machine learning* dan *deep learning* untuk pengenalan SIBI. Perbandingan penting dilakukan untuk melihat efektifitas tiap algoritma klasifikasi dalam hal performa klasifikasi dan waktu komputasi. Algoritma klasifikasi *machine learning* memiliki waktu komputasi lebih rendah sedangkan *Deep learning* memiliki performa klasifikasi lebih tinggi. Penelitian ini menganalisis *time to performance* dari algoritma *machine learning* dan *deep learning* dalam melakukan klasifikasi citra SIBI huruf A hingga Z. *K-Nearest Neighbors* (KNN), *Support Vector Machine* (SVM), dan *Convolutional neural network* (CNN) dengan *transfer learning* adalah tiga algoritma klasifikasi populer yang dibandingkan dalam penelitian ini. Arsitektur *transfer learning* yang digunakan adalah Xception, ResNet50, VGG15, dan MobileNetV2. Dari hasil penelitian yang dilakukan menggunakan 5 *cross validation*, CNN dengan arsitektur Xception memiliki nilai *F1 Score* tertinggi yaitu 99,57% dengan waktu *training* rata-rata 1.387 detik. Sedangkan KNN dengan nilai K = 1 memiliki waktu *training* tercepat yaitu 0,03 detik dan memiliki nilai *F1 Score* 86,95%.

Kata kunci: klasifikasi SIBI, KNN, SVM, CNN, machine learning, deep neural network

COMPARISON ANALYSIS OF MACHINE LEARNING AND DEEP LEARNING ALGORITHMS FOR IMAGE CLASSIFICATION OF INDONESIAN LANGUAGE SIGNING SYSTEMS (SIBI)

Abstract

The person who has a disorder such as deaf or dumb are unable to communicate verbally, which causes difficulties in communicating. They can only communicate through sign language, one of which is the Indonesian Language Sign System or Sistem Isyarat Bahasa Indonesia (SIBI). Sign Language Recognition is a complex classification problem to solve. Each sign language has its syntax and grammar. Computer vision is a technique used by computers to classify images. Computer vision helps automatically recognize SIBI images, making it easier for normal people to communicate with deaf or mute people. In previous studies, no one has compared machine learning and deep learning classification algorithms for the classification of SIBI. Therefore, a meaningful comparison is made to see each classification algorithm's effectiveness in classification performance and computation time. Machine learning classification algorithms have lower computation time, while Deep learning has higher classification performance. This study analyzes the time to performance of machine learning and deep learning algorithms in classifying SIBI images of letters A to Z. *K-Nearest Neighbors* (KNN), *Support Vector Machine* (SVM), and *Convolutional neural network* (CNN) with *transfer learning* are three popular classification algorithms compared in this study. The transfer learning architectures used are Xception, ResNet50, VGG15, and MobileNetV2. The results of research conducted using 5 cross-validation, CNN with the Xception architecture has highest *F1 Score* of 99.57%, with an average training time of 1.387 seconds. KNN, with a value of K = 1, has the fastest training time of 0.03 seconds and an *F1 Score* of 86.95%.

Keywords: SIBI Classification, KNN, SVM, CNN, machine learning, deep neural network

1. PENDAHULUAN

Kehidupan sehari-hari lebih sulit bagi orang yang memiliki gangguan pendengaran. Orang yang memiliki gangguan pendengaran dan komunikasi atau orang tuli dan bisu kesulitan berkomunikasi dengan orang normal. Teknik komunikasi berbasis penglihatan dalam bahasa isyarat memudahkan komunikasi bagi orang tuli dan bisu. Bahasa isyarat adalah metode untuk berkomunikasi antara sesama orang tuli dan bisu atau dengan orang normal. Bahasa isyarat terdiri dari gerakan jari, tangan, gerakan, tubuh, dan wajah yang jika digabungkan akan mewakili kata yang ingin disampaikan (Biasa, 2002). Salah satu Bahasa isyarat yang berlaku di Indonesia adalah Sistem Isyarat Bahasa Indonesia (SIBI). Namun, bahasa SIBI awam dipahami oleh kebanyakan orang. Hal ini membuat sulit bagi orang normal untuk berkomunikasi dengan orang tuli atau bisu. Untuk memudahkan orang normal dalam memahami bahasa isyarat, diperlukan suatu sistem yang dapat menerjemahkan bahasa isyarat secara otomatis dan akurat.

Terdapat beberapa penelitian yang pernah dilakukan sebelumnya. (Suharjito, et al. 2021) mengimplementasikan model *Convolutional neural network* (CNN) yang dikombinasikan dengan *transfer learning* dari ImageNet dan Kinectic untuk mengatasi masalah dataset dengan jumlah kecil. Dataset dikumpulkan dengan menggunakan kamera ponsel Samsung Galaxy S6, sensor Sony Exmore RS IMX240 16 megapiksel. Dataset yang digunakan 200 video yang terdiri dari 10 kata (kelas) dan 2 penanda tangan. Akurasi validasi tertinggi adalah 100% dan akurasi pengujian tertinggi adalah 97,50%. Hasil terbaik diperoleh dengan menggunakan model *transfer learning* Inception.

(Erdefi Rakun et al, 2022) menggunakan fitur skeleton dan *hand shape* untuk mengklasifikasikan SIBI. Hasil akurasi yang dicapai dengan metode fusi fitur adalah akurasi dengan urutan menurun: 88,016% saat menggunakan *sequence feature vector concatenation*, 85,448% saat menggunakan *Conneau feature vector concatenation*, 83,723% saat menggunakan *feature-vector concatenation*, dan 49,618% saat menggunakan penggabungan fitur sederhana. Model klasifikasi yang digunakan adalah 2-layer bidirectional LSTM.

(Yusnita et al, 2017) mengimplementasikan pengenalan gerakan tangan statis SIBI abjad dari "A" hingga "Z", angka dari "0" hingga "9", dan tanda baca tambahan seperti "Titik", "Tanda Tanya", dan "Spasi". Metode klasifikasi yang digunakan adalah Artificial Neural Network (ANN) Dataset yang digunakan terdiri dari 100 citra untuk setiap gestur dengan akurasi 90%.

(Sindarto et al, 2022) melakukan klasifikasi SIBI menggunakan CNN dengan arsitektur EfficientDet-Lite 4 terhadap 15 kosakata SIBI.

Akurasi yang didapatkan dari penelitian ini adalah 88%.

(Darmatasia et al, 2021) melakukan klasifikasi citra SIBI yang terdiri dari huruf abjad A hingga E menggunakan CNN dengan dan tanpa filter gradient. Akurasi yang didapat dengan menggunakan filter adalah 98%, namun jika tanpa menggunakan filter hanya 25%.

(Sholawati et al, 2022) melakukan klasifikasi menggunakan CNN untuk citra SIBI abjad A hingga Z. Akurasi, *recall*, *Specificity*, dan *sensitivity* yang dihasilkan dari penelitian ini adalah 80.76%. Dataset citra yang digunakan berjumlah 416 yang terdiri dari 384 citra dari hasil pengambilan foto dan 32 hasil pengambilan dari frame video.

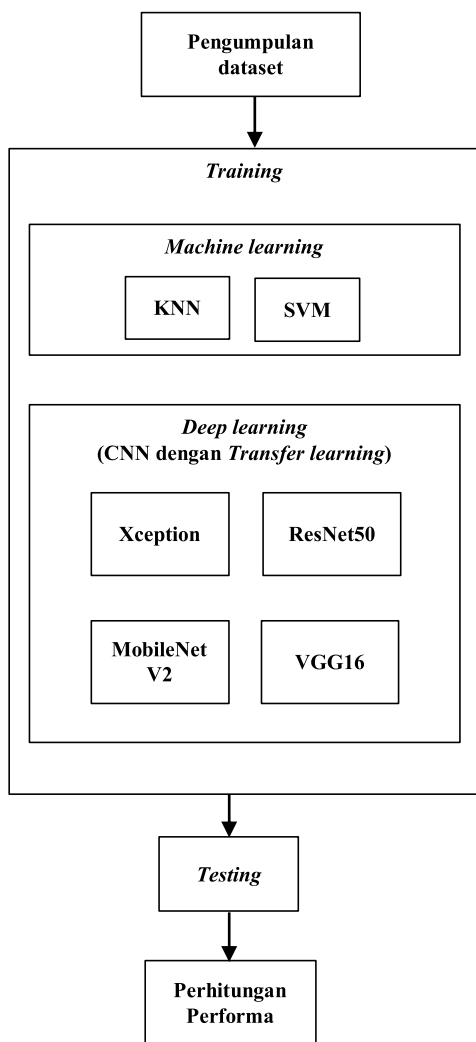
(Legowo, Sumpeno and Premunanto, 2017) melakukan klasifikasi 10 gerakan SIBI menggunakan *naïve bayes* dengan ekstraksi fitur ciri sebanyak 19. Akurasi yang didapatkan dari penelitian ini adalah 70.7%. 10 gerakan SIBI yang diklasifikasikan adalah huruf A, B, C, D, U, "halo", "ajak", "kacau", "mana", dan "bukan".

Dari beberapa penelitian sebelumnya belum ada yang melakukan analisis perbandingan algoritma *machine learning* dan *deep learning* untuk klasifikasi citra SIBI huruf A hingga Z. Citra SIBI yang berupa gambar pergelangan tangan merupakan citra yang sederhana dan mudah untuk diklasifikasikan. Namun perlu dianalisis perbandingan performa dari *machine learning* dan *deep learning* untuk melihat efektifitas tiap metode klasifikasi dengan mempertimbangkan waktu *training* algoritma dan perhitungan *metric* performa *accuracy*, *precision*, *recall*, dan *F1 Score*. Penelitian ini membandingkan waktu *training* dan performa algoritma klasifikasi *machine learning* (KNN dan SVM) dan *deep learning* (CNN dengan arsitektur *transfer learning* Xception, ResNet50, VGG16, MobileNetV2) yang berguna untuk pertimbangan pemilihan algoritma yang cocok untuk klasifikasi citra SIBI.

Sistematika penulisan pada penelitian ini terdiri dari 4 bagian. Bab 1 menjelaskan mengenai pendahuluan dan latar belakang permasalahan, bab 2 menjelaskan metode penelitian, bab 3 menjelaskan hasil dan pembahasan, dan bab 4 dijelaskan menjelaskan kesimpulan penelitian.

2. METODE PENELITIAN

Pada penelitian ini, terdapat empat tahapan penelitian yang dilakukan yaitu dimulai dari pengumpulan dataset, *training* model klasifikasi, *testing* model klasifikasi, dan perhitungan performa klasifikasi. Pada perhitungan performa klasifikasi juga dibandingkan performa dari tiap algoritma. Gambar 1 menunjukkan tahapan penelitian yang dilakukan. Semua tahapan tersebut dilakukan secara berurutan sesuai dengan kaidah.



Gambar 1. Tahapan penelitian

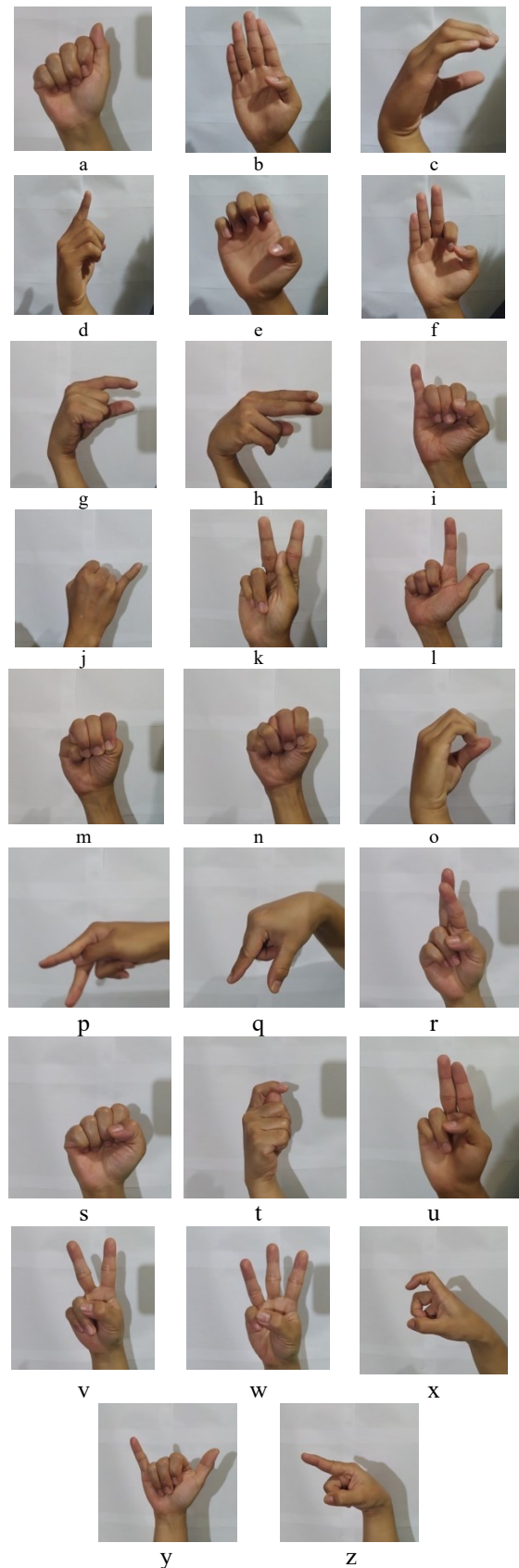
2.1. Pengumpulan Dataset

Dataset citra SIBI dalam penelitian ini didapatkan dari sebuah dataset publik kaggle (Afkaar, 2022). Alasan penggunaan dataset ini adalah bersifat publik dan dapat diakses oleh peneliti lain. Penggunaan dataset publik memudahkan perbandingan performa metode lain yang akan diusulkan oleh peneliti lain ke depan. Dataset terdiri dari 26 kelas yaitu huruf alfabet A hingga Z yang masing-masing kelas terdiri dari 19 citra. Tabel 1 menunjukkan detail dataset.

Tabel 1. Detail dataset SIBI

Jumlah Kelas	Jumlah citra tiap kelas
26 (Huruf A hingga Z)	19
Total jumlah citra	494

Tiap citra sebelum digunakan untuk proses *training* dan *testing* ukurannya *resize* dengan ukuran 64x64 piksel. Hal ini dilakukan agar semua data memiliki ukuran piksel yang sama.



Gambar 2. Contoh dataset alfabet SIBI dari huruf (a) hingga (z)

2.2. Training Model Klasifikasi

Setiap algoritma pada tahapan ini diatur parameter yang digunakan. Tujuannya adalah untuk

menganalisis parameter yang mempengaruhi performa klasifikasi. Jenis *distance* dan jumlah tetangga K adalah parameter yang diuji di algoritma KNN. Jenis jarak yang digunakan adalah Euclidean (Dokmanic et al., 2015) dan jumlah tetangga yang dipilih adalah 1, 2, 3, 4, dan 5. Parameter KNN yang digunakan dalam penyelidikan ini ditampilkan pada Tabel 2.

Tabel 2. Parameter KNN

Parameter	Deskripsi
Jumlah Neighbors	1,2,3,4,5
Tipe <i>Distance</i>	Euclidean

Parameter yang datur pada algoritma SVM adalah jenis *kernel*. *Kernel* yang digunakan adalah linear, poly, dan rbf. Parameter SVM dapat dilihat pada Tabel 3.

Tabel 3. Parameter *Kernel* SVM

Parameter	Deskripsi
<i>Kernel</i>	Linear, Poly, RBF

Pada algoritma CNN parameter yang diujicobakan adalah jumlah *epoch*, tipe *convolution*, tipe *activation function*, dan jumlah dense layer. Tabel 4 menunjukkan parameter dari CNN. Arsitektur CNN yang digunakan adalah LeNet (Lecun et al, 1998). Arsitektur *transfer learning* yang digunakan di CNN adalah Xception, ResNet50, VGG16 dan MobileNetv2.

Xception dipilih karena Xception memiliki jumlah parameter yang sama dengan Inceptionv3, dan telah terbukti lebih baik daripada InceptionV3 saat diujicobakan pada database ImageNet. Inception v3 dianggap kurang efektif dibandingkan Xception (Chollet, 2016).

Arsitektur ResNet50 dipilih karena telah menunjukkan kemampuannya dengan menjuarai kompetisi ILSVRC pada tahun 2015. ResNet50 adalah versi VGG19 yang menyertakan jaringan residual, menjadikannya model yang rumit dengan parameter yang dapat dikelola yang dapat berjalan lebih cepat selama komputasi (Victor Ikechukwu et al., 2021).

VGG16 dipilih karena telah menunjukkan dirinya dengan menempati posisi kedua dalam kompetisi *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC) 2014. VGG16 adalah model yang kompleks dengan waktu perhitungan yang lambat, namun dapat menghasilkan hasil klasifikasi yang akurat, menjadikannya pilihan yang sangat baik untuk perbandingan (Simonyan and Zisserman, 2015).

MobileNetV2 dipilih karena memiliki model yang tidak kompleks dan ringan dengan latensi rendah dan waktu komputasi yang cepat (Sandler et al., 2018). MobileNetV2 sesuai untuk digunakan dengan perangkat elektronik portabel.

Arsitektur *transfer learning* tersebut digunakan untuk mengekstraksi fitur yang selanjutnya akan diklasifikasikan oleh *top layer*. *Top layer* terdiri dari

satu hidden layer yang berisikan 128 *neuron* dengan *activation function* ReLu dan satu *output layer* yang berisikan 26 *neuron* dengan *activation function* Softmax. Jumlah *neuron* pada *output layer* menyesuaikan dengan jumlah label yang diklasifikasikan. Optimizer yang digunakan adalah adam dan loss function yang digunakan adalah categorical crossentropy. Tabel 4 menunjukkan parameter yang digunakan untuk CNN.

Tabel 4. Parameter CNN

Parameter	Deskripsi
Arsitektur	Xception, ResNet50, VGG16, MobileNetv2
Top layer	1 hidden layer (128 neurons, ReLu), 1 output layer (26 neurons, SoftMax)
Optimizer	Adam
Loss Function	Categorical Crossentropy
<i>Epoch</i>	100

Activation function yang digunakan pada *neuron* yang ada di hidden layer pada top layer adalah *Rectified Linear Unit* (ReLU) (Agarap et al, 2018). Rumus *activation function* ReLu dapat dilihat pada Persamaan (1), $R(z)$ adalah output *activation function* dan z adalah nilai input.

$$R(z) = \max(0, z) \quad (1)$$

Activation function yang digunakan pada *output layer* yang ada di top layer adalah softmax (Zeiler and Fergus, 2014). Jumlah *neuron* yang ada di *output layer* sama dengan jumlah kelas SIBI yang diklasifikasikan yaitu 26. Rumus *activation function* softmax dapat dilihat pada persamaan (2), dengan input x_i sebagai input, n adalah jumlah label, dan j adalah label kelas.

$$S(x_i) = \frac{e^{x_i}}{\sum_j^n e^{x_j}} \quad (2)$$

Proses klasifikasi pada penelitian ini dilakukan dengan menggunakan perangkat keras AI workstation. Tabel 5 memberikan informasi yang tepat tentang spesifikasi komputer yang digunakan.

Training model dilakukan sebanyak 5 kali karena model evaluasi yang digunakan adalah *K-Fold Cross validation* dengan nilai $K=5$. Dataset yang digunakan untuk proses *training* adalah 80% dan 20% untuk *testing* dari dataset total. Persebaran data yang digunakan di tiap kelas dipastikan seimbang karena menggunakan metode *stratification*. Performa yang didapatkan dari setiap *cross validation* dicari rata-ratanya untuk semua algoritma dengan parameternya masing-masing.

Tabel 5. Spesifikasi komputer untuk proses *training* dan *testing*

Parameter	Spesifikasi
CPU	Intel® Xeon W-1390 (8C / 16T, 2.8 / 5.2GHz, 16MB)
RAM	32 GB
Space of Disk	1 TB
GPU Model Name	NVIDIA® RTX™ A4000 16GB GDDR6 (4x DP 1.4a)

2.4. Testing Model Klasifikasi

Pada tahapan ini, 20% dari dataset total digunakan untuk menguji tiap model klasifikasi dengan berbagai macam parameternya di tiap *cross validation*. Untuk model klasifikasi CNN dengan *transfer learning*, di tiap *epoch* saat proses *training*, langsung diujicobakan dengan dataset *testing*. Hal ini dilakukan untuk melihat performa klasifikasi di tiap *epoch*. Namun dataset testin gini tidak digunakan untuk memperbarui bobot pada arsitektur *transfer learning*. Performa yang terbaik di *epoch* tertentu yang digunakan sebagai model klasifikasi yang final.

2.5. Perhitungan Performa

Perhitungan performa dilakukan di setiap tahapan *cross validation* dan nilai performanya dirata-rata. Metric performa yang digunakan adalah *accuracy*, *precision*, *recall*, *F1 Score*, dan waktu komputasi total *training* dan *testing*. Rumus perhitungan metric performa *accuracy*, *precision*, *recall*, dan *F1 Score* untuk setiap algoritma klasifikasi dapat dilihat secara berturut turut pada persamaan (3), (4), (5), dan (6). TP adalah jumlah True Positive, TN adalah jumlah True Negative, FP adalah jumlah False Positive, dan FN adalah jumlah False Negative.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

$$F1\ Score = \frac{Precision \times Recall}{Precision+Recall} \tag{6}$$

3. HASIL DAN PEMBAHASAN

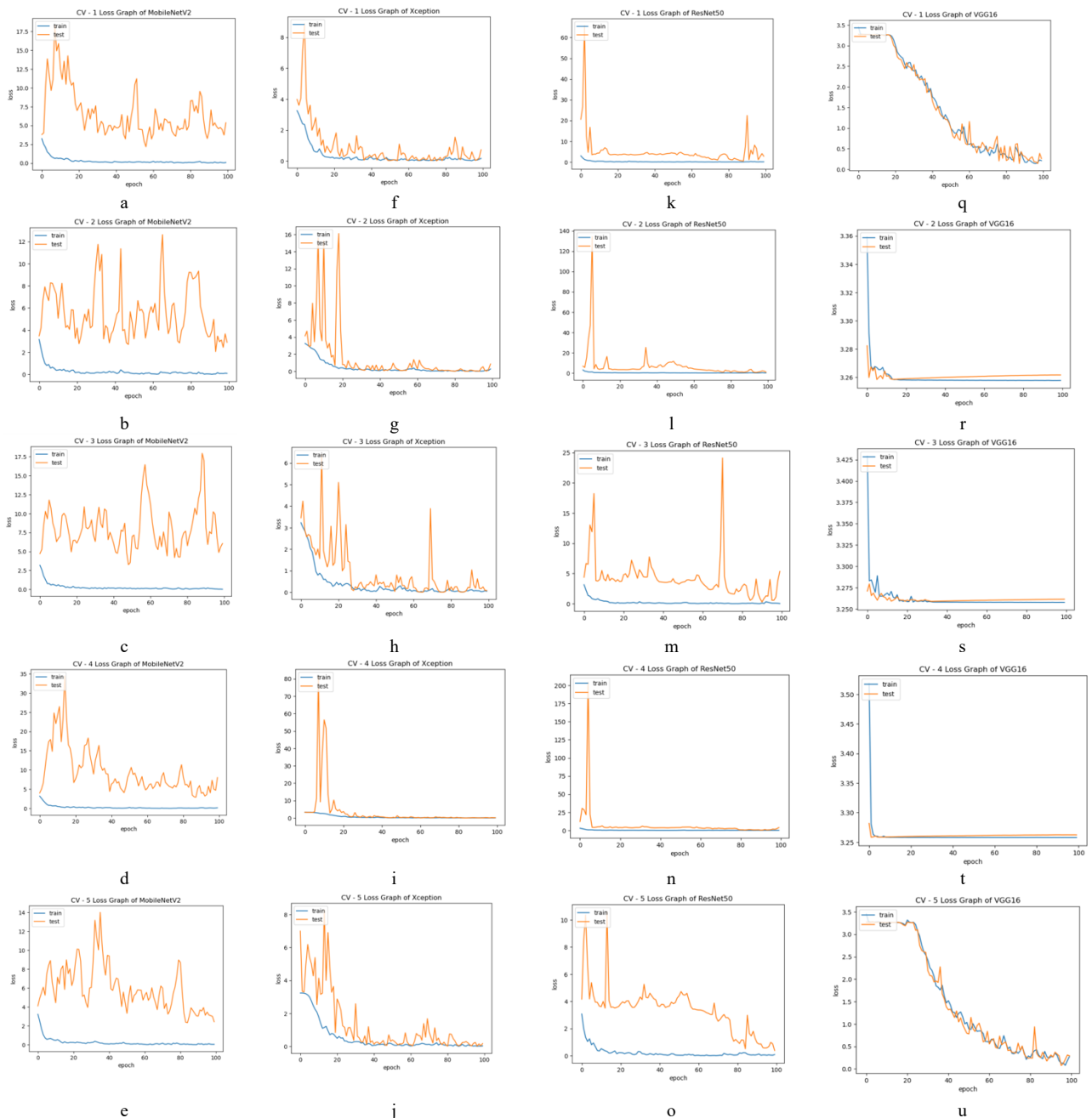
Hasil *metric* performa setiap algoritma ditampilkan pada Tabel 6. Pada penelitian ini menggunakan *5-Fold Cross validation* sehingga terdapat *metric* performa yang dihasilkan dari CV1, CV2, CV3, CV4, dan CV5. AVG adalah rata-rata performa di tiap *metric* untuk semua *Cross validation*.

Tabel 6. *Metric* Performa Algoritma Klasifikasi SIBI

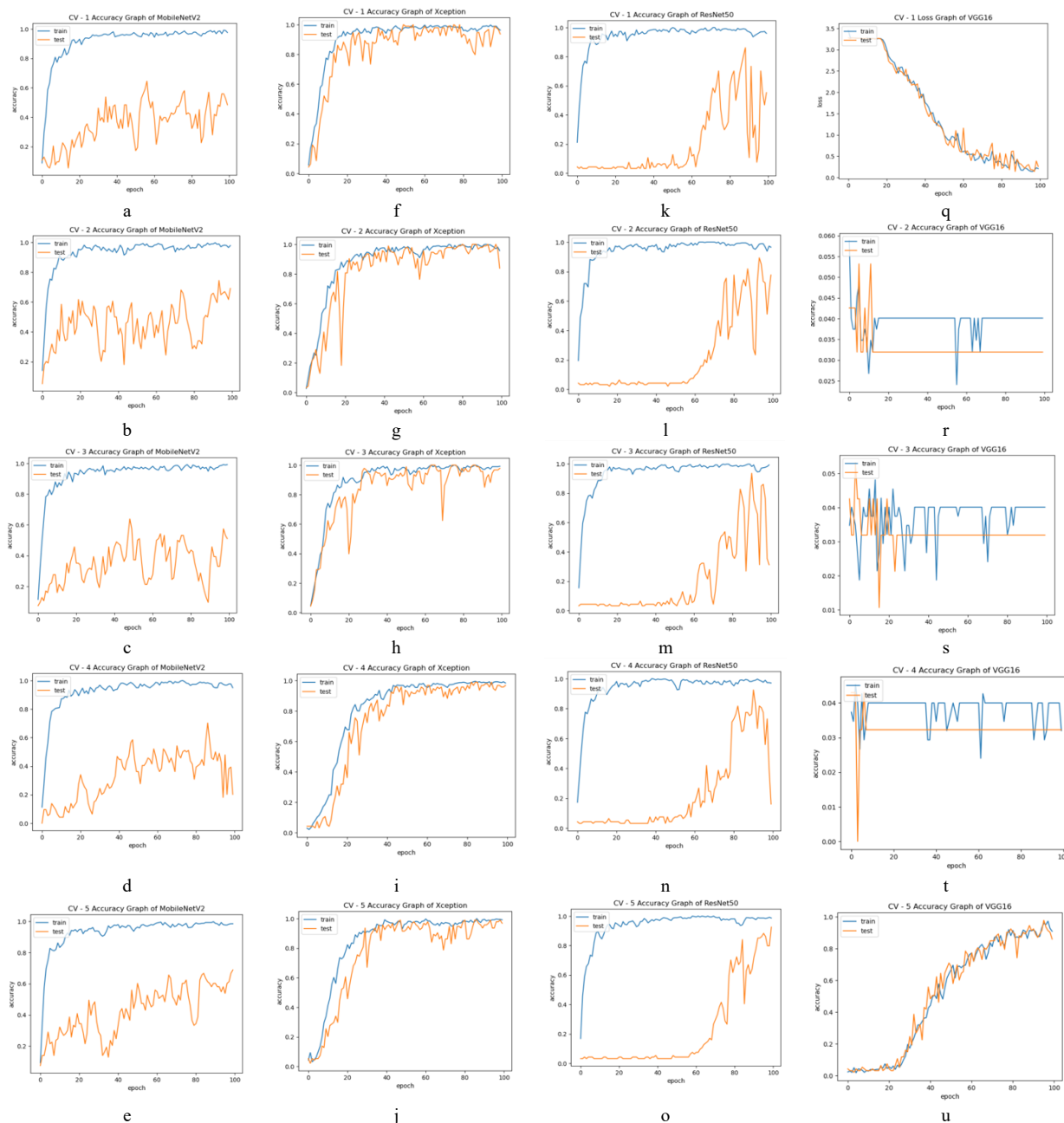
	Metrics	CV 1	CV 2	CV 3	CV 4	CV 5	AVG
KNN K = 1	Acc	84.04%	90.43%	88.30%	87.10%	87.10%	87.39%
	Prec	86.81%	92.22%	91.03%	91.29%	90.36%	90.34%
	Rec	84.04%	90.43%	88.30%	87.10%	87.10%	87.39%
	<i>F1 Score</i>	83.57%	90.19%	87.63%	86.47%	86.89%	86.95%
	Time (s)	0.03	0.03	0.03	0.03	0.03	0.03
KNN K = 2	Acc	73.40%	64.89%	72.34%	69.89%	68.82%	69.87%
	Prec	81.81%	59.83%	71.49%	68.03%	72.67%	70.76%
	Rec	73.40%	64.89%	72.34%	69.89%	68.82%	69.87%
	<i>F1 Score</i>	72.71%	59.01%	69.50%	66.18%	66.98%	66.88%
	Time (s)	0.04	0.03	0.03	0.03	0.03	0.03
KNN K = 3	Acc	68.09%	64.89%	67.02%	64.52%	63.44%	65.59%
	Prec	76.03%	60.76%	67.79%	66.43%	68.79%	67.96%
	Rec	68.09%	64.89%	67.02%	64.52%	63.44%	65.59%
	<i>F1 Score</i>	67.26%	59.56%	64.08%	61.97%	62.66%	63.11%
	Time (s)	0.04	0.04	0.03	0.03	0.03	0.03
KNN K = 4	Acc	59.57%	59.57%	64.89%	63.44%	60.22%	61.54%
	Prec	57.26%	58.90%	66.84%	65.95%	62.24%	62.24%
	Rec	59.57%	59.57%	64.89%	63.44%	60.22%	61.54%
	<i>F1 Score</i>	56.01%	55.21%	62.69%	60.76%	58.68%	58.67%
	Time (s)	0.04	0.03	0.03	0.03	0.03	0.03
KNN K = 5	Acc	57.45%	57.45%	60.64%	62.37%	59.14%	59.41%
	Prec	61.33%	57.44%	62.48%	67.31%	63.17%	62.35%
	Rec	57.45%	57.45%	60.64%	62.37%	59.14%	59.41%
	<i>F1 Score</i>	56.35%	53.27%	58.15%	60.80%	58.95%	57.51%
	Time (s)	0.03	0.03	0.03	0.03	0.04	0.03
SVM Linear	Acc	81.91%	94.68%	90.43%	89.25%	90.32%	89.32%
	Prec	83.56%	95.80%	93.01%	89.05%	91.99%	90.68%
	Rec	81.91%	94.68%	90.43%	89.25%	90.32%	89.32%
	<i>F1 Score</i>	80.93%	94.38%	89.71%	88.17%	90.12%	88.66%
	Time (s)	11.55	11.48	11.31	11.56	11.22	11.42
SVM Poly	Acc	85.11%	93.62%	90.43%	89.25%	90.32%	89.74%
	Prec	87.15%	94.95%	93.30%	89.50%	91.99%	91.38%
	Rec	85.11%	93.62%	90.43%	89.25%	90.32%	89.74%
	<i>F1 Score</i>	84.05%	93.30%	89.80%	88.33%	90.12%	89.12%
	Time (s)	11.94	11.93	11.66	11.79	11.90	11.84
SVM RBF	Acc	70.21%	74.47%	74.47%	66.67%	75.27%	72.22%
	Prec	74.22%	77.39%	79.24%	65.15%	75.81%	74.36%
	Rec	70.21%	74.47%	74.47%	66.67%	75.27%	72.22%
	<i>F1 Score</i>	69.36%	72.40%	73.24%	62.42%	73.09%	70.10%
	Time (s)	14.75	14.77	14.64	14.68	15.22	14.81
Xception	Acc	100.00%	98.94%	98.94%	100.00%	100.00%	99.57%
	Prec	100.00%	98.94%	98.94%	100.00%	100.00%	99.57%
	Rec	100.00%	98.94%	98.94%	100.00%	100.00%	99.57%
	<i>F1 Score</i>	100.00%	98.94%	98.94%	100.00%	100.00%	99.57%
	Time (s)	1359.35	1406.15	1378.35	1398.28	1393.90	1387.21
ResNet50	Acc	89.36%	86.17%	92.55%	92.47%	93.55%	90.82%
	Prec	89.36%	86.17%	92.55%	92.47%	93.55%	90.82%
	Rec	89.36%	86.17%	92.55%	92.47%	93.55%	90.82%
	<i>F1 Score</i>	89.36%	86.17%	92.55%	92.47%	93.55%	90.82%
	Time (s)	1394.64	1390.48	1426.80	1392.26	1376.35	1396.11
VGG16	Acc	5.32%	5.32%	97.87%	97.85%	4.30%	42.13%
	Prec	5.32%	5.32%	97.87%	97.85%	4.30%	42.13%
	Rec	5.32%	5.32%	97.87%	97.85%	4.30%	42.13%
	<i>F1 Score</i>	5.32%	5.32%	97.87%	97.85%	4.30%	42.13%
	Time (s)	1331.49	1306.79	1359.11	1355.14	1312.15	1332.94
MobileNetV2	Acc	70.21%	63.83%	74.47%	64.52%	68.82%	68.37%
	Prec	70.21%	63.83%	74.47%	64.52%	68.82%	68.37%
	Rec	70.21%	63.83%	74.47%	64.52%	68.82%	68.37%
	<i>F1 Score</i>	70.21%	63.83%	74.47%	64.52%	68.82%	68.37%
	Time (s)	1327.40	1320.02	1347.35	1338.73	1332.09	1333.12

Berdasarkan hasil perhitungan performa yang telah dilakukan, CNN dengan arsitektur Xception memiliki performa *F1 Score* yang terbaik yaitu 99,57%, sedangkan VGG16 memiliki performa terburuk, yaitu dengan *F1 Score* 42,13%. Waktu komputasi total proses *training* dan *testing* di algoritma *machine learning* jauh lebih cepat jika dibandingkan dengan *transfer learning*. KNN untuk semua nilai K memiliki waktu komputasi sebesar 0,03 detik. SVM membutuhkan waktu komputasi paling cepat 11,42 detik yaitu pada *kernel linear* dan paling lama 14,81 detik yaitu pada *kernel RBF*. *Transfer learning* membutuhkan waktu komputasi paling cepat 1332,94 detik yaitu pada arsitektur VGG16 dan paling lama 1396,11 detik yaitu pada arsitektur ResNet50 dengan jumlah *epoch* 100. Grafik *loss* dan *accuracy* dari *transfer learning* dapat dilihat pada Gambar 3 dan 4.

Xception memiliki performa *F1 Score* yang paling baik, namun membutuhkan waktu komputasi yang lama, sedangkan KNN dengan nilai K = 1 memiliki waktu komputasi paling cepat namun memiliki performa *F1 Score* 12,62% lebih rendah dibandingkan Xception. Perbedaan nilai *F1 Score* ini signifikan. Xception memiliki kemampuan untuk menghindari *overfitting* pada permasalahan klasifikasi alfabet SIBI. Nilai *F1 Score* Xception stabil di setiap *cross validation*. Sedangkan KNN memiliki nilai *F1 Score* yang tidak stabil di setiap *Cross validation* meskipun masih memiliki rata-rata nilai *F1 Score* yang baik yaitu 86,95%.



Gambar 3. Grafik Loss dari MobileNetV2 (a)-(e), Xception (f)-(j), ResNet50 (k)-(o), VGG16 (q)-(u)



Gambar 4. Grafik *accuracy* dari MobileNetV2 (a)-(e), Xception (f)-(j), ResNet50 (k)-(o), VGG16 (q)-(u)

Peneliti yang memiliki keterbatasan spesifikasi komputer untuk melatih model klasifikasi alfabet SIBI lebih baik menggunakan model *machine learning* KNN atau SVM, sedangkan peneliti yang membutuhkan performa yang baik dan tidak memiliki masalah dengan spesifikasi komputer, maka disarankan menggunakan Xception.

4. KESIMPULAN

Belum terdapat perbandingan performa *machine learning* dan *deep learning* pada penelitian sebelumnya. Penelitian ini memberikan komparasi yang detail antara *machine learning* dan *deep learning* untuk klasifikasi citra huruf SIBI.

Klasifikasi huruf SIBI menggunakan citra sebetulnya bukan hal yang kompleks karena citra huruf SIBI sederhana dan tidak mengandung tekstur yang rumit. Hal ini menimbulkan permasalahan penelitian yaitu seberapa efektif penggunaan algoritma *machine learning* dan *deep learning* untuk mengklasifikasikan citra alfabet SIBI. Metrics performa waktu komputasi dijadikan salah satu perhitungan dalam penelitian ini

Klasifikasi citra alfabet SIBI pada dataset kaggle menggunakan CNN dengan arsitektur Xception menunjukkan performa terbaik. *F1 Score Xception* adalah 99,57%, Namun, Xception membutuhkan waktu eksekusi lama (1396,11 detik) untuk melakukan *training* dan *testing*.

Pemilihan metode classifier tergantung kebutuhan dari peneliti. Jika membutuhkan waktu komputasi yang cepat dengan mempertimbangkan keterbatasan spesifikasi komputer yang dimiliki, maka pemilihan metode *machine learning* KNN atau SVM dapat digunakan. Hal ini dikarenakan KNN dan SVM memiliki waktu komputasi yang cepat dan memiliki performa *F1 Score* tinggi namun tidak sebaik *transfer learning*. Jika peneliti membutuhkan performa yang maksimal dan spesifikasi komputer tidak menjadi masalah, maka pemilihan *transfer learning* dengan arsitektur Xception sangat disarankan.

Penggunaan fitur selain warna RGB serta teknik ekstraksi fitur seperti *edge detection*, dan *dimensionality reduction* menggunakan *Principal Component Analysis* (PCA) dapat dipertimbangkan untuk penelitian selanjutnya. Metode tersebut dikombinasikan dengan *classifier machine learning* dan diharapkan dapat meningkatkan performa klasifikasi mendekati *transfer learning* dengan waktu komputasi yang rendah.

DAFTAR PUSTAKA

- AFKAAR, L., 2022. *Datasets SIBI Sign Language Alphabets | Kaggle*. [online] Available at: <<https://www.kaggle.com/datasets/mlanangafkkaar/datasets-lemlitbang-sibi-alphabets>> [Accessed 30 November 2022].
- AGARAP, A.F.M., 2018. Deep Learning using Rectified Linear Units (ReLU). *arXiv*, (1), pp.2–8.
- BIASA, D.P.L., 2002. *Kamus Sistem Isyarat Bahasa Indonesia*. Jakarta.
- CHOLLET, F., 2016. Xception: Deep Learning with Depthwise Separable Convolutions. *Computer Vision Foundation*. <https://doi.org/10.4271/2014-01-0975>.
- DARMATASIA, 2021. PENGENALAN SISTEM ISYARAT BAHASA INDONESIA (SIBI) MENGGUNAKAN GRADIENT-CONVOLUTIONAL NEURAL NETWORK. *Jurnal Instek*, 6(1), pp.56–65.
- DOKMANIC, I., PARHIZKAR, R., RANIERI, J. AND VETTERLI, M., 2015. Euclidean Distance Matrices: Essential Theory, Algorithms and Applications. *IEEE Signal Processing Magazine*, [online] 32(6), pp.12–30. <https://doi.org/10.1109/MSP.2015.2398954>.
- ERDEFI, R. AND NOER, F.P.S., 2022. Improving Recognition of SIBI Gesture by Combining Skeleton and Hand Shape Features. *Journal of Computer Science and Information*, 2, pp.69–79. <https://doi.org/10.14341/conf05-08.09.22-191>.
- LECUN, Y., BOTTOU, L., BENGIO, Y. AND HA, P., 1998. LeNet. *Proceedings of the IEEE*, (November), pp.1–46.
- LEGOWO, R.S., SUMPENO, S. AND PRAMUNANTO, E., 2017. Klasifikasi gerakan tangan SIBI (sistem isyarat bahasa indonesia) menggunakan leap motion dengan metode klasifikasi naive bayes. [online] Available at: <https://repository.its.ac.id/46150/%0Ahttps://repository.its.ac.id/46150/1/2913100028-Undergraduate_Thesis.pdf>.
- SANDLER, M., HOWARD, A., ZHU, M., ZHMOGINOV, A. AND CHEN, L.-C., 2018. MobileNetV2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, [online] pp.4510–4520. Available at: <<http://arxiv.org/abs/1801.04381>> [Accessed 27 May 2021].
- SHOLAWATI, M., AULIASARI, K. AND ARIWIBISONO, F., 2022. Pengembangan Aplikasi Pengenalan Bahasa Isyarat Abjad Sibi Menggunakan Metode Convolutional Neural Network (Cnn). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(1), pp.134–144. <https://doi.org/10.36040/jati.v6i1.4507>.
- SIMONYAN, K. AND ZISSERMAN, A., 2015. Very deep convolutional networks for large-scale image recognition. In: *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. [online] International Conference on Learning Representations, ICLR. Available at: <<http://www.robots.ox.ac.uk/>> [Accessed 27 May 2021].
- SINDARTO, S.S., RATNAWATI, D.E. AND ARWANI, I., 2022. Klasifikasi Citra Sistem Isyarat Bahasa Indonesia (SIBI) dengan Metode Convolutional Neural Network pada Perangkat Lunak berbasis Android. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, [online] 6(5), pp.2129–2138. Available at: <<http://j-ptiik.ub.ac.id>>.
- SUHARJITO, THIRACITTA, N. AND GUNAWAN, H., 2021. SIBI Sign Language Recognition Using Convolutional Neural Network Combined with Transfer Learning and non-trainable Parameters. *Procedia Computer Science*, [online] 179(2019), pp.72–80. <https://doi.org/10.1016/j.procs.2020.12.011>.
- VICTOR IKECHUKWU, A., MURALI, S., DEEPU, R. AND SHIVAMURTHY, R.C., 2021. ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images. *Global Transitions Proceedings*, 2(2), pp.375–381. <https://doi.org/10.1016/J.GLTP.2021.08.027>.
- YUSNITA, L., ROESTAM, R. AND WAHYU, R.B., 2017. Implementation of Real-Time Static Hand. *CommIT (Communication & Information Technology)*, 11(2), pp.85–91.

ZEILER, M.D. AND FERGUS, R., 2014. Visualizing and understanding convolutional networks. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8689 LNCS(PART 1), pp.818–833. https://doi.org/10.1007/978-3-319-10590-1_53.

Halaman ini sengaja dikosongkan