



PERBANDINGAN ARSITEKTUR VGG16 DAN RESNET50 UNTUK REKOGNISI TULISAN TANGAN AKSARA LAMPUNG

Rikendry¹, Azzin Maharil²

Fakultas Teknik dan Ilmu Komputer Universitas Teknokrat Indonesia

Corresponding email : rikendry@teknokrat.ac.id

Received: 10 January 2022 Accepted: 23 March 2022 Published: 30 June 2022

Abstract

Handwriting recognition or commonly called Hand Writing Recognition (HWR) is a computer system that can be used to recognize letters that come from handwriting. HWR itself is a system developed from Optical Character Recognition (OCR), with the HWR system, it is hoped that the computer can read and recognize the letters or characters entered by the user in handwritten form. The data used in this research is in the form of handwritten image data of Lampung script as many as 20 characters. In this study, the author will compare two deep learning models, namely VGG16 and ResNet50. The results of this study indicate that the VGG16 architectural model provides 91% accuracy results and better training time, while for ResNet50 it provides 65% accuracy results and requires longer computational time and small accuracy results.

Keywords: Lampung characters, handwriting s, deep learning

Abstrak

Pengenalan tulisan tangan atau biasa disebut Hand Writing Recognition (HWR) adalah sebuah sistem komputer yang dapat digunakan untuk mengenali huruf yang berasal dari tulisan tangan. HWR sendiri merupakan sistem yang dikembangkan dari Optical Character Recognition (OCR), dengan adanya sistem HWR ini diharapkan komputer dapat membaca dan mengenali huruf-huruf ataupun karakter yang dimasukkan oleh user dalam bentuk tulisan tangan. Data yang digunakan pada penelitian ini berupa data gambar tulisan tangan aksara lampung sebanyak 20 aksara. Pada penelitian kali ini penulis akan membandingkan dua model deep learning yaitu VGG16 dan ResNet50. Hasil dari penelitian ini menunjukkan bahwa model arsitektur VGG16 memberikan hasil akurasi 91% serta waktu training yang lebih baik, sedangkan untuk ResNet50 memberikan hasil akurasi sebesar 65% dan memerlukan waktu komputasi yang lebih lama dan hasil akurasi yang kecil.

Kata kunci : Aksara Lampung, Tulisan Tangan, Deep Learning

To cite this article:

Rikendry and Maharil. (2022). Perbandingan Arsitektur Vgg16 Dan Resnet50 Untuk Rekognisi Tulisan Tangan Aksara Lampung. *Jurnal Informatika dan Rekayasa Perangkat Lunak*, Vol. 3, No. 2, 236-243

PENDAHULUAN

Indonesia merupakan negara dengan budaya yang memiliki adat istiadat yang beragam. Semua itu bisa dilihat dari banyaknya budaya yang telah di wariskan dari zaman dahulu oleh masing masing daerah di Indonesia. Salah satu yang bisa kita lihat suku Lampung. Suku lampung memiliki beragam warisan budaya, salah satunya yaitu kain tapis, tarian, aksara lampung, dan masih banyak lagi. Media tulisan aksara lampung pun bermacam macam. Bisa menggunakan kulit kayu, bilah bambu, daun lontar, kulit hewan, dan lain-lain.

Dengan mengingat betapa pentingnya budaya dan usaha dalam mengenali, pelestarian, dan penyebaran maka pada penelitian kali ini akan membahas mengenai upaya digital dan pengolahan pada hasil digital, misalnya pengenalan untuk aksara Lampung. Pengenalan tulisan tangan atau biasa disebut Hand Writing Recognition (HWR) adalah sebuah sistem komputer yang dapat digunakan mengenali huruf yang berasal dari tulisan tangan. HWR sendiri merupakan sistem yang dikembangkan dari Optical Character Recognition (OCR) yang dapat mengenali suatu huruf atau angka yang dapat langsung menterjemahkannya ke dalam bentuk dokumen teks. Dengan adanya sistem HWR ini diharapkan komputer dapat membaca huruf-huruf ataupun karakter yang dimasukkan oleh user dalam bentuk tulisan tangan menjadi dokumen text tanpa harus mengetikkan ulang. pada tugas akhir kali akan di lakukan analysis untuk pengenalan tulisan tangan pada sisitem komputer menggunakan metode Convolution Neutral Network.

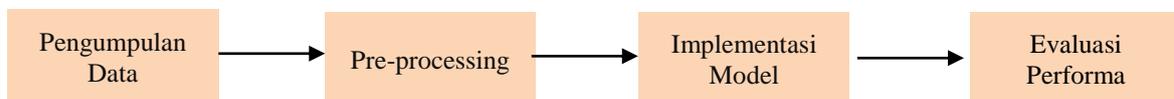
Banyaknya hidden layer yang digunakan antara input layer dan output layer, maka jaringan ini dapat dikatakan sebagai deep neural network. Beberapa tahun belakangan Deep Learning telah menunjukkan kinerja yang sangat baik. Sebagian besar disebabkan oleh komputasi yang lebih baik, kumpulan data yang berukuran besar, dan teknik untuk melatih suatu jaringan. (Goodfellow, Bengio, Y, dan Courville, A., 2016). Kemampuan CNN dianggap sebagai model terbaik untuk memecahkan masalah pengenalan atau pendeteksian suatu objek. Pada tahun 2012 dengan akurasi yang baik bisa bersaing dengan manusia pada data-data tertentu. (A. Coates, H.Lee, A.Y. Ng, 2011). Convolutional Neural Network atau bisa disingkat dengan (CNN) merupakan sebuah pengembangan lebih lanjut dari algoritma terdahulu yaitu Artificial Neural Network (ANN) yang menerima sebuah input berupa sebuah citra gambar. Namun, pengembangan algoritma CNN masih cukup mahal, dari waktu yang dihabiskan untuk melatih model dan juga perangkat keras yang dibutuhkan untuk melatih model. Jadi untuk menyelamatkan proses pengembangan Model CNN dan juga fokus sebagian besar pengguna menggunakan Model CNN. model Convolutional Neural Network (CNN) yang telah mengalami proses pelatihan disimpan dalam bentuk pra- model terlatih. Keras yang merupakan API tingkat tinggi untuk jaringan saraf tiruan menyediakan beberapa model pra-terlatih seperti: VGG16 dan RESNET50.

Berdasarkan latar belakang yang telah di sebutkan di atas, penelitian ini menerapkan penerapan teknik deep learning yang menggunakan model arsitektur dari CNN untuk membantu mengidentifikasi huruf-huruf aksara lampung. Penelitian ini berfokus pada pengklasifikasian citra tulisan tangan aksara lampung dengan membandingkan dua pre-trained model. VGG16 merupakan model CNN yang memanfaatkan convolutional layer dengan spesifikasi convolutional filter kecil (3x3). Dengan ukuran filter convolutional, kedalaman jaringan saraf dapat ditingkatkan dengan lebih banyak lapisan convolutional. Hasilnya, model CNN menjadi lebih akurat dibandingkan model CNN sebelumnya. Model VGG16 memiliki 19 lapisan yang terdiri dari 16 lapisan konvolusi dan 3 lapisan yang terhubung penuh.

Sedangkan Model ResNet Model ResNet adalah model yang menggunakan kerangka pembelajaran residual yang mendalam. Menggunakan kerangka kerja ini, setiap lapisan jaringan memiliki referensi ke lapisan jaringan sebelumnya; ini membuat proses optimasi lebih mudah daripada lapisan jaringan yang tidak memiliki konektivitas. Karena proses optimasi yang lebih mudah, jaringan syaraf tiruan yang dibentuk dapat memiliki banyak lapisan hingga 34 lapisan dan akibatnya akurasi dari jaringan saraf tiruan menjadi meningkat dari jaringan syaraf tiruan yang tidak menggunakan jaringan residual.

METODE PENELITIAN

Secara umum pada penelitian ini terdapat empat tahap penelitian yaitu, pengumpulan data, pre-processing, implementasi model arsitektur, analisis hasil.



1. Pengumpulan Data (Data Collection)

Dataset yang digunakan pada penelitian ini di dapat dari, penelitian terdahulu yang telah dikumpulkan, Dataset terdiri dari 20 kelas, yaitu ka,ga,nga,pa,ba,ma,ta,da,na,ca,ja,nya,ya,a,la,ra,sa,wa,ha,gha, berikut adalah contoh data yang akan digunakan, dapat dilihat pada gambar 2.



Gambar 1 dataset aksara lampung

2. **Pemrosesan (Pre-processing)**

Pre-processing yang dilakukan pada penelitian kali ini, Cropping dan augmentasi citra. Cropping dilakukan secara otomatis dengan ukuran yang telah ditetapkan sebesar 224x224 pixel, adapun yang dilakukan saat proses cropping adalah untuk meningkatkan performa saat komputasi. Sedangkan augmentasi dilakukan agar dataset semakin banyak, hal ini bertujuan untuk membuat hasil komputasi semakin baik.

3. **Penerapan Model (Implementasi Model)**

Pada penelitian ini akan di implementasikan model arsitektur sebanyak dua model, yaitu VGG16 dan ResNet50

a. VGG16

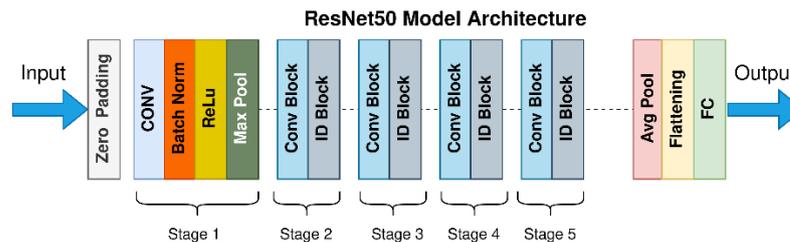
VGG16 merupakan model arsitektur jaringan dasar yang menggunakan model *Convolutional Neural Network* (CNN) dengan jumlah lapisan layer sebanyak 16 lapisan. Model VGG16.



Gambar 2 arsitektur VGG16

b. ResNet50

ResNet50 merupakan model arsitektur jaringan dasar yang menggunakan *Convolutional Neural Network* (CNN) dengan jumlah layer atau lapisan sebanyak 50 lapisan.



Gambar 3 arsitektur resnet 50

Dua model deep learning yang telah dijelaskan akan diuji menggunakan pendekatan deep learning model yang akan digunakan sudah dilatih dahulu. Adapun parameter yang digunakan yaitu learning rate 0.001, batch size 64, dan jumlah epoch untuk Vgg16 berjumlah 25 epoch sedangkan untuk ResNet50 hanya menggunakan 25 epoch Fungsi aktivasi yang digunakan di penelitian ini pada layer konvolusi adalah Rectified Linear Unit (ReLU), optimizer yang digunakan Adam, dan untuk loss function menggunakan categorical_crossentropy dikarenakan jenis dataset yang memiliki lebih dari 1 label. Setiap model dilatih dengan menggunakan 80% data gambar dan 20% untuk testing pada masing-masing kelas. Model yang diimplementasikan menggunakan library Tensorflow dan Keras.

4. **Evaluasi Performa**

Setiap model deep learning dievaluasi menggunakan performa akurasi, presisi, recall, F1-score, dan waktu komputasi. Performa akurasi dihitung dengan menggunakan persamaan (1), nilai presisi dihitung dengan menggunakan persamaan (2), recall dihitung dengan menggunakan persamaan (3), dan F1-score dihitung dengan menggunakan persamaan (4).

$$1. \text{ Akurasi} = \frac{TP+TN}{TP+TN+FP+FN}$$

2. $Presisi = \frac{TP}{TP+FP}$
3. $Recall = \frac{TP}{TP+FN}$
4. $F1 - Score = 2x \frac{presisi \times recall}{presisi+recall}$

Suatu kondisi dikatakan True Positive (TP) atau True Negative (TN) ketika suatu observasi dikenali dengan benar, sedangkan False Positive (FP) dan False Negative (FN) merupakan kondisi suatu observasi salah diidentifikasi.

HASIL DAN PEMBAHASAN

1. Implementasi VGG16

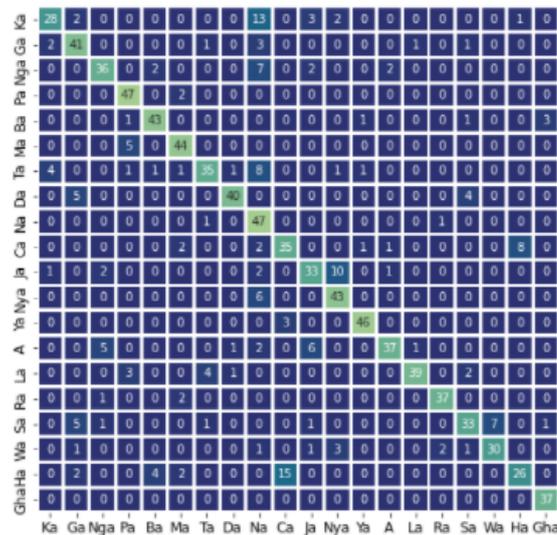
Pada tahapan ini akan dilakukan implementasi model CNN dengan arsitektur VGG16 dan ResNet50 sebanyak 3 kali pada masing-masing arsitektur bertujuan untuk memperoleh nilai rata-rata akurasi dari setiap model Performa untuk model deep learning yang telah latih dengan pengukuran performa yang ada pada pembahasan sebelumnya dapat dilihat pada Gambar 4.

```
Accuracy : 81.276%
Loss      : 1.474
```

Classification Report :

	precision	recall	f1-score	support
Ka	0.78	0.63	0.70	49
Ga	0.77	0.88	0.82	49
Nga	0.79	0.69	0.74	49
Pa	0.84	0.94	0.88	49
Ba	0.90	0.88	0.89	49
Ma	0.82	0.86	0.84	49
Ta	0.86	0.79	0.82	53
Da	0.95	0.82	0.88	49
Na	0.63	0.90	0.74	49
Ca	0.74	0.71	0.73	49
Ja	0.66	0.82	0.73	49
Nya	0.75	0.84	0.79	49
Ya	0.94	0.94	0.94	49
A	0.88	0.69	0.77	52
La	0.96	0.88	0.91	49
Ra	0.93	0.93	0.93	40
Sa	0.77	0.76	0.76	49
Wa	0.88	0.72	0.79	39
Ha	0.78	0.65	0.71	49
Gha	0.86	1.00	0.92	37

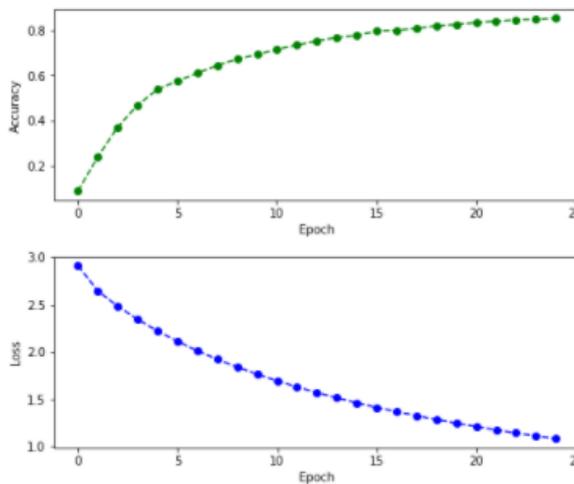
Gambar 4 Pengujian Performa VGG16



Gambar 5 Gambar hasil prediksi VGG16

Berdasarkan Gambar 4 dapat dilihat bahwa untuk model VGG16 memperoleh akurasi performa sebesar 81% dari keseluruhan data testing, Nilai Precision tertinggi terdapat pada kelas La dengan tingkat sensitivitas sebesar 96%. dalam hal membedakan objek dan Precision terendah pada kelas Na sebesar 63%. Nilai Recall tertinggi terdapat pada kelas Gha, sebesar 100% untuk mengenali objeknya sendiri dan Recall terendah pada kelas Ka sebesar 63%.

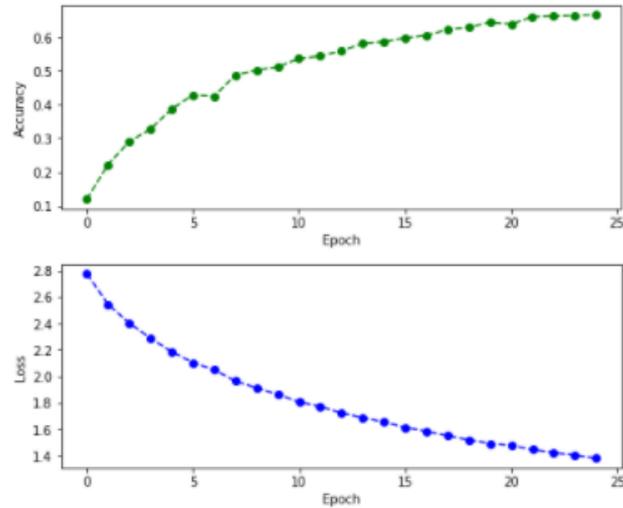
Grafik akurasi pelatihan model VGG16, pada penelitian kali ini tidak menggunakan validasi data, dikarenakan menghasilkan banyak terjadinya kesalahan prediksi, untuk grafik hasil dapat dilihat pada Gambar 6. dapat dilihat bahwa model VGG16



Gambar 6 Performa akurasi dan loss VGG16.

2. Implementasi ResNet50

Untuk hasil akurasi dari model ResNet50 mencapai 96% jika di dibandingkan dengan VGG16 maka dapat disimpulkan bahwa hasil dari performa ResNet50 lebih baik di dibandingkan VGG16, dapat dilihat pada Gambar 8.



Gambar 7 Performa akurasi dan loss ResNet50.

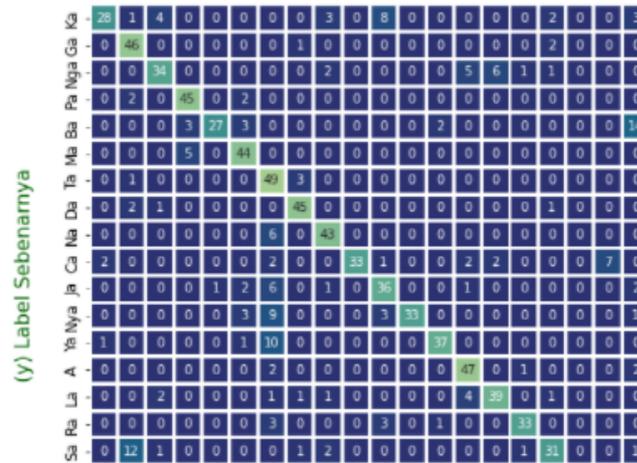
Accuracy : 70.293%
 Loss : 1.324

Classification Report :

	precision	recall	f1-score	support
Ka	0.87	0.41	0.56	49
Ga	0.59	0.90	0.71	49
Nga	0.88	0.43	0.58	49
Pa	0.54	0.98	0.70	49
Ba	0.96	0.53	0.68	49
Ma	0.77	0.84	0.80	49
Ta	0.51	0.85	0.63	53
Da	0.85	0.71	0.78	49
Na	0.89	0.63	0.74	49
Ca	1.00	0.63	0.78	49
Ja	0.42	0.80	0.55	49
Nya	0.93	0.57	0.71	49
Ya	1.00	0.71	0.83	49
A	0.55	0.92	0.69	52
La	0.97	0.61	0.75	49
Ra	1.00	0.57	0.73	40
Sa	0.67	0.63	0.65	49
Wa	0.93	0.72	0.81	39
Ha	0.82	0.76	0.79	49
Gha	0.63	0.84	0.72	37

Gamabar 8 Pengujian Performa ResNet50

Berdasarkan Gambar 5 dapat dilihat bahwa model ResNet50 memperoleh akurasi performa sebesar 70%, Untuk nilai *Precision* tertinggi terdapat pada kelas Ca dan La adengan tingkat sensitivitas sebesar 100%. dalam membedakan objek lain dan *Precision* terendah pada kelas Ja sebesar 42%. Nilai *Recall* tertinggi terdapat pada kelas Pa, sebesar 98% dalam mengenali objeknya sendiri dan *Recall* terendah pada kelas Ka sebesar 41%.



Gambar 9 Gambar hasil prediksi ResNet50

Meskipun akurasi dari VGG16 hanya sebesar 92% akan tetapi waktu komputasi yang diberikan lebih cepat dibandingkan Resnet50.

HASIL PERBANDINGAN ARSITEKTUR

Pada tahap ini merupakan hasil perbandingan dari arsitektur VGG16 dan ResNet50 dari hasil pengujian data dapat dilihat pada tabel 1

Tabel 1 perbandingan Vgg16 dan ResNet50

Arsitektur	VGG16	ResNet50
Time/sec	6min 23s	5min 21s
Train loss	1.0843	1.4046
Train accuracy	0.8552	0.6667
Precision dan Recall overall	81.276%	70.293%
Trainable parameter	10,260	40,980
Kapasitas Penyimpanan	57.704MB	109.217MB

Pada Tabel 1 dapat di lihat perbandingan dari masing-masing model arsitektur yang menghasilkan

Kesimpulan Dan Saran

Dari penelitian ini dapat disimpulkan bahwa untuk pengenalan aksara lampung, dengan arsitektur VGG16 adalah arsitektur yang cukup baik dalam hal akurasi maupun jumlah parameter, sedangkan untuk ResNet50 sendiri kurang dapat memberikan hasil yang baik untuk proses training, tetapi memberikan hasil yang baik untuk waktu komputasi.

REFERENSI/DAFTAR PUSTAKA

- [1] Nurfitra, R. D. & Ariyanto, G. Implementasi Deep Learning Berbasis Tensorflow Untuk Pengenalan Sidik Jari. *Emitor: Jurnal Teknik Elektro* 18, 22–27 (2018).
- [2] Lorentius, C. A., Adipranata, R. & Tjondrowiguno, A. Pengenalan Aksara Jawa dengan Menggunakan Metode Convolutional Neural Network. *e-Proceeding of Engineering* 7, 2558–2567 (2020).
- [3] Rohim, A., Sari, Y. A. & Tibyani. Convolution neural network (cnn) untuk pengklasifikasian citra makanan tradisional. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* 3, 7038–7042 (2019).
- [4] Aryantio, A. & Munir, R. Pengenalan Aksara Lampung Menggunakan Jaringan Syaraf Tiruan. *Jurnal Konferensi Nasional Informatika* 34–38 (2015).

- [5] Handhayani, T. Identifikasi Penulis Melalui Pola Tulisan Tangan. *Jurnal Muara Sains, Teknologi, Kedokteran, dan Ilmu Kesehatan* 1, 210–217 (2017).