# **BAB IV PENGUJIAN DAN ANALISIS**

Bab ini membicarakan tentang tahap-tahap dalam proses pengenalan aksara Jawa dan hasil dari pengujian pengenalan aksara Jawa dengan menggunakan metode *Transfer Learning*.

## Skenario Simulasi

Skenario simulasi dilakukan dengan model VGG, Inception, dan Xception dengan masing – masing parameter yang berbeda. Berikut adalah gambaran model secara umum untuk memperjelas arsitektur model *transfer learning* yang digunakan.

A group of colorful rectangular objects

Description automatically generated with medium confidence

Masing – masing model yaitu VGG, Inception, dan Exception, akan memiliki 3 layer terakhir yang sama (*pooling, tuned classifier, softmax)*. Perbedaan dari arsitektur tiap model yang diuji, hanya terdapat pada bagian *Feature Extractor*-nya saja, sesuai dengan arsitektur yang disebutkan di bab II.

## Parameter Simulasi

Terdapat beberapa parameter yang akan diuji oleh penulis, rinciannya sebagai berikut:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Augmentasi** | ***Freeze*** | **Learning Rate** | **Layer** | **Optimizer** |
| VGG | *Yes, No* | *Full, ½, None* | 0.001, 0.0001 | 1, 2 | Adam, SGD |
| Inception | *Yes, No* | *Full, ½, None* | 0.001, 0.0001 | 1, 2 | Adam, SGD |
| Xception | *Yes, No* | *Full, ½, None* | 0.001, 0.0001 | 1, 2 | Adam, SGD |

Dari rincian parameter tersebut, akan dilakukan uji coba dengan seluruh kombinasi yang ada, menghasilkan 144 skenario yang berbeda.

Dalam proses *training* masing – masing model, digunakan juga *early stopping callback* untuk menghindarinya *resource* komputasi yang terbuang sia – sia.

## *Performance Matrix*

Hasil pengujian tiap skenario akan menghasilkan waktu *training* dan akurasinya. Untuk menentukan model terbaik dari seluruh pengujian, akan ditentukan berdasarkan Akurasi pada *test set* tertinggi, serta penyelesaian waktu *training* tercepat.

## Analisis Hasil Pengujian

Pelatihan masing – masing model akan disimpan dan dianalisis di dalam bentuk tabel seperti berikut:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Id** | **Model** | **Augmentasi** | ***Freeze*** | **Learning Rate** | **Layer** | **Optimizer** | **Result Train** | **Validation** | **Test** | **Test Loss** | **Time (min)** | **Epoch** |
| 1 | VGG | *Yes* | *Full* | 0.001 | 1 | SGD | 0.21 | 0.55 | 0.51 | 2.63 | 67 | 50 |
| 2 | VGG | *Yes* | *½* | 0.001 | 1 | SGD | 1.00 | 0.99 | 0.99 | 0.04 | 62 | 30 |
| 3 | VGG | *Yes* | None | 0.001 | 1 | SGD | 0.99 | 0.99 | 0.99 | 0.03 | 50 | 13 |
| 4 | VGG | *Yes* | *Full* | 0.0001 | 1 | SGD | 0.07 | 0.12 | 0.11 | 2.95 | 66 | 50 |
| 5 | VGG | *Yes* | *½* | 0.0001 | 1 | SGD | 0.97 | 0.98 | 0.98 | 0.07 | 107 | 50 |
| 6 | VGG | *Yes* | None | 0.0001 | 1 | SGD | 0.96 | 0.98 | 0.98 | 0.06 | 118 | 33 |
| 7 | VGG | *Yes* | *Full* | 0.001 | 2 | SGD | 0.23 | 0.53 | 0.49 | 2.38 | 65 | 50 |
| 8 | VGG | *Yes* | *½* | 0.001 | 2 | SGD | 0.99 | 0.99 | 0.98 | 0.07 | 59 | 28 |
| 9 | VGG | *Yes* | None | 0.001 | 2 | SGD | 0.99 | 0.99 | 0.98 | 0.09 | 72 | 19 |
| 10 | VGG | *Yes* | *Full* | 0.0001 | 2 | SGD | 0.07 | 0.22 | 0.23 | 2.92 | 67 | 50 |
| … | … | … | … | … | … | … | … | … | … | … | … | … |
| 141 | Xception | *No* | *None* | 0.001 | 2 | Adam | 0.99 | 1.00 | 0.98 | 0.13 | 15 | 14 |
| 142 | Xception | *No* | *Full* | 0.0001 | 2 | Adam | 0.80 | 0.82 | 0.75 | 0.87 | 16 | 50 |
| 143 | Xception | *No* | *½* | 0.0001 | 2 | Adam | 1.00 | 0.97 | 0.93 | 0.21 | 8 | 14 |
| 144 | Xception | *No* | *None* | 0.0001 | 2 | Adam | 1.00 | 1.00 | 0.98 | 0.04 | 14 | 15 |

Untuk mempermudah analisis, penulis akan menggunakkan grafik serta bantuan dari *tools* D-tale. Analisis akan dilakukan secara menyeluruh, mulai dari lingkup model serta dampak dari tiap parameter yang digunakan.

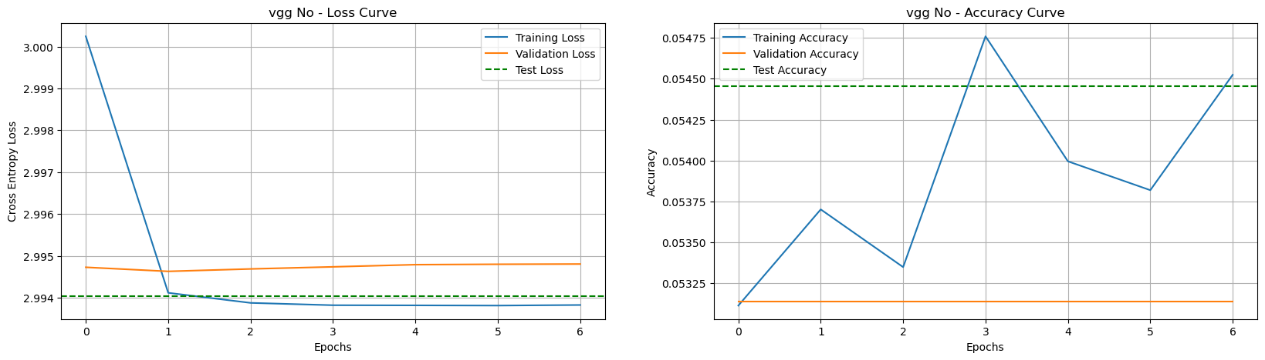
### Evaluasi *Performance Matrix*

A graph of different colored lines

Description automatically generated

Gambar diatas merupakan visualisasi keseluruhan skenario yang telah diuji. *Plotting* dilakukan dengan nilai *index* pada tabel sebagai *x-axis* dan *test accuracy* sebagai *y-axis.*

Dari hasil *plot* tersebut, terlihat bahwa skenario model VGG menghasilkan akurasi *test* yang stagnan hanya menyentuh 5% saja di beberapa index. Setelah diteliti, ternyata model tersebut tidak mengalami peningkatan performa seiring berulangnya tiap *epoch.* Dibawah ini merupakan hasil grafik training dari model salah satu model tersebut.



Hal ini bisa terjadi jika parameter, data, maupun aristektur model yang digunakan memang tidak cocok digunakan untuk kasus ini. Hasil skenario menunjukkan bahwa *optimizer* Adam pada model VGG berpotensi menghasilkan stagnansi *training*.

A graph of different colored bars

Description automatically generated

Dalam *bar-chart* di atas, terlihat bahwa model VGG memiliki rata – rata akurasi terendah dibandingkan dengan model Inception dan Xception. Namun, hal ini bisa saja bias dikarenakan adanya beberapa hasil model yang mengalami *training* stagnan. Di bawah ini merupakan grafik untuk melihat jumlah masing – masing model yang menyentuh akurasi *test* diatas 50%.

A graph of different colored bars

Description automatically generatedA graph with blue rectangular bars

Description automatically generated with medium confidence

Berbeda dengan sebelumnya, grafik di atas ini menunjukkan bahwa performa VGG tidak kalah dengan model Inception dan Xception. Namun, jumlah model yang memasuki kategori akurasi terbanyak, terdapat di model Inception sebanyak 40 model.

A graph of different colored lines

Description automatically generatedA graph of a number and number

Description automatically generated with medium confidence

Grafik … (kiri), menunjukkan waktu yang dibutuhkan seluruh model untuk menyelesaikan *training*-nya. Grafik … (kanan), merata – rata waktu *training* dari masing – masing model. Berdasarkan informasi tersebut, secara umum Inception telah dibuktikan memiliki waktu *training* tercepat dibandingkan dengan model VGG dan Xception.

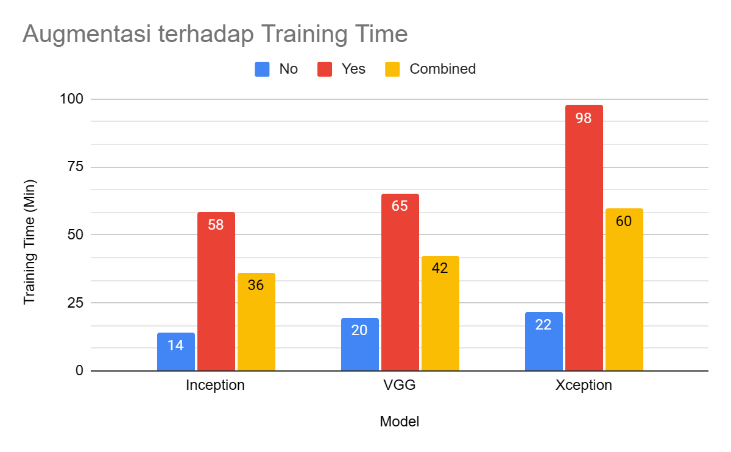
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Id** | **Model** | **Augmentasi** | ***Freeze*** | **Learning Rate** | **Layer** | **Optimizer** | **Result Train** | **Validation** | **Test** | **Test Loss** | **Time (min)** | **Epoch** |
| 95 | Inception | *No* | *½* | 0.0001 | 2 | Adam | 1.00 | 0.98 | 0.95 | 0.17 | 3 | 11 |
| 86 | Inception | *No* | *½* | 0.001 | 1 | Adam | 0.99 | 0.99 | 0.99 | 0.05 | 3 | 11 |
| 44 | VGG | *No* | *½* | 0.001 | 2 | Adam | 0.05 | 0.05 | 0.05 | 3.00 | 4 | 8 |
| 139 | Xception | *No* | *Full* | 0.001 | 2 | Adam | 0.75 | 0.79 | 0.74 | 0.91 | 5 | 15 |
| 134 | Xception | *No* | *½* | 0.001 | 1 | Adam | 0.98 | 0.99 | 0.93 | 0.16 | 5 | 9 |
| … | … | … | … | … | … | … | … | … | … | … | … | … |

Informasi di atas merupakan hasil uji skenario tabel yang telah diurutkan berdasarkan kolom *Time* secara *ascending*. Inception masih menjadi model tercepat yang menyelesaikan proses *training-*nya hanya dengan 11 *epoch*.

### Evaluasi Parameter terhadap *Performance Matrix*

Dalam bagian ini, penulis akan menjabarkan parameter – parameter yang digunakan dan pengaruhnya masing – masing parameter tersebut terhadap performa model dari segi waktu *training* dan akurasi *test set.*

A graph with different colored bars

Description automatically generated

A graph of different colored bars

Description automatically generatedA graph of numbers and a number

Description automatically generated

A graph of different colored bars

Description automatically generatedA graph of numbers and a number of people

Description automatically generated with medium confidence

A graph of different colored bars

Description automatically generatedA graph of numbers and a number of people

Description automatically generated with medium confidence

A graph of different colored bars

Description automatically generatedA graph of numbers and a number

Description automatically generated with medium confidence

Augmentasi: Parameter ini berpengaruh terhadap akurasi test dan time. Pengaruh time sangat besar jika augmentasi dilakukan

Freeze: Parameter freeze dengan value “None” paling berdampak signifikan terhadap akurasi test untuk model Inception dan Xception. Sedangkan untuk model VGG, akurasi tertinggi justru terdapat pada value “Full”. Freeze juga berpengaruh terhadap training time. Semakin banyak layer yang tidak di-freeze, maka semakin lama waktu pelatihan model dilakukan.

Learning rate: Default learning rate yaitu 0.001 menghasilkan akurasi yang lebih besar dibandingkan dengan value 0.0001. Namun, hal ini hanya terjadi untuk model Inception dan Xception saja. Di sisi lain, VGG tidak mengalami dampak yang signifikan terhadap perubahan parameter learning rate. Terkait waktu pelatihan model, learning rate dengan value 0.0001 akan menghasilkan waktu yang lebih lama. Hal ini dikarenakan akan memperlambat proses konvergensi dalam pelatihan model.

Layer: Parameter layer terhadap ketiga model tidak memberikan dampak yang signifikan terhadap hasil akurasi test maupun lama pelatihan waktu model.

Optimizer: Parameter ini berdampak signifikan terhadap akurasi test maupun lama waktu pelatihan model. Terjadi keunikan untuk model VGG, hasil akurasi test yang dihasilkan dengan nilai parameter SGD jauh lebih baik dibandingkan dengan parameter Adam. Untuk kedua model lainnya yaitu Inception dan Xception, hasil akurasi test yang lebih baik, dihasilkan dari parameter Adam. Waktu training time untuk nilai parameter SGD, secara umum menghabiskan waktu lebih banyak dibandingkan parameter Adam.

A colorful squares with text

Description automatically generated with medium confidence

Didukung dengan gambar *correlation matrix*, hal – hal yang sudah disebutkan sebelumnya, memberikan statement yang sama.

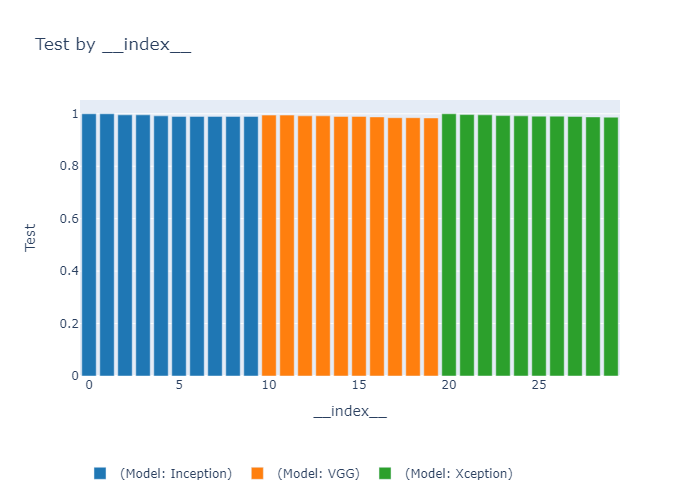
### Evaluasi Top 10 Models

Pada bagian ini, penulis akan membahas masing – masing 10 model terbaik tiap jenis model yang berbeda, berdasarkan nilai akurasi testnya. Dikarenakan terdapat 3 jenis model, maka total data yang akan didapat adalah 30 model berbeda.

Hal ini dilakukan untuk melihat lebih detail terkait parameter yang dapat digunakan untuk menghasilkan model terbaik. Penulis akan menganalisis secara garis besar, “parameter apa yang terdapat di dalam 30 model terbaik ini”.

Di bawah ini merupakan rincian daftar model yang diambil.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Id** | **Model** | **Augmentasi** | ***Freeze*** | **Learning Rate** | **Layer** | **Optimizer** | **Result Train** | **Validation** | **Test** | **Test Loss** | **Time (min)** | **Epoch** |
| 65 | Inception | *Yes* | None | 0.0001 | 1 | Adam | 0.9982 | 0.9961 | 1 | 0.002035 | 38 | 14 |
| 89 | Inception | *No* | None | 0.0001 | 1 | Adam | 0.9991 | 0.9952 | 1 | 0.004126 | 11 | 19 |
| 62 | Inception | *Yes* | None | 0.001 | 1 | Adam | 0.9956 | 1 | 0.996 | 0.00795 | 74 | 27 |
| 68 | Inception | *Yes* | None | 0.001 | 2 | Adam | 0.9949 | 0.9942 | 0.996 | 0.009453 | 52 | 19 |
| 50 | Inception | *Yes* | None | 0.001 | 1 | SGD | 0.999 | 0.9903 | 0.9921 | 0.02624 | 65 | 23 |
| 56 | Inception | *Yes* | None | 0.001 | 2 | SGD | 0.9996 | 0.9961 | 0.9901 | 0.03117 | 79 | 31 |
| 74 | Inception | *No* | None | 0.001 | 1 | SGD | 0.9988 | 0.9952 | 0.9901 | 0.05131 | 28 | 50 |
| 86 | Inception | *No* | None | 0.001 | 1 | Adam | 0.9882 | 0.9952 | 0.9901 | 0.04759 | 9 | 16 |
| 91 | Inception | *No* | ½ | 0.001 | 2 | Adam | 0.9962 | 0.9903 | 0.9901 | 0.03443 | 5 | 17 |
| 95 | Inception | *No* | None | 0.0001 | 2 | Adam | 0.9941 | 1 | 0.9901 | 0.03589 | 8 | 14 |
| 25 | VGG | *No* | ½ | 0.001 | 1 | SGD | 0.9838 | 0.9903 | 0.995 | 0.1023 | 22 | 44 |
| 40 | VGG | *No* | ½ | 0.0001 | 1 | Adam | 0.9971 | 1 | 0.995 | 0.01521 | 18 | 35 |
| 1 | VGG | *Yes* | ½ | 0.001 | 1 | SGD | 0.9952 | 0.9942 | 0.9921 | 0.04196 | 62 | 30 |
| 2 | VGG | *Yes* | None | 0.001 | 1 | SGD | 0.9871 | 0.9913 | 0.9921 | 0.02506 | 50 | 13 |
| 31 | VGG | *No* | ½ | 0.001 | 2 | SGD | 0.9471 | 0.9855 | 0.9901 | 0.0664 | 19 | 39 |
| 46 | VGG | *No* | ½ | 0.0001 | 2 | Adam | 0.9927 | 0.9952 | 0.9901 | 0.02785 | 24 | 50 |
| 11 | VGG | *Yes* | None | 0.0001 | 2 | SGD | 0.9454 | 0.9826 | 0.9881 | 0.04981 | 185 | 50 |
| 23 | VGG | *Yes* | None | 0.0001 | 2 | Adam | 0.9931 | 0.9961 | 0.9851 | 0.07386 | 66 | 19 |
| 26 | VGG | *No* | None | 0.001 | 1 | SGD | 0.9888 | 0.9807 | 0.9851 | 0.1034 | 26 | 28 |
| 7 | VGG | *Yes* | ½ | 0.001 | 2 | SGD | 0.9918 | 0.9932 | 0.9842 | 0.06673 | 59 | 28 |
| 134 | Xception | *No* | None | 0.001 | 1 | Adam | 0.9815 | 0.9855 | 1 | 0.01162 | 10 | 9 |
| 113 | Xception | *Yes* | None | 0.0001 | 1 | Adam | 0.9992 | 0.9981 | 0.997 | 0.01017 | 45 | 10 |
| 104 | Xception | *Yes* | None | 0.001 | 2 | SGD | 0.9991 | 0.9865 | 0.996 | 0.0247 | 221 | 50 |
| 116 | Xception | *Yes* | None | 0.001 | 2 | Adam | 0.9912 | 0.9971 | 0.9931 | 0.02588 | 44 | 10 |
| 119 | Xception | *Yes* | None | 0.0001 | 2 | Adam | 0.999 | 1 | 0.9921 | 0.0427 | 83 | 19 |
| 98 | Xception | *Yes* | None | 0.001 | 1 | SGD | 0.9994 | 0.9894 | 0.9911 | 0.03065 | 220 | 50 |
| 115 | Xception | *Yes* | ½ | 0.001 | 2 | Adam | 0.9949 | 0.9981 | 0.9911 | 0.02447 | 37 | 13 |
| 137 | Xception | *No* | None | 0.0001 | 1 | Adam | 0.9982 | 0.9903 | 0.9901 | 0.02281 | 11 | 11 |
| 110 | Xception | *Yes* | None | 0.001 | 1 | Adam | 0.9949 | 0.9971 | 0.9881 | 0.0348 | 52 | 12 |
| 109 | Xception | *Yes* | ½ | 0.001 | 1 | Adam | 0.992 | 0.9971 | 0.9871 | 0.041 | 23 | 8 |



Grafik di atas merupakan 30 model beserta performa akurasi testnya.

## Percobaan Model dengan Data Baru

* 1. Parameter Pengambilan Model
  2. Data

**DAFTAR PUSTAKA**

# **DAFTAR PUSTAKA BUKU**