

KLASIFIKASI AWAN DENGAN CONVOLUTIONAL NEURAL NETWORK

Bagus Nuryasin^{1, a)}, Riser Fahdiran²

^{1,2}*Program Studi Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta, Jl. Rawamangun Muka, Jakarta 13220, Indonesia*

Email: ^{a)}nuryasinbagus@gmail.com

ABSTRAK

Seperti yang diketahui daerah Jakarta ialah daerah yang cukup rawan banjir. Untuk mengantisipasinya maka salah satu hal yang diperlukan adalah mengembangkan model prediksi cuaca dengan pendekatan klasifikasi jenis awan menggunakan Convolutional Neural Network. Pada penelitian ini, kami mengumpulkan dan memproses data citra 10 jenis awan yang akan digunakan untuk pelatihan model. Tabel klasifikasi dan tabel pengukuran hasil klasifikasi digunakan untuk mengukur akurasi klasifikasi model. Hasil klasifikasi model mencapai akurasi hampir sempurna untuk semua jenis awan dengan akurasi sebesar 95%. Adapun sejumlah kecil kesalahan dalam klasifikasi di beberapa jenis awan karena kemiripan bentuk dan tekstur.

PENDAHULUAN

Seperti yang diketahui daerah Jakarta ialah daerah yang cukup rawan banjir. Dalam hitungan jam, hujan deras yang mengguyur di tempat tertentu menyebabkan banjir dimana-mana. Untuk mengantisipasinya maka salah satu hal yang diperlukan adalah mengembangkan model prediksi cuaca yang cepat dan tepat. Awan merupakan indikator utama dalam menentukan keadaan cuaca dan masing-masing jenis awan mempunyai arti yang berbeda. Informasi tentang jenis awan sangat penting untuk mengamati dan memprediksi keadaan cuaca. Untuk mendapatkan informasi tentang awan biasanya menggunakan citra satelit dan citra radar stasiun cuaca.

Menariknya mengembangkan model prediksi cuaca merupakan topik yang sangat kompleks yang membutuhkan pemahaman mendalam tentang Fisika, Matematika dan di era modern, komputasi juga. Beberapa pendekatan prediksi mulai dari pendekatan yang sifatnya deterministik sampai dengan pendekatan secara stokastik. Beberapa tahun terakhir, pendekatan prediksi cuaca dengan jaringan syaraf buatan berhasil diaplikasikan seperti reduksi dimensi dan prediksi jangka pendek akan tetapi membutuhkan ekstraksi fitur. Dengan menggunakan Convolutional Neural Network yang mampu meniru sistem pengenalan citra seperti visual cortex manusia sehingga memiliki kemampuan mengolah informasi citra tanpa mengekstraksi fitur. Di dalam eksperimen ini Convolutional Neural Network digunakan untuk mengklasifikasi objek citra jenis awan. Difokuskan pada pembuatan model dalam melakukan klasifikasi objek jenis awan.

LANDASAN TEORI

Awan

Awan adalah sekumpulan tetes yang mempunyai konsentrasi berorde 100 per cm^3 dan mempunyai jejari sekitar 10 μm . Tetes hujan akan tumbuh jika populasi awan menjadi tidak stabil. Pertama tumbukan langsung dan penangkapan tetes-tetes air. Kedua interaksi antara tetes air dan kristal es yang terbatas pada awan yang puncaknya di atas paras 0°C . Jika sebuah kristal es berada bersama sejumlah tetes air kelewat dingin, maka situasi menjadi tidak stabil. Keseimbangan tekanan uap di atas es lebih kecil daripada di atas air pada temperatur yang sama, karenanya kristal es tumbuh dengan difusi uap dan tetes menguap untuk mengimbangnya.

Genus Awan

Tiap genus (golongan utama) awan dibagi menjadi jenis awan, dan tiap jenis awan dibagi lagi menjadi varietas awan. Awan dapat digolongkan menjadi sepuluh genus yaitu Cirrus (Ci), Cirrocumulus (Cc), Cirrostratus (Cs), Altocumulus (Ac), Altostratus (As), Nimbostratus (Ns), Stratocumulus (Sc), Stratus (St), Cumulus (Cu), dan Cumulonimbus (Cb),

Citra

Citra bisa didefinisikan sebagai suatu fungsi dua dimensi, $f(x,y)$, dimana x dan y adalah koordinat spasial, dan $f(x,y)$ adalah nilai pada koordinat (x,y) yang sering disebut intensitas. (Gonzales dan Wood, 2008). Citra digital adalah citra $f(x,y)$ yang telah digitalisasi baik dari segi koordinat area maupun pada nilai intensitasnya. Sebuah citra digital terdiri dari sejumlah elemen yang disebut picture element atau pixel. Setiap nilai pixel diwakili oleh dua buah bilangan bulat untuk menunjukkan lokasi dalam bidang citra.

Gambar menunjukkan citra dipandang sebagai dua dimensi dengan jumlah kolom N dan jumlah baris M . Masing-masing dari nilai intensitas *pixel* dianggap sebagai elemen matriks dua dimensi. Banyaknya cara mempresentasikan citra berwarna diantaranya dengan format RGB (red, green, blue). Pada RGB, citra berwarna dinyatakan sebagai gabungan dari tiga buah citra monokromatik merah, hijau, dan biru berskala sama.

Praproses citra

Praproses citra dilakukan untuk menyiapkan citra untuk diproses lebih lanjut, baik untuk kebutuhan ekstraksi fitur maupun kebutuhan klasifikasi. Beberapa teknik preprocessing yang umum digunakan antara lain penghilang noise dengan filter spasial, segmentasi, cropping atau pemotongan region of interest dan perubah ukuran citra

Convolutional Neural Network

Convolutional Neural Network merupakan jaringan syaraf yang dikhususkan untuk memproses data yang memiliki grid. Sebagai contoh adalah berupa citra dua dimensi. Nama konvolusi sendiri merupakan operasi aljabar linear yang mengkalikan matriks dari filter pada citra yang akan di proses. Model jaringan seperti ini sudah terbukti sangat ampuh dalam menangani permasalahan klasifikasi citra.

Operasi konvolusi

Operasi konvolusi adalah operasi pada dua fungsi argument bernilai nyata (Goodfellow, Bengio, dan Courville, 2016). Operasi ini menerapkan fungsi output sebagai Feature Map dari input citra. Input dan output ini dapat dilihat sebagai dua argumen bernilai riil. Secara formal operasi konvolusi dapat ditulis dengan rumus berikut.

$$s(t) = (x * w)(t)$$

Fungsi $s(t)$ memberikan output tunggal berupa Feature Map, argument pertama adalah input yang merupakan x dan argument kedua w sebagai kernel atau filter. Jika kita melihat input sebagai citra dua dimensi, maka kita bisa mengasumsikan t sebagai pixel dan menggantinya dengan i dan j . Operasi untuk konvolusi ke input dengan lebih dari satu dimensi dapat ditulis sebagai berikut.

$$S(i, j) = (K * I)(i, j) = \sum_m \sum_n I(i - m, j - n) K(m, n)$$

Persamaan diatas adalah perhitungan dasar dalam operasi operasi konvolusi dimana i dan j adalah piksel dari citra. Perhitungannya bersifat komutatif dan muncul saat K sebagai kernel, I sebagai input dan kernel yang dapat dibalik relative terhadap input. Sebagai alternatif, operasi konvolusi dapat dilihat sebagai perkalian matriks antaran citra masukan dan kernel dimana keluarannya dapat dihitung dengan dot product.

0	1	1	1	0	0	0
0	0	1	1	1	0	0
0	0	0	1	1	1	0
0	0	0	1	1	0	0
0	0	1	1	0	0	0
0	1	1	0	0	0	0
1	1	0	0	0	0	0

I

1	0	1
0	1	0
1	0	1

K

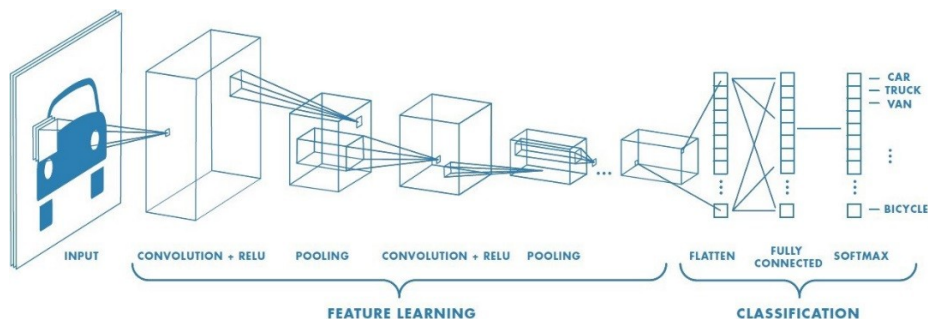
*

1	4	3	4	1
1	2	4	3	3
1	2	3	4	1
1	3	3	1	1
3	3	1	1	0

I * K

Arsitektur Jaringan CNN

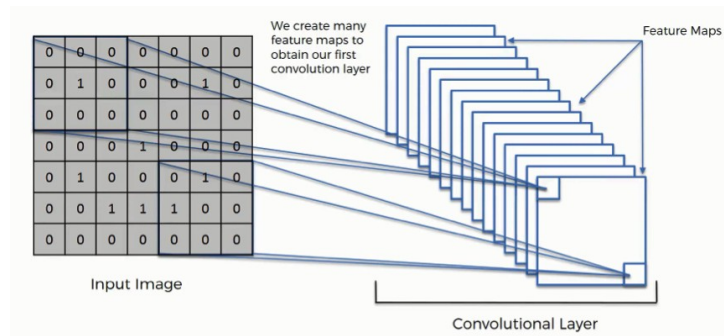
Arsitektur dari CNN juga terbagi menjadi 2 bagian besar yaitu Feature Extraction Layer dan Fully-Connected Layer (FC). Seperti pada gambar berikut.



Lapisan Ekstraksi Fitur (Feature Extraction Layer)

Proses pada bagian ini adalah melakukan “Encoding” dari sebuah gambar menjadi fitur yang berupa angka-angka yang mempresentasikan gambar tersebut. Lapisan Ekstraksi Fitur terdiri dari dua bagian, Lapisan Konvolusional dan Lapisan Penggabungan.

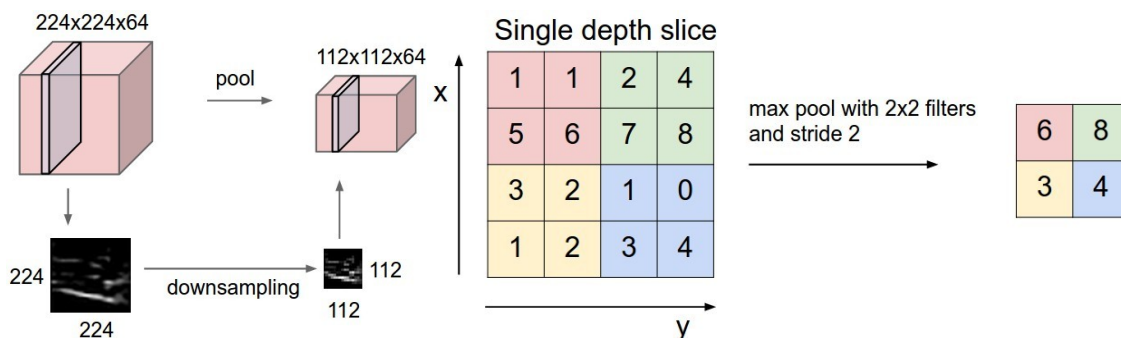
Lapisan Konvolusional (Convolutional Layer)



Lapisan Konvolusional merupakan lapisan dimana operasi konvolusi antara input dan nilai dari filter tersebut sehingga menghasilkan sebuah output atau biasa disebut sebagai Peta Aktivasi (Activation Map) atau Peta Fitur (Feature Map). Saringan (Filter) akan digeser ke seluruh bagian dari gambar. Proses dari Convolutional layer seperti pada gambar berikut.

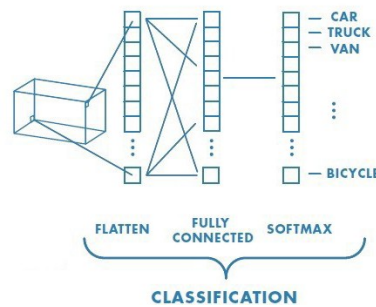
Lapisan Penggabungan (Pooling Layer)

Lapisan Penggabungan biasanya berada setelah Lapisan Konvolusional. Pada prinsipnya Lapisan Penggabungan terdiri dari sebuah filter dengan ukuran tertentu yang akan bergeser pada seluruh area Peta Fitur. Penggabungan yang biasa digunakan adalah Penggabungan nilai maksimum (Max Pooling). Max pooling membagi output dari convolution layer menjadi beberapa grid kecil lalu mengambil nilai maksimal dari setiap grid untuk menyusun matriks citra yang telah direduksi seperti yang ditunjukkan pada Gambar .



Tujuan dari penggunaan pooling layer adalah mengurangi dimensi dari feature map sehingga mempercepat komputasi.

Lapisan Terhubung Seluruhnya (Fully Connected Layer)



Lapisan Terhubung Seluruhnya adalah lapisan yang biasanya digunakan dalam penerapan Lapisan simpul (neuron) banyak yang terhubung (Multilayer Perceptron) dan bertujuan untuk melakukan transformasi pada dimensi data agar data dapat diklasifikasikan secara linear. Setiap simpul pada lapisan konvolusional perlu ditransformasi menjadi data satu dimensi terlebih dahulu sebelum dapat dimasukkan ke dalam sebuah Lapisan Terhubung Seluruhnya. Karena hal tersebut menyebabkan data kehilangan informasi spasialnya dan tidak reversibel, Lapisan Terhubung Seluruhnya hanya dapat diimplementasikan di akhir jaringan.

Parameter yang mempengaruhi CNN (Hyperparameter)

Kecepatan Pembelajaran (Learning Rate)

Skalar yang digunakan untuk melatih model melalui penurunan gradien. Selama setiap iterasi, algoritme penurunan gradien mengalikan kecepatan pembelajaran dengan gradien. Produk yang dihasilkan disebut langkah gradien.

Iterasi Pelatihan (Epoch)

Iterasi Pelatihan adalah ketika seluruh dataset sudah melalui proses pelatihan model sampai dikembalikan ke awal untuk sekali putaran, karena satu Iterasi Pelatihan terlalu besar untuk dimasukkan (feeding) kedalam model maka dari itu kita perlu membaginya ke dalam satuan kecil (batches).

Ukuran Tumpukan (Batch Size)

Ukuran Tumpukan adalah jumlah sampel data yang disembarkan ke model.

Penelitian yang relevan

[1] pada penelitian Klasifikasi Citra Menggunakan *Convolutional Neural Network* (CNN) pada *Caltech 101* melakukan klasifikasi citra menggunakan *feedforward* dan selanjutnya dilakukan pembelajaran metode *backpropagation*. Penelitian menggunakan 5 kategori unggas, yaitu : *Emu*, *Flamingo*, *Ibis*, *Pigeon* dan *Rooster* yang terdiri dari 150 citra. Selain itu kategori lainnya adalah

Cougar, *Crocodile* dan *Face*. Klasifikasi citra yang dilakukan menghasilkan presentase kebenaran antara 20% sampai 50%.

[2] pada penelitian CloudNet: Ground-Based Cloud Classification With Deep Convolutional Neural Network melakukan klasifikasi awan menggunakan *Convolutional Neural Network* (CNN). Penelitian menggunakan 11 kategori jenis awan pada fitur tekstur, struktur, dan bentuk secara bersamaan.

METODE

Pengumpulan Data

Pengumpulan data dilakukan dengan menggunakan data sekunder yang digunakan [2] untuk mengklasifikasi jenis awan. Datanya berupa data citra sepuluh jenis awan.

Data latih digunakan algoritma klasifikasi untuk membentuk sebuah model klasifikasi, model ini merupakan representasi pengetahuan yang akan digunakan untuk mengklasifikasi kelas data yang ada, semakin besar data latih yang akan digunakan, maka akan semakin baik model dalam memahami pola data. Data uji yang ada digunakan untuk mengetahui sejauh mana klasifikasi berhasil melakukan klasifikasi dengan tepat. Perbandingan jumlah data latih, data validasi, dan data uji dapat dilihat pada tabel untuk data citra yang digunakan.

Data Citra	Data Latih	Data Valid	Jumlah
Alto cumulus (Ac)	176	45	221
Alto stratus (As)	150	38	188
Cumulonimbus (Cb)	193	49	242
Cirrocumulus (Cc)	214	54	268
Cirrus (Ci)	111	28	139
Cirrostratus (Cs)	229	58	287
Cumulus (Cu)	146	37	183
Nimbostratus (Ns)	219	55	274
Stratocumulus (Sc)	272	68	340
Stratus (St)	161	41	202
Jumlah	1871	473	2344

Praproses data

Pada praproses data, citra dimuatkan pada wadah sementara kemudian citra diperbanyak dan dimodifikasi untuk pelatihan dan validasi model CNN. Di library Fastai, pemuatan citra pada wadah sementara dapat menggunakan script `ImageDataBunch.from_folder()` dan modifikasi citra dapat menggunakan script `get_transform()`. Script `get_transform()` memodifikasi citra diantaranya :

- Mengubah resolusi citra menjadi 224×224
- membalik citra secara horizontal
- memutar citra sebesar 10°
- memperbesar citra sebesar 10% dari citra aslinya
- mengubah kecerahan dan kontras sebesar 20% dari citra aslinya
- menggeser citra sebesar 20% dari lokasi citra aslinya

Pelatihan Model

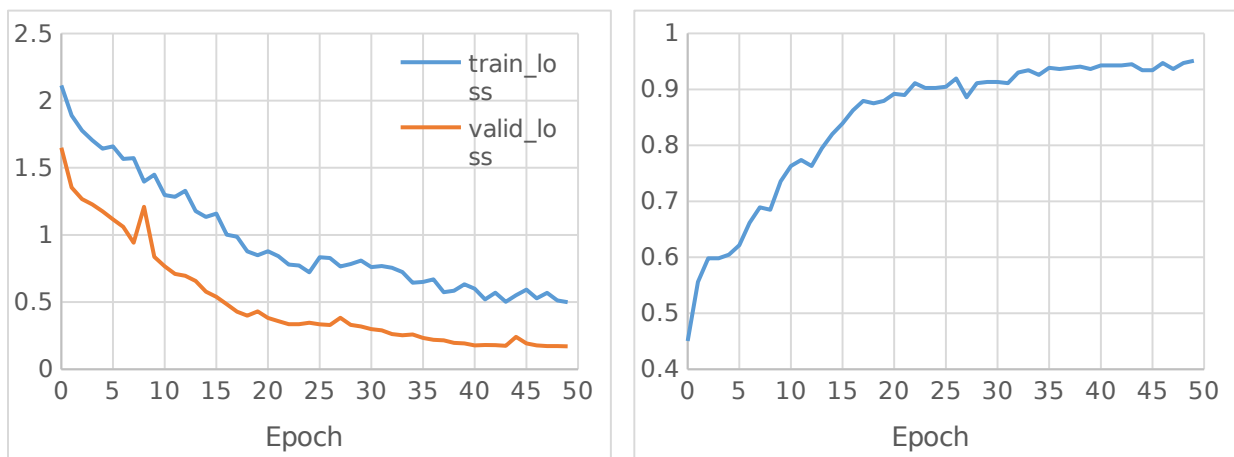
Pada pelatihan model, inisiasi model dilakukan terlebih dahulu. Pelatihan model menggunakan data latih dan validasi. Dalam eksperimen ini, kami menggunakan 25 epoch pertama dengan learning rate 3×10^{-3} dan 25 epoch kedua dengan learning rate dari 1×10^{-4} sampai 1×10^{-5} , serta batch size sebesar 8. Dari pelatihan model akan menghasilkan nilai akurasi (accuracy) dan kesalahan klasifikasi (loss). Nilai loss terbagi menjadi dua yaitu, train_loss (kesalahan klasifikasi data latih) dan valid_loss (kesalahan klasifikasi data latih).

ANALISA DAN PEMBAHASAN

Hasil Pelatihan Model

Hasil Pelatihan Model berupa nilai akurasi, train_loss, dan valid_loss.

Berikut hasil plot nilai akurasi dan loss.



Seperti yang terlihat pada dua gambar diatas plot yang dihasilkan dari pelatihan model, nilai train_loss dan valid_loss semakin menurun untuk setiap epoch-nya menunjukkan model berusaha memahami pola data citra. Puncak-puncak pada garis nilai loss menunjukkan kesulitan pola dalam pelatihan. Nilai akurasi model juga semakin meningkat untuk setiap epoch-nya, adapun lembah-lembah pada garis nilai akurasi juga menunjukkan kesulitan pola dalam pelatihan.

Validasi Model

Untuk mengevaluasi model maka diperlukan validasi model, biasanya validasi model menggunakan data validasi. Untuk melihat hasil evaluasi model biasanya membutuhkan tabel pengukuran hasil klasifikasi dan tabel klasifikasi.

Berikut tabel klasifikasi setiap jenis awan. Pada tabel klasifikasi dapat dilihat sejumlah kecil kesalahan klasifikasi model mendekati label asli dalam kriteria bentuk dan tekstur. Sebagai contoh kesalahan klasifikasi Cs yang seharusnya Cc karena kemiripan bentuk dan formasi awan dan kesalahan klasifikasi Ns yang seharusnya Sc karena kemiripan keabuan awan.

PENUTUP

Kesimpulan

Berdasarkan hasil dari analisis yang telah dilakukan, maka peneliti dapat menarik beberapa kesimpulan seperti berikut:

- Metode Convolutional Neural Network (CNN) berhasil diimplementasikan untuk klasifikasi 10 jenis awan dengan menggunakan library fastai.
- Dengan menggunakan 25 epoch pertama dengan learning rate 3×10^{-3} dan 25 epoch kedua dengan learning rate dari 1×10^{-4} sampai 1×10^{-5} , serta batch size sebesar 8. Pada tabel pengukuran hasil klasifikasi, didapatkan klasifikasi jenis awan dengan model CNN ini mencapai akurasi hampir sempurna untuk semua jenis awan dengan akurasi sebesar 95%.
- Model mengalami sejumlah kecil kesalahan dalam klasifikasi jenis awan. Kesalahan klasifikasi model mendekati label asli karena kemiripan kriteria bentuk dan tekstur.

Saran

Berdasarkan hasil dari analisis dan kesimpulan, dapat diberikan saran sebagai berikut:

- Mengubah kegunaan klasifikasi sebagai deteksi. Jenis-jenis awan tertentu seperti cumulonimbus dapat memberi tanda akan terjadinya badai. Sehingga model CNN ini dapat lebih berguna untuk kepentingan umum.
- Memperhatikan arsitektur model CNN di Google Colaboratory dengan library fastai. Maksimal waktu pelatihan model yang diberikan secara gratis oleh Google ialah 12 jam. Gunakan fitur simpan untuk setiap epoch agar tidak mengulangi pelatihan model lagi.

DAFTAR PUSTAKA

- [1] I. W. S. E. P, A. Y. Wijaya dan R. Soelaiman, "Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101," *JURNAL TEKNIK ITS*, vol. 5, pp. A65-A69, 2016.
- [2] J. Zhang, P. Liu, F. Zhang dan Q. Song, "CloudNet: Ground-Based Cloud Classification With Deep Convolutional Neural Network," *Geophysical Research Letters*, p. 45, 2018.
- [3] K. P. Danukusumo, "Implementasi Deep Learning Menggunakan Convolutional Neural Network untuk Klasifikasi Citra Candi Berbasis GPU," Program Studi Teknologi Industri Universitas Atma Jaya, Yogyakarta, 2017.
- [4] R. C. Gonzalez dan R. E. Woods, Digital Image Processing, 3rd penyunt., United Kingdom: Prentice Hall, 2008.
- [5] I. Goodfellow, . Y. Bengio dan . A. Courvil, Deep Learning, Massachusetts: MIT Press, 2016.
- [6] B. . T. HK, Mikrofisika Awan dan Hujan, 2 penyunt., H. Widiyatmoko, R. Satyaningsih dan W. Fitria, Penyunt., Jakarta, Jakarta: Badan Meteorologi Klimatologi dan Geofisika, 2012.
- [7] A. Karpathy, "Convolutional Neural Networks (CNNs / ConvNets)," [Online]. Available: <http://cs231n.github.io>.
- [8] Rismayati, "Implementasi Convolutional Neural Network untuk Sortasi Mutu Salak Ekspor Berbasis Citra Digital.," Program Studi S2 Ilmu Komputer Departemen Ilmu Komputer dan Elektronika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Gajah Mada, Yogyakarta, 2016.
- [9] A. Handhoko, M. I. Zul dan F. Yuli, "Sistem Deteksi Kondisi Cuaca Berdasarkan Citra Awan Berbasis Pengolahan Citra Digital Menggunakan Algoritma k-Nearest Neighbor (k-NN)," *3rd Applied Business and Engineering Conference 2015 - Batam*, 2015.
- [10] Arfian, "Implementasi Convolutional Neural Network Terhadap Transportasi Tradisional Menggunakan Keras (Studi Kasus : Data Citra Transportasi Tradisional Andong, Becak dan Pedati)," Jurusan Statistika Fakultas Matematika Dan Ilmu Pengetahuan Alam Universitas Islam Indonesia, Yogyakarta, 2018.