



Seminar Nasional Keinsinyuran (SNIP)

Alamat Prosiding: snip.eng.unila.ac.id



Pengenalan Aksara Lampung Menggunakan Metode CNN (Convolutional Neural Network)

Nanda Putri Sazqiah^{a,*}, Yessi Mulyani^b, dan Meizano Ardhi Muhammad^c, dan Martinus^d, Irza Sukmana^e, Gigih Forda Nama^f, Zulmiftah Huda^g, Trisya Septiana^h, Panji Kurniawanⁱ, Gita Paramita Djausal^j

^{a,b,c,e,f,g,h}Program Studi Teknik Informatika, Universitas Lampung, Bandar Lampung, Jl. Prof. Sumantri Brojonegoro, Bandar Lampung 35145

ⁱJurusan Arsitektur, Universitas Lampung, Bandar Lampung, Jl. Prof. Sumantri Brojonegoro, Bandar Lampung 35145

^{d,e}Jurusan Teknik Mesin, Universitas Lampung, Jl. Prof. Soemantri Brojonegoro, Bandar Lampung 35145

^jJurusan Administrasi Bisnis, Universitas Lampung, Jl. Prof. Soemantri Brojonegoro, Bandar Lampung 35145

INFORMASI ARTIKEL

ABSTRAK

Riwayat artikel:
Diterima 02 Maret 2022
Direvisi 16 Maret 2022
Diterbitkan 22 April 2022

Kata kunci:
Aksara Lampung
Convolutional Neural Network
CNN

Penelitian ini bertujuan untuk membuat sistem yang mampu mengenali Aksara Lampung dengan metode CNN (Convolutional Neural Network). Sebelum penelitian ini, telah ada beberapa penelitian yang bertujuan serupa dengan metode yang berbeda. Penelitian ini memiliki tingkat kesulitan sendiri dalam hal mengumpulkan dataset yang akan digunakan untuk proses pelatihan dikarenakan dataset tidak tersedia secara publik sehingga harus dibangun sendiri. Secara garis besar, proses pengenalan Aksara Lampung dapat dibagi menjadi 3 bagian, yaitu : Pengumpulan bahan (dataset), Ekstraksi fitur dari setiap aksara Lampung dan Image Classification. Sistem pengenalan aksara Lampung ini dibangun menggunakan tools Google Colaboratory dengan Bahasa pemrograman Python. Berdasarkan hasil perhitungan menggunakan confusion matrix didapatkan hasil akurasi sebesar 100%. Hal ini dikarenakan pola yang digunakan sangat sederhana karena aksara Lampung memiliki pola yang sederhana dan bentuk yang hampir sama hanya yang membedakan adalah rotasinya. Sistem ini sudah dilakukan proses Save Model dimana bertujuan untuk meminimalisir waktu untuk training ulang dan hasil save model berupa file bisa digunakan untuk pengembangan proyek baru seperti android atau embedded system dengan ukuran file yaitu 211 KB

1. Pendahuluan

Indonesia merupakan Negara kepulauan terbesar di dunia yang terdiri dari 5 pulau besar dan 4 kepulauan. Secara keseluruhan Indonesia memiliki 17.504 pulau. Tercatat pada tahun 2015, jumlah penduduk Indonesia mencapai 255 juta jiwa dengan bahasa, etnis, adat istiadat, dan agama yang berbeda-beda. Indonesia merupakan negara yang kaya akan budaya. Salah satu dimensi kebudayaan yang sangat berpengaruh adalah Bahasa Daerah. Menurut sumber dari Buku Pedoman Pencacah Sensus Penduduk BPS 2010 dan buku Kewarganegaraan, Suku Bangsa, Agama, dan Bahasa Sehari-hari Penduduk Indonesia, terdapat 1211 bahasa (1158 bahasa daerah) di Indonesia. Seiring dengan berkembangnya teknologi dan komunikasi global kondisi budaya kita semakin terkikis. Ratusan bahasa daerah yang ada di Indonesia terancam punah. Salah satu bahasa daerah yang terancam punah yaitu Bahasa Lampung.

Dalam bahasa daerah dikenal juga huruf sebagai bentuk penulisan atau representasi dari bahasa tersebut. Salah satu bahasa yang memiliki huruf khusus sebagai bentuk penulisan dari bahasa tersebut adalah bahasa Lampung, dengan tulisan lampung atau yang lebih dikenal dengan Aksara Lampung. Huruf Lampung dikenal juga sebagai Had Lampung adalah salah satu aksara tradisional Nusantara yang digunakan untuk menulis bahasa Lampung.

Convolutional Neural Network (CNN) merupakan salah satu algoritma dari Deep Learning yang merupakan pengembangan dari Multilayer Perceptron (MLP) yang dirancang untuk mengolah data dua dimensi, misalnya gambar/citra atau suara. CNN digunakan untuk mengklasifikasi data yang terlabel dengan menggunakan metode supervised learning.

Aksara Lampung memiliki bentuk yang unik, namun saat ini minat masyarakat akan aksara Lampung sangat memprihatinkan. Melihat bagaimana pentingnya nilai / eksistensi sebuah budaya,

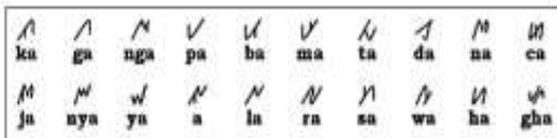
*Penulis korespondensi.

E-mail: nanda@gmail.com

penulis mengusulkan sebuah sistem untuk mengenali huruf Lampung. Dengan adanya sistem ini maka akan membantu pemerintah Indonesia dalam mengembangkan pelestarian huruf Lampung. Sistem ini menggunakan metode CNN (Convolutional Neural Network) karena sebelumnya belum pernah ada penelitian pengenalan aksara Lampung menggunakan metode ini. Secara garis besar, proses pengenalan huruf Lampung dapat dibagi menjadi 3 bagian, yaitu : Pengumpulan bahan (dataset), Ekstraksi fitur dari setiap aksara Lampung dan Image Classification atau pengenalan aksara Lampung.

1.1. Aksara Lampung

Aksara Lampung terdiri dari induk huruf yang berjumlah 20 huruf, yakni ka-ga-nga-pa-ba-ma-ta-da-na-ca-ja-nya-ya-a-la-ra-sa-wa-ha-gha. Serta atribut lain seperti anak huruf, anak huruf ganda dan gugus konsonan, juga terdapat lambang, angka dan tanda baca. Aksara Lampung atau sering disebut Kaganga dibaca dan ditulis dari kiri ke kanan. Aksara Lampung diperkirakan masuk pada era Kerajaan Sriwijaya (700-1.000 Masehi).



Gambar 1. Aksara Lampung

1.2. Deep Learning

Deep Learning (Pembelajaran Dalam) merupakan salah satu cabang ilmu pembelajaran mesin (Machine Learning) yang terdiri dari algoritma pemodelan abstraksi tingkat tinggi pada data yang menggunakan sekumpulan fungsi transformasi non-linier yang ditata secara berlapis dan mendalam. Teknik dan algoritma dalam pembelajaran dalam dapat digunakan baik untuk kebutuhan pembelajaran terarah (*supervised learning*), pembelajaran tak terarah (*unsupervised learning*), dan semi-terarah (*semi-supervised learning*) dalam berbagai aplikasi seperti pengenalan citra, pengenalan suara, klasifikasi teks, dan sebagainya.

1.3. Jaringan Syaraf Tiruan

Jaringan Syaraf Tiruan atau Artificial Neural Network (ANN) adalah model yang terinspirasi oleh bagaimana neuron dalam otak manusia bekerja. Tiap neuron pada otak manusia saling berhubungan dan informasi mengalir dari setiap neuron tersebut. Gambar di bawah adalah ilustrasi neuron dengan model matematisnya.



Gambar 2. Ilustrasi Neuron Dengan Model Matematisnya

2. Sistem Pengenalan Aksara Lampung

Secara garis besar, proses pengenalan huruf Lampung dapat dibagi menjadi 3 bagian, yaitu : Pengumpulan bahan (dataset), Ekstraksi fitur dari setiap aksara Lampung dan Image Classification atau pengenalan aksara Lampung. Sistem keseluruhan dapat dilihat pada Gambar 3.

1.4. Convolutional Neural Network

Convolutional Neural Network (CNN) merupakan salah satu algoritma dari Deep Learning yang merupakan pengembangan dari Multilayer Perceptron (MLP) yang dirancang untuk mengolah data dua dimensi, misalnya gambar/citra atau suara. CNN digunakan untuk mengklasifikasi data yang terlabel dengan menggunakan metode supervised learning. Cara kerja supervised learning yaitu terdapat data yang dilatih dan terdapat variabel yang ditargetkan sehingga tujuan dari metode ini adalah mengelompokkan suatu data kepada yang sudah ada.

1.5. Google Colaboratory

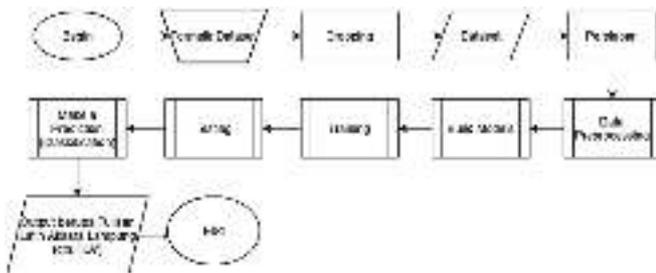
Google Colaboratory atau Google Colab adalah tools baru yang dikeluarkan oleh Google Internal Research yang dibuat guna membantu para Researcher dalam mengolah data untuk keperluan belajar maupun eksperimen pada pengolahan data khususnya pada bidang Machine Learning. Tools ini secara penggunaan hampir sama seperti Jupyter Notebook dan dibuat diatas environment Jupyter yang tidak memerlukan pengaturan atau setup terlebih dahulu sebelum digunakan dan berjalan sepenuhnya pada Cloud dengan memanfaatkan media penyimpanan Google Drive.

1.6. Tensorflow

Tensorflow merupakan perpustakaan perangkat lunak yang dikembangkan oleh Tim Google Brain dalam organisasi penelitian Mesin Cerdas Google, dengan tujuan melakukan pembelajaran mesin dan penelitian jaringan syaraf tiruan. Tensorflow merupakan pengembangan aljabar komputasi teknik pengoptimalan kompilasi. Mempermudah penghitungan dengan banyak ekspresi matematis dimana masalahnya adalah waktu yang dibutuhkan untuk melakukan perhitungan.

1.7. Python

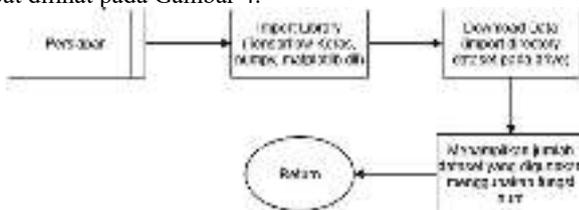
Python adalah salah satu dari bahasa pemrograman tingkat tinggi yang bersifat interpreter, interaktif, object-oriented dan dapat beroperasi di hampir semua platform, seperti keluarga Linux, Windows, Mac, dan platform lainnya. Python adalah salah satu bahasa pemrograman tingkat tinggi yang mudah dipelajari karena sintaks yang jelas dan elegan, yang dikombinasikan dengan penggunaan modul-modul yang mempunyai struktur data tingkat tinggi, efisien, dan siap langsung digunakan.



Gambar 3. Flowchart Tahapan Penelitian

2.1 Persiapan

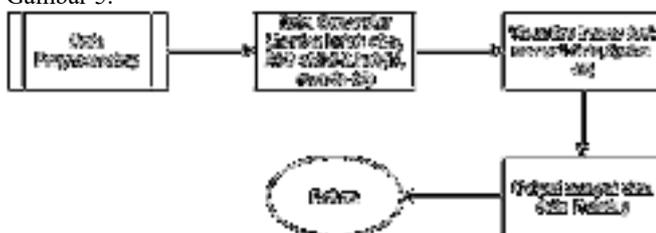
Pada proses ini dilakukan 2 tahapan yaitu Import Package dan Download Data guna untuk mendeklarasikan library yang digunakan dan menghubungkan dataset yang diunggah pada google drive. Adapun jumlah dataset yang digunakan yaitu 3000 gambar/citra yang terdiri dari 20 kelas. Garis besar proses ini dapat dilihat pada Gambar 4.



Gambar 4. Flowchart Persiapan

2.2 Data Preprocessing

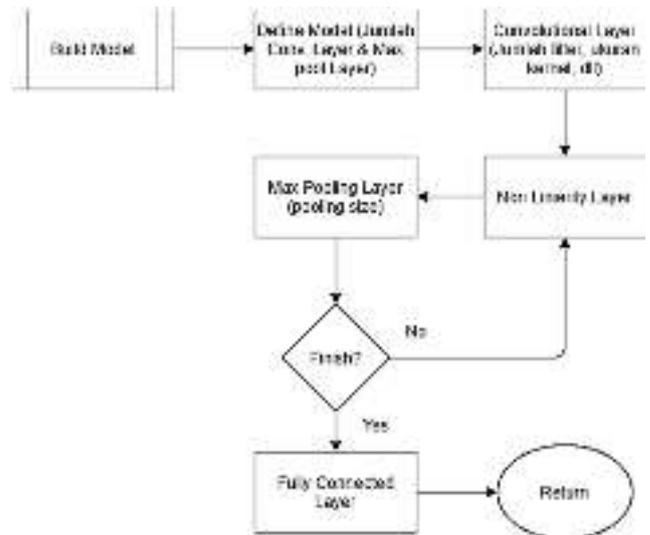
Pada proses ini dilakukan 2 tahapan yaitu Data Generator dan Visualize Image guna untuk mendeklarasikan ukuran batch size=100, ukuran citra 180x180, rescale citra 1/225 dan memvisualisasikan citra. Garis besar proses ini dapat dilihat pada Gambar 5.



Gambar 5. Flowchart Data Preprocessing

2.3 Build Models

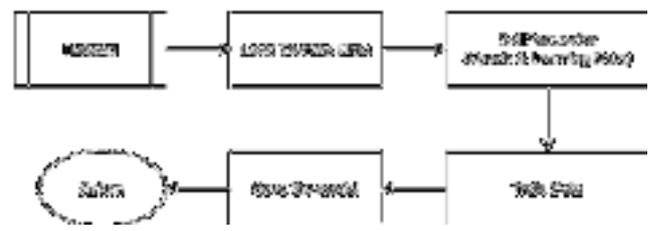
Proses ini adalah proses untuk mendeklarasikan semua variabel untuk model yang akan digunakan pada training, testing, dan klasifikasi. Dimana pada sistem ini menggunakan 12 layer yang terdiri dari Conv. Layer, Max pooling layer 2D atau 2x2, flatten layer, dan fully connected layer dengan kernel size 3x3 dan fungsi aktivasi yang digunakan yaitu ReLU. Garis besar proses ini dapat di lihat pada Gambar 6.



Gambar 6. Flowchart Build Models

2.4 Training

Proses training ini adalah proses untuk melatih model yang telah dibuat sebelumnya dengan menggunakan sejumlah data training yang telah disiapkan. Dengan epoch sebanyak 10 dan learning rate 0.0001. Secara garis besar proses training dapat digambarkan seperti flowchart pada Gambar 7.



Gambar 7. Flowchart Training

2.5 Testing

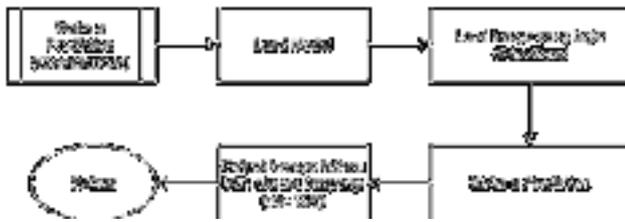
Proses testing ini adalah proses untuk menguji / meng-evaluasi model yang telah di-training dengan menggunakan sejumlah data yang telah disiapkan. Secara garis besar proses testing dapat digambarkan seperti flowchart pada Gambar 8.



Gambar 8. Flowchart Testing

2.6 Image Classification

Proses image classification ini adalah proses untuk mengklasifikasi data image baru menggunakan model yang telah di-training. Secara garis besar proses image classification dapat digambarkan seperti flowchart pada Gambar 9.



Gambar 9. Flowchart Image Classification

2.7 Save Model

Pada proses ini dilakukan proses save model yaitu menyimpan model setelah dilakukan proses training. Tujuannya yaitu untuk meminimalisir waktu yang digunakan untuk pelatihan atau training yang memakan waktu cukup lama bisa sampai berjam-jam, sehari-hari atau bahkan berminggu-minggu tergantung dari batch size dan epoch yang digunakan. Dengan dilakukannya save model memungkinkan untuk membagikan file hasil save model untuk pengembangan proyek baru seperti android dan embedded system dengan ukuran file save model/weight yaitu 211 KB.

3. Hasil dan pembahasan

3.1 Hasil Training

Pada Tabel 1 adalah hasil dari pelatihan data train dan data validation menggunakan epoch sebanyak 10. Berdasarkan tabel dapat diketahui bahwa epoch menghasilkan nilai accuracy dan nilai loss untuk data train dan data validation. Nilai akurasi merupakan suatu nilai yang digunakan untuk mengetahui tingkat keberhasilan model yang telah dibuat. Dan nilai loss adalah suatu ukuran dari sebuah error yang dibuat oleh network, dan tujuannya adalah untuk meminimalisirnya.

Tabel 1 Hasil Training Data Train dan Validation

| Epoch | Data Training | | Data Validation | |
|-------|---------------|--------|-----------------|--------|
| | Acc | Loss | Acc | Loss |
| 1 | 0.1043 | 2.9185 | 0.5471 | 1.4939 |
| 2 | 0.6078 | 1.3204 | 0.8157 | 0.6225 |
| 3 | 0.8352 | 0.5585 | 0.9171 | 0.3094 |
| 4 | 0.9437 | 0.1926 | 0.9457 | 0.2094 |
| 5 | 0.9788 | 0.0796 | 0.9457 | 0.2588 |
| 6 | 0.9774 | 0.0540 | 0.9629 | 0.1776 |
| 7 | 0.9917 | 0.0317 | 0.9571 | 0.1859 |
| 8 | 0.9938 | 0.0225 | 0.9700 | 0.1651 |
| 9 | 0.9947 | 0.0188 | 0.9471 | 0.2397 |
| 10 | 0.9941 | 0.0237 | 0.9600 | 0.2079 |

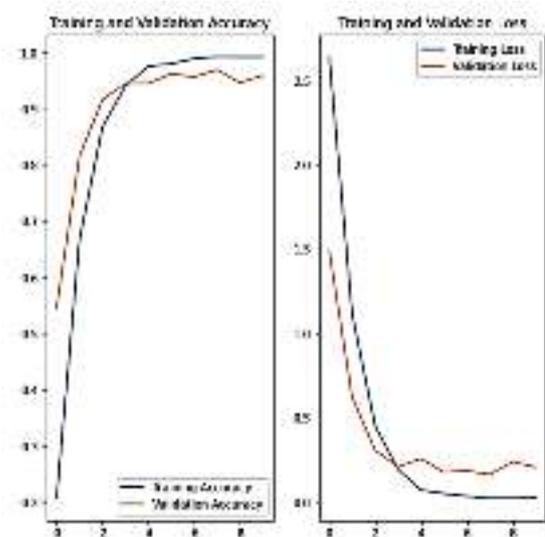
Gambar 10 menunjukkan grafik pergerakan nilai akurasi dan nilai loss untuk data train dan data validation yang dihasilkan pada setiap epoch. Berdasarkan gambar 10, untuk garis berwarna biru menunjukkan pergerakan untuk data train, sedangkan untuk garis berwarna orange menunjukkan pergerakan nilai akurasi dan nilai loss untuk data validation. Pada grafik disebelah kanan menunjukkan nilai loss untuk kedua data, dapat dilihat bahwa untuk data train, nilai loss yang didapatkan terus turun hingga

epoch ke-10, dan nilai loss untuk data validation yang awalnya turun kemudian pada sebuah titik nilai loss naik dan kembali turun hingga epoch ke-10.

Sedangkan pada grafik disebelah kiri menunjukkan nilai akurasi dari kedua data. Berdasarkan gambar IV.1 dapat dilihat bahwa nilai akurasi hasil pelatihan yang dihasilkan baik untuk data train maupun data validation sama-sama naik hingga epoch ke-10. Untuk data train, pada akhir epoch didapatkan nilai akurasi sebesar 0,9941. Sedangkan untuk data validation nilai akurasi pada epoch terakhir yang didapatkan yaitu sebesar 0,9600.

Cara penulisan tabel adalah seperti pada Tabel 1. Semua tabel diletakkan dapat diletakkan dalam teks, dengan *caption* (judul) di atasnya (center); sedangkan gambar pada bagian bawah halaman dengan *caption* di bawahnya (rata kiri/kanan). Usahakan tabel/gambar berukuran cukup, bila berwarna (minimal 300 DPI), tanpa garis tepi. Gambar dan tabel harus diletakkan di bagian terdekat dari teks dan wajib dijelaskan seperti pada Gambar 1 dan 2. *Caption* tabel dan gambar ditulis dengan Dutch801 Rm BT 9.5 point.

Persamaan ditulis dengan tab seperti contoh. Size font pada equation adalah 9.5 dengan tulisan dan angka menggunakan font Dutch801 Rm BT.



Gambar 10. Grafik Hasil Pelatihan

3.2 Hasil Prediksi

Tahap ini bertujuan untuk menguji sistem atau model yang telah dibuat yaitu untuk image classification atau mengelompokkan huruf induk aksara Lampung yang telah diekstraksi pada tahap sebelumnya ke dalam kelas-kelas yang sesuai. Dengan output berupa tulisan latin aksara Lampung (contoh "Ka").

Hasil prediksi: Gha

Gambar 10. Hasil Prediksi

3.3 Pengujian

Berdasarkan Tabel 2 pengujian peforma pada huruf induk aksara Lampung, dilakukan uji coba sebanyak 20 kali dengan masing-masing data uji yaitu satu untuk setiap kelasnya setelah dilakukan model pembelajaran. Didapatkan hasil yaitu TP (true positive) bernilai 20 yang berarti bahwa 20 data uji berhasil dideteksi dengan benar. FP (false positive) bernilai 0, FN (false negative) bernilai 0 dan TN (true negative) bernilai 20 yang berarti bahwa sebanyak 20 nilai berhasil dideteksi secara benar.

Tabel 2 Pengujian Peforma pada Huruf Induk Aksara Lampung

| | | | |
|--------|-------------------|-------------------|-------------------|
| Aktual | Prediksi | | |
| | | Aksara : Benar | Aksara : Salah |
| | Aksara : Benar | TP : 20 | FN : 0 |
| | Aksara : Salah | FP : 0 | TN : 20 |

Perhitungan :

$$Akurasi = \frac{TP+TN}{TP+FN+FP+TN} = \frac{20+20}{20+0+0+20} = 1 = 100\%$$

$$Recall = \frac{TP}{TP+FN} = \frac{20}{20+0} = 1 = 100\%$$

$$Precision = \frac{TP}{TP+FP} = \frac{20}{20+0} = 1 = 100\%$$

Berdasarkan hasil perhitungan menggunakan rumus, didapatkan akurasi sebesar 100%, recall 100% dan precision 100%. Dikarenakan pada saat uji coba sebanyak 20 sampel yang berbeda sistem mampu mendeteksi dengan benar secara keseluruhan.

Berdasarkan hasil perhitungan menggunakan rumus, didapatkan akurasi sebesar 100%. Hal ini dikarenakan pola yang digunakan sangat sederhana karena aksara Lampung memiliki pola yang sederhana dan bentuk yang hampir sama hanya yang membedakan adalah rotasinya.

4. Kesimpulan

Berdasarkan hasil pengujian, dapat ditarik beberapa kesimpulan yaitu, Berdasarkan hasil perhitungan menggunakan confusion matrix, didapatkan akurasi sebesar 100%. Hal ini dikarenakan pola yang digunakan sangat sederhana karena aksara Lampung memiliki pola yang sederhana dan bentuk yang hampir sama hanya yang membedakan adalah rotasinya. Pada proses training untuk epoch sebanyak 10 dibutuhkan waktu sebanyak 25 menit 51 detik atau kurang lebih 26 menit dengan memori yang terpakai sebesar 0.95 GB. Sistem “Pengenalan Aksara Lampung Menggunakan Metode CNN (Convolutional Neural Network) yang berjalan pada tools Google Colab sudah dilakukan proses Save Model dimana bertujuan untuk meminimalisir waktu untuk training ulang dan hasil save model berupa file bisa digunakan untuk pengembangan proyek baru seperti android atau embedded system dengan ukuran file yaitu 211 KB.

Ucapan terima kasih

Penulis berterima kasih kepada Ibu Yessi Mulyani, S.T., M.T. dan Bapak Ir. Meizano Ardhi Muhammad, S.T., M.T. selaku dosen pembimbing yang telah memberikan bimbingan dan masukan dalam pengerjaan makalah.

Daftar pustaka

Christopher Albert Lorentius et al, "Pengenalan Aksara Jawa dengan Menggunakan Metode Convolutional Neural Network", Program Studi Teknik Informatika, Universitas Kristen Petra, 2019.

Muhammad Dahria, "Kecerdasan Buatan (Artificial Intelligence)", Jurnal Sistem dan Teknologi Informasi Indonesia, Vol.5, No. 2, Pp.185, Agustus 2008.

Putri Navia Rena, "Penerapan Metode Convolutional Neural Network Pada Pendeteksi Gambar Notasi Balok", Program Studi Teknik Informatika, Universitas Islam Negeri Syarif Hidayatullah, Jakarta, 2019.

Jimmy Pujoseno, "Implementasi Deep Learning Menggunakan Convolutional Neural Network Untuk Klasifikasi Alat Tulis", Jurusan Statistika, Universitas Islam Indonesia, Yogyakarta, 2018.

Christopher Albert Lorentius et al, "Pengenalan Aksara Jawa dengan Menggunakan Metode Convolutional Neural Network", Program Studi Teknik Informatika, Universitas Kristen Petra, 2019.

Sarirotul Ilahiyah and Agung Nilorigi, " Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network", Jurnal Sistem dan Teknologi Informasi Indonesia, Vol.3, No. 2, Pp.50, 2 Agustus 2018.

Kris Sivam, "Rancang Bangun Alat Identifikasi Jenis Daging Dengan Pegolahan Citra Digital Menggunakan Python 2.7 dan OpenCV Berbasis Raspberry Pi 3", Jurusan Teknik Elektro, Universitas Lampung, 2018.

Syarifah Rosita Dewi, "Deep Learning Object Detection Pada Video Menggunakan Tensorflow dan Convolution Neural Network", Program Studi Statistika, Universitas Islam Indonesia, Yogyakarta, 2018.

Arfian, "Implementasi Convolution Neural Network Terhadap Transportasi Tradisional Menggunakan Keras", Program Studi Statistika, Universitas Islam Indonesia, Yogyakarta, 2018.

Benny Parrani, "Pengantar Google Colaboratory", Academy Digital Talent Scholarship, Universitas Mataram, 2019.

Rommy Oktaviandi et al, "Pemanfaatan Layanan Google Drive Untuk Menjalankan Aplikasi Pengontrolan Peralatan Listrik Menggunakan USB COMM Port", Jom FTEKNIK, Vol.2, No.2, Oktober 2015.

Muhammad Rifqi Mukhlas, "Perancangan Kampanye Pelestarian Aksara Lampung", Program Studi Desain Komunikasi Visual, Universitas Telkom, 2015.

Tri Lestari, "Pengenalan Karakter Tulisan Tangan Aksara Lampung Menggunakan Ekstraksi Fitur Chain Code 8 Direction Dan Projection Profile", Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung, 2018.

Christopher Albert Lorentius et al, "Pengenalan Aksara Jawa dengan Menggunakan Metode Convolutional Neural Network", Program Studi Teknik Informatika, Universitas Kristen Petra, 2019.

Cahyo Prarian et al, "Analisis Pengenalan Aksara Lampung Menggunakan Modified Direction Feature (MDF) Dan Hidden Markov Model", Teknik Telekomunikasi, Universitas Telkom, 2013.

Adhika Aryantio and Rinaldi Munir, "Pengenalan Aksara Lampung Menggunakan Jaringan Syaraf Tiruan", School of Electrical Engineering and Informatics, Institute Technology of Bandung, 2015.

Muath Alali et al, "Narrow Convolutional Neural Network for Arabic Dialects Polarity Classification", Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 2019.

I Wayan Suartika E. P et al, "Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101", Jurnal Teknik ITS, Vol.5, No. 1, Pp.A65, 2016.

Siwi Prihatiningsih et al, "Analisa Performa Pengenalan Tulisan Tangan Angka Berdasarkan Jumlah Iterasi Menggunakan Metode Convolutional Neural Network", Jurnal Ilmiah Teknologi dan Rekayasa, Vol.24, No. 1, Pp.58, 1 April 2019.