

## Comparison of Transfer Learning Algorithm Performance in Hand Sign Language Digits Image Classification

### Perbandingan Kinerja Algoritma *Transfer Learning* dalam Klasifikasi Citra *Hand Sign Language Digits*

A. Muh. Amil Siddik\*

\* Program Studi Sistem Informasi, Universitas Hasanuddin, Indonesia

Email: [amilsiddik@unhas.ac.id](mailto:amilsiddik@unhas.ac.id)

#### Abstract

Image hand sign classification has become an interesting topic in image processing and machine learning. However, to achieve optimal performance in hand sign image classification tasks, a large and diverse dataset as well as powerful learning algorithms are required. One popular technique for improving the performance of classification models is transfer learning, which allows the use of knowledge learned from previous models and applies it to new tasks. In this study, the performance of two different transfer learning algorithms, ResNet-50 and VGG-16, was compared on the Sign Language Digits Dataset, which consists of 10 different types of handwriting images. The results of the experiment showed that both tested transfer learning algorithms had good performance. However, VGG-16 provided the best results with an accuracy of 97,29%, precision of 97,38%, recall of 97,45%, and an F1 score of 97,36%, while ResNet-50 achieved an accuracy of 94,57%, precision of 94,75%, recall of 94,96%, and an F1 score of 94,78%. In conclusion, transfer learning algorithms are effective techniques for improving the performance of hand sign image classification models. Choosing the appropriate transfer learning algorithm and dataset can help generate more accurate classification models.

**Keywords:** Transfer Learning, VGG-16, ResNet-50.

#### Abstrak

Klasifikasi citra *hand sign* telah menjadi topik yang menarik dalam pengolahan citra dan pembelajaran mesin. Namun, untuk mencapai kinerja yang optimal dalam tugas klasifikasi citra *hand sign*, diperlukan dataset yang besar dan beragam, serta algoritma pembelajaran yang kuat. Salah satu teknik yang populer untuk meningkatkan kinerja model klasifikasi adalah *transfer learning*, yang memungkinkan untuk menggunakan pengetahuan yang telah dipelajari dari model sebelumnya dan diterapkan pada tugas baru. Pada penelitian ini, dibandingkan kinerja dua algoritma *transfer learning* yang berbeda, yaitu ResNet-50 dan VGG-16 pada klasifikasi citra dataset *Sign Language Digits Dataset* yang terdiri dari 10 jenis



citra handwriting yang berbeda. Hasil dari pengujian menunjukkan bahwa kedua algoritma *transfer learning* yang diuji memiliki kinerja yang baik. Namun, VGG-16 memberikan hasil terbaik dengan akurasi sebesar 97,29%, presisi sebesar 97,38% recall sebesar 97,45% dan F1 *score* sebesar 97,36%, sedangkan pada ResNet-50 memberikan akurasi sebesar 94,57%, presisi sebesar 94,75%, *recall* 94,96% dan F1 *score* sebesar 94,78%. Sebagai kesimpulan, algoritma *transfer learning* merupakan teknik yang efektif dalam meningkatkan kinerja model klasifikasi citra *hand sign*. Pemilihan algoritma *transfer learning* yang tepat dan dataset yang sesuai dapat membantu menghasilkan model klasifikasi yang lebih akurat.

**Kata kunci:** *Transfer Learning*, VGG-16, ResNet-50.

## 1. PENDAHULUAN

Klasifikasi citra adalah salah satu tugas penting dalam pengolahan citra dan memiliki banyak aplikasi dalam berbagai bidang, seperti pengenalan wajah, deteksi objek, dan pengenalan tulisan tangan. Saat ini, terdapat banyak algoritma yang telah dikembangkan untuk tugas klasifikasi citra, termasuk Convolutional Neural Network (CNN). ResNet-50 dan VGG-16 adalah dua arsitektur CNN yang sangat populer dan telah banyak digunakan dalam berbagai aplikasi klasifikasi citra [9,12].

Penelitian terdahulu oleh [2,5,14,15] menunjukkan bahwa ResNet-50 memiliki performa yang lebih baik daripada VGG-16 dalam klasifikasi citra. Para peneliti tersebut menggunakan dataset yang berbeda, namun menghasilkan kesimpulan yang sama yaitu ResNet-50 memiliki performa yang lebih baik jika dibandingkan dengan VGG-16.

Penelitian sebelumnya oleh [11] menunjukkan bahwa VGG-16 memiliki akurasi yang lebih tinggi daripada ResNet-50 pada klasifikasi citra untuk dataset ImageNet menggunakan teknik transfer learning. Begitu pula, penelitian yang dilakukan oleh [8] yang membandingkan beberapa arsitektur jaringan VGG-16 dan ResNet-50, dalam klasifikasi citra *Diabetes Retinopathy*. Hasil penelitian tersebut menunjukkan bahwa VGG-16 memiliki akurasi yang lebih baik dibandingkan ResNet50 dalam klasifikasi citra *Diabetes Retinopathy*.

Dari penjelasan di atas, terlihat bahwa masih belum ada konsensus tentang algoritma klasifikasi citra terbaik antara ResNet-50 dan VGG-16. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja ResNet-50 dan VGG-16 dalam klasifikasi citra pada dataset yang berbeda, dengan harapan dapat memberikan rekomendasi tentang algoritma yang lebih baik pada kasus klasifikasi citra.

## 2. DATASET DAN METODE PENELITIAN

### 2.1 Deskripsi Dataset

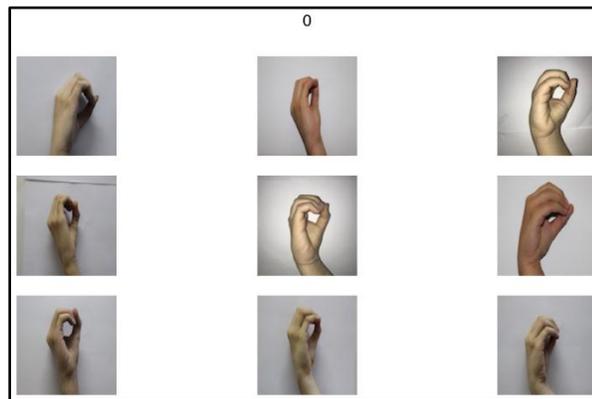
Dataset dalam penelitian ini merupakan citra *Sign Language Digits Dataset* yang berisi gambar tangan manusia yang membentuk angka 0 sampai 9 [6]. Dataset ini terdiri atas 2062 citra. Masing-masing citra pada dataset ini berukuran 100×100 pixel. Distribusi dataset setiap kelas dapat dilihat pada Tabel 2. 1

**Tabel 2. 1.** Distribusi Kelas Dataset

Label	Jumlah
0	205
1	206
2	206
3	206

4	207
5	207
6	207
7	206
8	208
9	204

Sebagai contoh, untuk citra yang berlabel 0 dapat dilihat pada Gambar 2. 1. Dari gambar diperlihatkan 9 buah citra *Sign Language Digits Dataset* yang membentuk angka 0.



Gambar 2. 1 Contoh Citra dengan Label 0

## 2.2 Metode Penelitian

Convolutional Neural Network (CNN) adalah sebuah tipe arsitektur neural network yang banyak digunakan untuk tugas pengolahan citra, seperti klasifikasi citra dan deteksi objek. CNN menggunakan operasi konvolusi untuk mengekstrak fitur dari citra secara otomatis, sehingga memungkinkan jaringan untuk belajar fitur-fitur yang penting secara hierarkis [3, 13].

Konvolusi pada CNN adalah operasi matematis yang melibatkan filter atau kernel yang digeser secara bertahap pada citra input. Setiap pergeseran akan menghasilkan nilai konvolusi antara filter dan citra di posisi tersebut. Hasil konvolusi ini kemudian disimpan dalam sebuah map fitur, yang akan digunakan sebagai input untuk *layer* selanjutnya [3]. Adapun rumus operasi konvolusi dapat dilihat pada Persamaan (2.1) berikut:

$$FM[i]_{j,k} = \left( \sum_m \sum_n N_{[j-m,k-n]} F_{[m,n]} \right) + bF \quad (2.1)$$

dimana

$FM[i]_{j,k}$  : Matriks map fitur ke- $i$  dengan posisi piksel  $(j, k)$

$N_{[j-m,k-n]}$  : Matriks citra masukan dengan posisi piksel  $((j - m), (k - n))$

$F_{[m,n]}$  : Matriks filter konvolusi dengan posisi piksel  $(m, n)$

$bF$  : Nilai bias filter.

CNN biasanya terdiri dari beberapa *layer*, seperti *layer* konvolusi, *pooling layer*, dan *fully connected layer*. *Layer* konvolusi bertanggung jawab untuk mengekstrak fitur dari citra, sementara *pooling layer* digunakan untuk mengurangi dimensi dari map fitur dengan memilih nilai maksimum atau rata-rata dalam suatu wilayah tertentu. Output dari lapisan sebelumnya dihubungkan ke output akhir melalui *fully connected layer*. *Fully connected layer* terdiri dari *input layer*, *hidden layer* dan *output layer*. Seluruh matriks map fitur yang diperoleh dari proses *pooling layer* digabungkan ke dalam *input layer*, kemudian setiap piksel direntangkan menjadi sebuah vektor sepanjang jumlah piksel matriks yang diperoleh pada *pooling layer*. Selanjutnya, nilai-nilai dari *input layer* digunakan untuk menghitung *hidden layer*. Seluruh nilai yang

## JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

A. Muh. Amil Siddik

dimasukkan ke *layer* ini dikalikan dengan bobot yang telah diinisialisasi sebelumnya, dan kemudian ditambahkan dengan nilai bias. Rumus yang digunakan untuk menghitung nilai masukan untuk *node hidden layer* dapat dilihat pada Persamaan (2.2) berikut:

$$Z_{in_i} = \sum_{j=1}^n X_j * V_{j,i} + V_{0,i} \quad (2.2)$$

dimana:

$Z_{in_i}$  : masukan untuk *node hidden layer* ke-i dengan jumlah node  $n$

$X_j$  : node  $X$  ke-j

$V_{j,i}$  : bobot  $V$  untuk  $X_j$  dan node  $Z_j$

$V_{0,i}$  : bias  $V$  untuk  $Z_{in_i}$ .

Setelah perhitungan selesai, semua hasil perhitungan dimasukkan ke dalam fungsi aktivasi *Rectified Linear Unit* (ReLU). Sehingga dihasilkan nilai keluaran  $Z$ . Proses perhitungan pada *output layer* akan menggunakan hasil perhitungan ini. Perhitungan pada *output layer* terdiri dari perkalian nilai-nilai hasil perhitungan pada *hidden layer* dengan bobot yang telah diinisialisasi sebelumnya, kemudian ditambahkan dengan nilai bias. Untuk rumus perhitungannya adalah sebagai berikut:

$$y_{in_i} = \sum_{j=1}^n Z_j * W_{j,i} + W_{0,i} \quad (2.3)$$

dimana:

$y_{in_i}$  : masukan untuk *node hidden layer* ke-i dengan jumlah *node*  $n$

$Z_j$  : node  $Z$  ke-j

$W_{j,i}$  : bobot  $W$  untuk  $X_j$  dan node  $Z_j$

$W_{0,i}$  : bias  $W$  untuk  $y_{in_i}$ .

Setelah dilakukan perhitungan, selanjutnya hasil perhitungan dimasukkan ke dalam fungsi aktivasi *Softmax*, maka akan diperoleh nilai keluaran  $y$ . Adapun fungsi aktivasi

$$y_i = \frac{e^{y_{in_i}}}{\sum_{i=1}^m e^M} \quad (2.4)$$

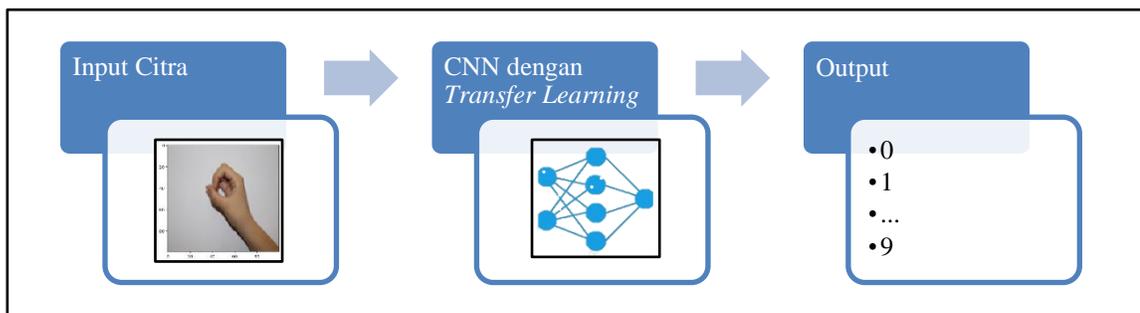
dimana:

$y_{out_i}$  : keluaran untuk *node hidden layer* ke-i dengan jumlah *node*  $n$

$y_{in_i}$  : masukan untuk *node hidden layer* ke-i dengan jumlah *node*  $n$

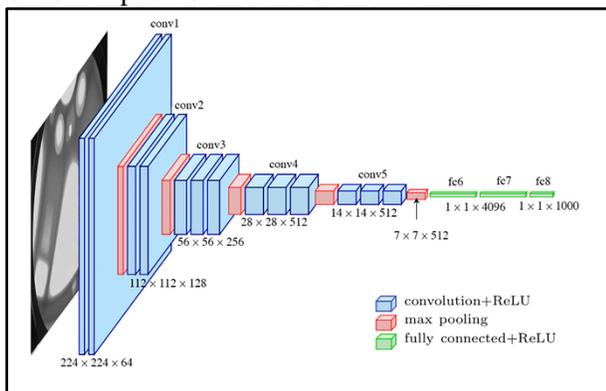
$M$  : semua masukan untuk lapisan *output* sejumlah  $M$  buah.

Pada penelitian ini diterapkan penambahan *transfer learning*, yang berarti model arsitektur dalam CNN diubah dengan model *pre-trained* dengan pelatihan pada dataset lain sebelumnya. Terdapat banyak model *transfer learning* yang telah diciptakan, tetapi dari beberapa penelitian sebelumnya ada dua metode yang sering digunakan untuk klasifikasi citra yaitu VGG-16 dan ResNet-50 [2,5,8,4]. Oleh karena itu dalam penelitian ini, akan dibandingkan kinerja antara VGG-16 dan ResNet-50. Arsitektur yang disajikan dalam penelitian ini dapat dilihat pada Gambar 2. 2



**Gambar 2. 2** Arsitektur CNN dengan Transfer Learning

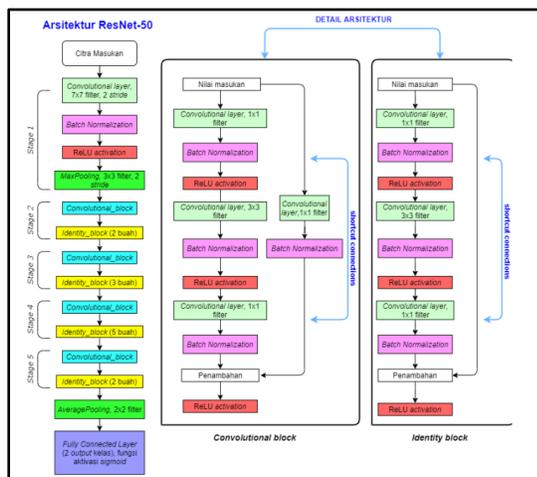
VGG-16 adalah versi *convolutional neural network* populer yang disebut VGG-Net. VGG-16 terdiri dari beberapa lapisan, termasuk 13 lapisan konvolusional dan 3 *fully connected layer*. Karena itu model ini harus mempelajari bobot 16 layer [1,11]. Model ini butuh input gambar berwarna berukuran  $224 \times 224$  piksel dan mengklasifikasikannya dalam salah satu dari 1000 kelas. Arsitektur VGG-16 diilustrasikan pada Gambar 2. 3 di bawah ini.



**Gambar 2. 3** Arsitektur VGG-16 [11].

ResNet-50 atau Residual Network adalah model yang dikembangkan oleh Microsoft dan memenangkan kompetisi tahunan ILSVRC (ImageNet Large Scale Recognition Visual Recognition Competition) tahun 2015 [5,7,10]. ILSVRC adalah kompetisi tahunan di mana beberapa tim bersaing untuk menghasilkan algoritma terbaik untuk tugas-tugas computer vision. ResNet-50 terdiri dari 50 *layer* dan lebih dari 25,6 juta parameter, arsitektur ini juga terdiri atas kombinasi dari konvolusi, blok identitas (input = output), dan *fully connected layer*.

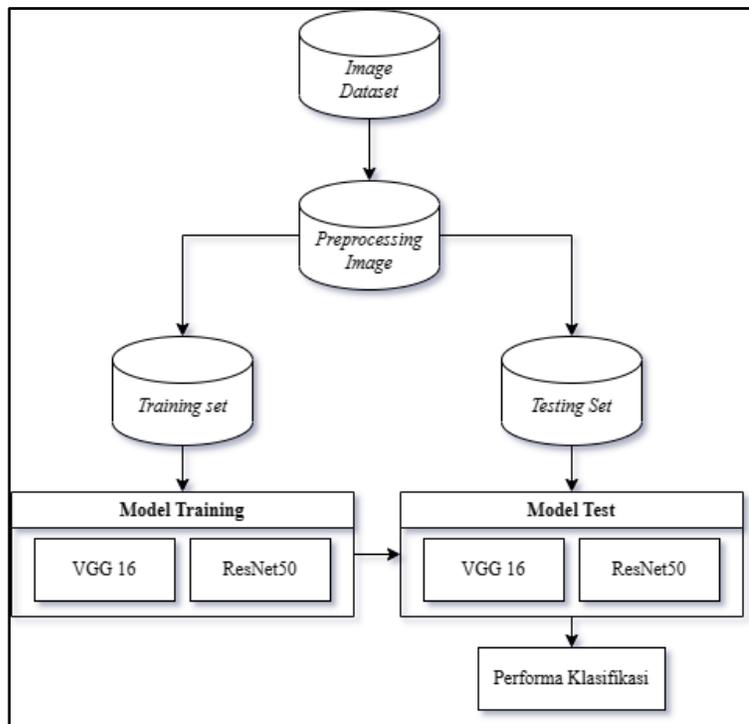
Dalam arsitektur ResNet-50, diperkenalkan konsep baru dengan istilah *shortcut connections*. Konsep ini muncul sebagai solusi dari masalah *vanishing gradient* yang muncul saat mencoba memperdalam struktur *network*, mengingat jumlah *layer* dari ResNet-50 yang sangat banyak. Dalam konsep *shortcut connections*, fitur yang merupakan input dari *layer* sebelumnya juga dijadikan sebagai input terhadap output dari *layer* tersebut. Cara ini dilakukan untuk meminimalisir hilangnya fitur-fitur penting pada saat proses konvolusi. Secara keseluruhan ResNet-50 terdiri dari 5 stage proses konvolusi yang kemudian dilanjutkan average pooling dan diakhiri dengan fully connected layer sebagai layer prediksi. Gambar 2. 4 di bawah mengilustrasikan arsitektur ResNet-50.



Gambar 2. 4 Arsitektur ResNet-50 [2].

Input awal dari penelitian ini adalah dataset citra yang ukurannya di-*resize* dan dimasukkan ke dalam model yang dirancang. Kemudian masuk ke *layer* CNN dengan *transfer learning*, pada bagian ini digunakan model VGG-16 atau ResNet-50. Setelah itu citra akan diklasifikasikan ke salah satu dari 10 angka pada output.

Pada penelitian ini, *dataset* citra yang digunakan dibagi menjadi data *training* dan *testing*, kemudian hasil dari *training* dan *testing* model digunakan untuk mengetahui performa dari klasifikasi citra. Gambar 2. 5 memperlihatkan alur atau tahapan pada penelitian ini.



Gambar 2. 5 Alur Penelitian

### 3. HASIL DAN PEMBAHASAN

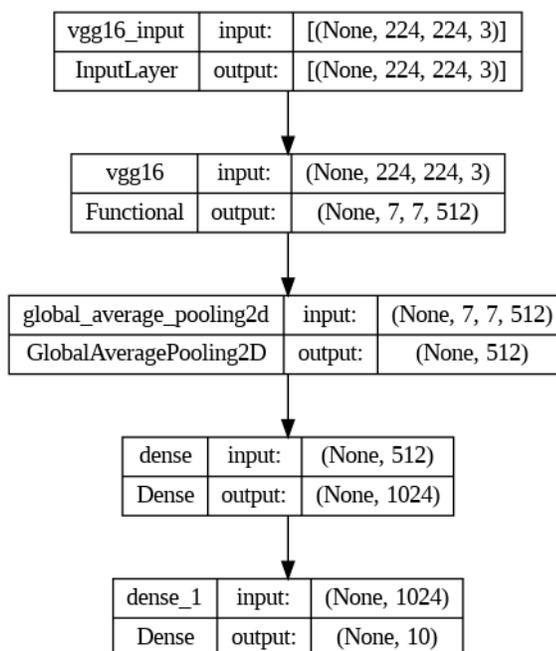
Pada bagian ini akan dibahas tentang hasil yang didapatkan menggunakan *transfer learning* VGG-16 dan ResNet-50. Sebelum dimasukkan ke dalam model, data citra *Sign Language Digits Dataset* di-*resize* ke ukuran  $224 \times 224$  px kemudian dibagi menjadi data *training* dan *testing* terlebih dahulu. Perbandingan data *training* dan *testing* yaitu 75% dari total data untuk data *training* dan 25% dari total data untuk data *testing*.

#### 3.1. Klasifikasi dengan VGG-16

Ringkasan model untuk klasifikasi citra dengan VGG-16 pada penelitian ini diilustrasikan pada Tabel 3. 1. dan Gambar 3. 1.

**Tabel 3. 1** Ringkasan Model VGG-16

Model: "sequential"		
Layer (type)	Output Shape	Param#
vgg16 (Functional)	(None, 7, 7, 512)	14714688
global_average_pooling2d (GlobalAveragePooling2D)	(None, 512)	0
dense (Dense)	(None, 1024)	525312
dense_1 (Dense)	(None, 10)	10250
Total params: 15,250,250		
Trainable params: 535,562		
Non-trainable params: 14,714,688		



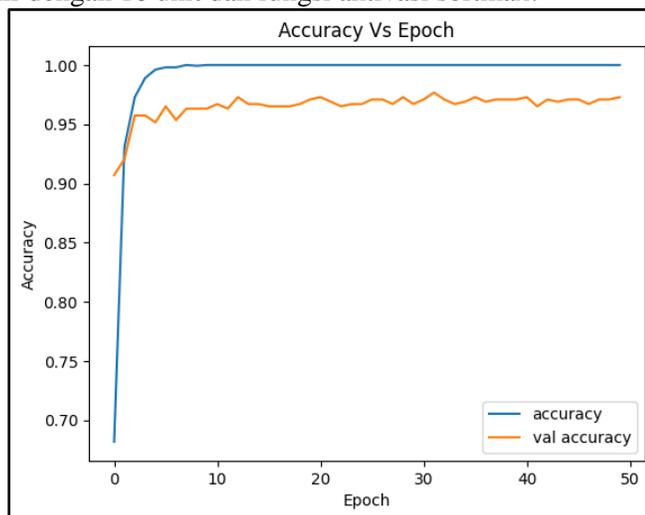
**Gambar 3. 1** Plot Model VGG-16

Dari Tabel 3. 1 dan Gambar 3. 1 dapat dilihat arsitektur dari *transfer learning* dengan VGG-16. Model CNN dibangun dengan menambahkan lapisan yang lebih dalam ke model dasar VGG16. Lapisan yang ditambahkan adalah GlobalAveragePooling2D yang bertindak sebagai

## JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

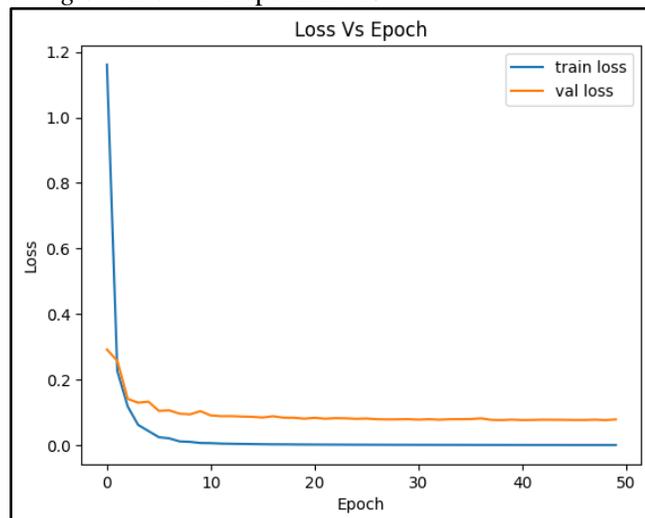
A. Muh. Amil Siddik

pengumpul fitur global, lapisan Dense dengan 1024 unit dengan fungsi aktivasi ReLU, dan lapisan Dense terakhir dengan 10 unit dan fungsi aktivasi softmax.



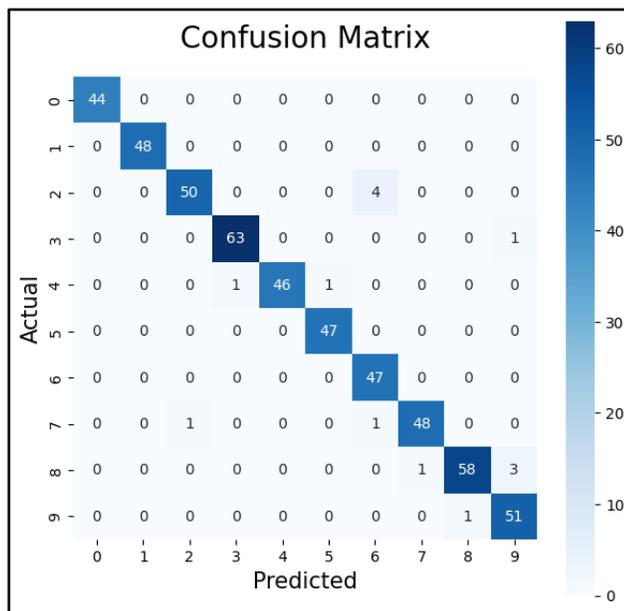
**Gambar 3. 2** *Training* dan validasi akurasi dengan model VGG-16

*Training* dan Validasi akurasi untuk VGG-16 dengan 50 epoch dapat dilihat pada Gambar 3. 2. Dari gambar terlihat bahwa akurasi validasi hampir stabil untuk setiap epoch pada klasifikasi, sedangkan akurasi *training* stabil setelah epoch ke 10.



**Gambar 3. 3** *Training* dan validasi *loss* dengan model VGG-16

Dari Gambar 3. 3 terlihat bahwa fungsi *loss* menurun secara signifikan selama fase validasi dan *training*.



Gambar 3. 4 Confusion Matrix VGG-16

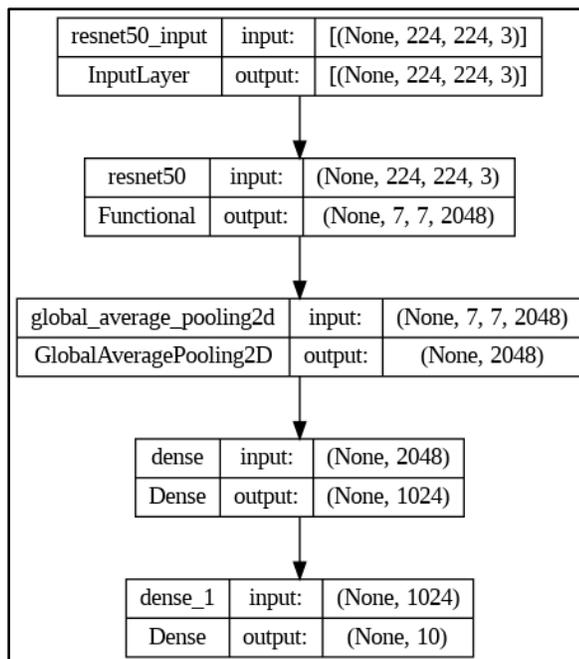
*Confusion matrix* klasifikasi citra *Sign Language Digits Dataset* menggunakan model VGG-16 diilustrasikan pada Gambar 3. 4. Gambar 3. 8, matriks ini digunakan untuk menggambarkan kinerja penuh model ResNet-50. Dari matrix terlihat bahwa sebagian besar citra telah terprediksi dengan tepat.

### 3.2. Klasifikasi dengan ResNet-50

Ringkasan model untuk klasifikasi citra dengan ResNet-50 yang digunakan dapat dilihat pada Tabel 3. 2.

Tabel 3. 2 Ringkasan Model ResNet-50

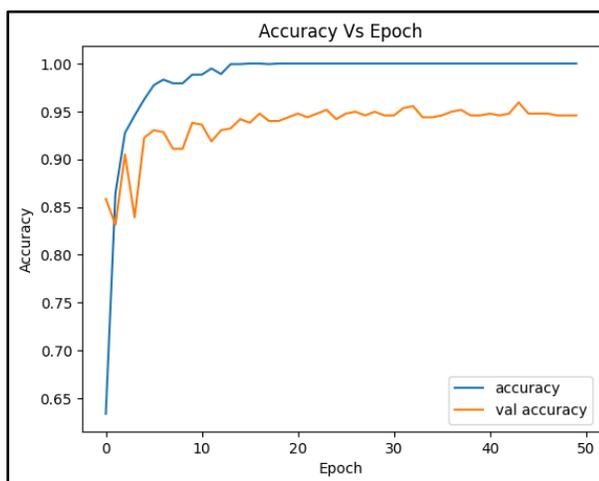
Model: "sequential"		
Layer (type)	Output Shape	Param#
resnet50 (Functional)	(None, 7, 7, 2048)	23587712
global_average_pooling2d (GlobalAveragePooling2D)	(None, 2048)	0
dense (Dense)	(None, 1024)	2098176
dense_1 (Dense)	(None, 10)	10250
Total params: 25,696,138		
Trainable params: 2,108,426		
Non-trainable params: 23,587,712		



**Gambar 3. 5** Plot Model ResNet-50

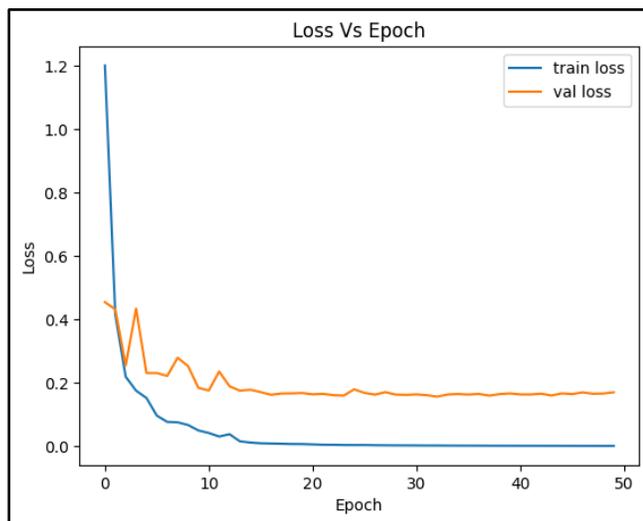
Seperti model VGG-16, dari Tabel 3. 2 dan Gambar 3. 5 dapat dilihat arsitektur dari *transfer learning* dengan ResNet-50. Model CNN dibangun dengan menambahkan lapisan yang lebih dalam ke model dasar ResNet-50. Lapisan yang ditambahkan adalah GlobalAveragePooling2D yang bertindak sebagai pengumpul fitur global, lapisan Dense dengan 1024 unit dengan fungsi aktivasi ReLU, dan lapisan Dense terakhir dengan 10 unit dan fungsi aktivasi softmax.

*Training* dan Validasi akurasi untuk ResNet-50 dengan 50 epoch dapat dilihat pada Gambar 3. 6. Dari gambar terlihat bahwa akurasi validasi hampir stabil untuk setiap epoch pada klasifikasi, sedangkan akurasi *traing* stabil setelah epoch ke 20.



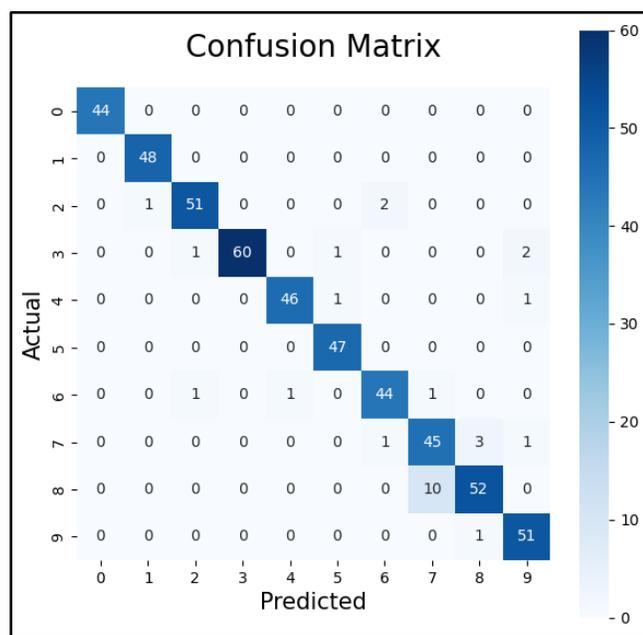
**Gambar 3. 6** *Training* dan validasi akurasi dengan model ResNet-50

Dari Gambar 3. 7 terlihat bahwa fungsi *loss* menurun secara signifikan selama fase validasi dan *training*.



**Gambar 3. 7** Training dan validasi *loss* dengan model ResNet-50

*Confusion matrix* klasifikasi citra *Sign Language Digits Dataset* menggunakan model ResNet-50 diilustrasikan pada Gambar 3. 8 untuk menggambarkan kinerja penuh model ResNet-50. Dari matriks terlihat bahwa sebagian besar citra telah terprediksi dengan tepat.



**Gambar 3. 8** Confusion Matrix ResNet-50

### 3.4. Perbandingan Performa VGG-16 dan ResNet-50

Dalam membandingkan algoritma dalam *transfer learning* biasanya, ada beberapa metrik yang umumnya digunakan.

1. Akurasi mengukur seberapa baik model klasifikasi dapat mengklasifikasikan data dengan benar secara keseluruhan. Akurasi dinyatakan dalam persentase dan dihitung dengan membagi jumlah prediksi yang benar dengan total jumlah sampel.

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN}, \quad (3.1)$$

- a) *TP* (*True Positive*) adalah kasus ketika model secara benar memprediksi sampel sebagai positif (klasifikasi benar), dan label sejati juga positif.
  - b) *TN* (*True Negative*) adalah kasus ketika model secara benar memprediksi sampel sebagai negatif (klasifikasi benar), dan label sejati juga negatif.
  - c) *FP* (*False Positive*) adalah kasus ketika model secara salah memprediksi sampel sebagai positif (klasifikasi salah), tetapi label sejati sebenarnya negatif.
  - d) *FN* (*False Negative*): Ini adalah kasus ketika model secara salah memprediksi sampel sebagai negatif (klasifikasi salah), tetapi label sejati sebenarnya positif.
2. Presisi menggambarkan sejauh mana hasil positif yang diprediksi adalah benar.

$$\text{Presisi} = \frac{TP}{TP+FP}, \quad (3.2)$$

3. *Recall* menggambarkan sejauh mana hasil positif yang sebenarnya ditemukan oleh algoritma.

$$\text{Recall} = \frac{TP}{TP+FN}, \quad (3.3)$$

4. *F1 Score* menggabungkan presisi dan recall dalam satu metrik tunggal. Ini adalah rata-rata harmonik antara presisi dan recall, yang memberikan perhatian seimbang pada kedua metrik tersebut.

$$\text{F1 Score} = \frac{2 \times \text{Recall} \times \text{Presisi}}{\text{Recall} + \text{Presisi}}. \quad (3.4)$$

Tabel 3. 3 di bawah merepresentasikan performa dari kedua model yang digunakan pada penelitian ini.

**Tabel 3. 3** Metrik Perbandingan VGG-16 dan ResNet-50

Algoritma	Akurasi	Presisi	Recall	F1 Score
VGG-16	0,9729	0,9738	0,9745	0,9736
ResNet-50	0,9457	0,9475	0,9496	0,9478

Berikut adalah pembahasan untuk masing-masing metrik:

1. Akurasi :
  - a. Model VGG-16 mencapai akurasi sebesar 0,9729, sedangkan model ResNet-50 memiliki akurasi 0,9457.
  - b. Dalam hal akurasi, performa yang ditunjukkan model VGG-16 lebih baik daripada model ResNet-50. Model VGG-16 mampu mengklasifikasikan dengan benar sekitar 97,29% dari seluruh sampel.
2. Presisi :
  - a. Model VGG-16 memiliki presisi sebesar 0,9738, sedangkan model ResNet-50 memiliki presisi sebesar 0,9475.
  - b. Dalam hal presisi, performa yang ditunjukkan model VGG-16 juga lebih baik. Presisi menggambarkan sejauh mana hasil positif yang diprediksi benar, dan model VGG-16 mampu mencapai presisi sekitar 97,38% untuk kelas-kelas yang diprediksi positif.

3. *Recall*:
  - a. Model VGG-16 memiliki *recall* sebesar 0,9745, sedangkan model ResNet-50 memiliki *recall* sebesar 0,9496.
  - b. Dalam hal *recall*, model VGG-16 sekali lagi menunjukkan performa yang lebih baik. *Recall* menggambarkan sejauh mana hasil positif yang sebenarnya ditemukan oleh model, dan model VGG-16 mampu mencapai *recall* sekitar 97,45% untuk kelas-kelas yang sebenarnya positif.
4. *F1 Score*:
  - a. Model VGG-16 memiliki *F1 score* sebesar 0,9736, sedangkan model ResNet-50 memiliki *F1 score* sebesar 0,9478.
  - b. *F1 score* merupakan rata-rata harmonik antara presisi dan *recall*, dan memberikan perhatian seimbang pada kedua metrik tersebut. Dalam hal *F1 score*, model VGG-16 sekali lagi menunjukkan performa yang lebih baik, dengan *F1 score* sekitar 97,36%.

Berdasarkan pembahasan di atas, secara keseluruhan, model VGG-16 menunjukkan performa yang lebih baik daripada model ResNet-50 dalam klasifikasi citra *Hand Sign Language Digits*, diukur berdasarkan akurasi, presisi, *recall*, dan *F1 score*. Namun, penting untuk dicatat bahwa evaluasi model harus selalu dipertimbangkan dalam konteks dan tujuan aplikasi yang spesifik, serta dengan mempertimbangkan faktor-faktor lain seperti kebutuhan komputasi, kecepatan inferensi, dan ukuran model.

#### 4. Kesimpulan

VGG-16 dan ResNet-50 adalah dua arsitektur jaringan saraf konvolusi yang banyak digunakan dalam aplikasi klasifikasi gambar. Kedua algoritma ini memiliki akurasi dan presisi yang tinggi, namun pada penelitian ini, algoritma VGG-16 mengungguli ResNet-50 dalam hal performa. VGG-16 mencapai akurasi, presisi, *recall*, dan *F1 score* yang lebih tinggi daripada ResNet-50.

Dengan akurasi sebesar 0,9729, VGG-16 memiliki kemampuan yang lebih baik dalam mengklasifikasikan sampel dengan benar dibandingkan dengan ResNet-50 yang memiliki akurasi sebesar 0,9457. Dalam hal presisi, VGG-16 mencapai nilai 0,9738, sementara ResNet-50 mencapai nilai 0,9475. Hal ini menunjukkan bahwa VGG-16 memiliki kecenderungan yang lebih rendah dalam memberikan prediksi yang salah positif dibandingkan dengan ResNet-50. VGG-16 juga menunjukkan *recall* yang lebih baik dengan nilai 0,9745, dibandingkan dengan ResNet-50 yang memiliki *recall* sebesar 0,9496. Ini menunjukkan bahwa VGG-16 lebih efektif dalam menemukan kembali sampel positif yang sebenarnya. *F1 score* yang dihasilkan oleh VGG-16 (0,9736) juga lebih tinggi daripada ResNet-50 (0,9478), menunjukkan bahwa VGG-16 mampu mencapai keseimbangan yang lebih baik antara presisi dan *recall*.

Oleh karena itu dapat disimpulkan bahwa dalam kasus klasifikasi citra *Hand Sign Language Digits*, VGG-16 mungkin menjadi pilihan yang lebih baik untuk digunakan dalam aplikasi atau tugas yang serupa. Namun, perlu diingat bahwa hasil ini dapat bervariasi tergantung pada dataset dan parameter lainnya yang digunakan dalam pelatihan dan evaluasi model. Oleh karena itu, pemilihan algoritma tergantung pada kebutuhan dan karakteristik dataset yang digunakan.

#### DAFTAR PUSTAKA

- [1] Guan, Q. Wang, Y., Ping, B., Li, D., Du, J., Qin, Yu, Lu, H., Wan, X. & J. Xiang. 2019. Deep convolutional neural network VGG-16 model for differential diagnosing of papillary

- thyroid carcinomas in cytological images: a pilot study. *Journal of Cancer*, Vol 10, No. 20, 4876-4882.
- [2] He, K., Zhang, X., Ren, S. & Sun, J., 2016. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Thn 2016, 770-778, Las Vegas, NV, USA.
- [3] Hijazi, S. L., Kumar, R. & Rowen, C., 2015. *Using Convolutional Neural Networks for Image Recognition*. Cadance, San Jose, CA, USA.
- [4] Kensert, A., Harrison, P. J. & Spjuth, O., 2019. Transfer Learning with Deep Convolutional Neural Networks for Classifying Cellular Morphological Changes. *SLAS Discov.*, Vol 24, No. 4, 466-475.
- [5] Mascarenhas, S. & Agarwal, M., 2021. A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification. *2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON)*. Thn 2021. 96-99. Bengaluru, India.
- [6] Mavi, A., 2011. *A New Dataset and Proposed Convolutional Neural Network Architecture for Classification of American Sign Language Digits*. ArXiv, abs/2011.08927.
- [7] Miranda, N., Novamizanti, L. & Rizal, S., 2020. CONVOLUTIONAL NEURAL NETWORK PADA KLASIFIKASI SIDIK JARI MENGGUNAKAN RESNET-50. *Jurnal Teknik Informatika (Jutif)*, Vol 1, 61-68 .
- [8] Mustapha, A., Mohamed, L., Hamid, H. & Ali, K., 2021. Diabetic Retinopathy Classification Using ResNet50 and VGG-16 Pretrained Networks. *International Journal of Computer Engineering and Data Science*, Vol 1, No. 1, 1-7.
- [9] Perdana, I. P. I., Putra, I. K. G. D. & Dharmadi, I. P. A., 2021. Classification of Sign Language Numbers Using the CNN Method. *JITTER- Jurnal Ilmiah Teknologi dan Komputer*, Vol 2, No.3, 485-493.
- [10] Rathi, P., Gupta, R., Agarwal, S. & Shukla, A., 2020. Sign Language Recognition Using ResNet50 Deep Neural Network Architecture. *5th International Conference on Next Generation Computing Technologies (NGCT-2019)*, 1-7. Uttarakhand, India.
- [11] Simonyan, K. & Zisserman, A., 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. *3rd International Conference on Learning Representations (ICLR 2015)*, Thn 2015, 1–14. California, US.
- [12] Stern, H., Shmueli, M. & Berman, S., 2010. Hand Gesture Classification. *The 3rd Israeli Conference on Robotics*. Thn 2010. Herzelia, Israel.
- [13] Suartika, I. W., Wijaya, A. Y. & Soelaiman, R., 2016. Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101. *JURNAL TEKNIK ITS*, Vol 5, No. 1, A65-A69.
- [14] Vatathanavaro, S., Tungjitnob, S. & Pasupa, K., 2018. White Blood Cell Classification: A Comparison between VGG-16 and ResNet-50 Models. *The 6th Joint Symposium on Computational Intelligence (JSCI6)*. Thn 2018, 1-2. Bangkok, Thailand.
- [15] Wonohadidjojo, D. M., 2021. Perbandingan Convolutional Neural Network pada Transfer Learning Method untuk Mengklasifikasikan Sel Darah Putih. *Jurnal Teknik Informatika*, Vol 13, No. 1, 51-57.