

Penerapan Metode Deep Learning Menggunakan Algoritma CNN Dengan Arsitektur VGG NET Untuk Pengenalan Cuaca

Rhyosvaldo Aurellio Tilasefana¹, Ricky Eka Putra²

^{1,2}Jurusan Teknik Informatika Fakultas Teknik Universitas Negeri Surabaya

¹rhyosvaldo.17051204072@mhs.unesa.ac.id

²rickyeka@unesa.ac.id

Abstrak — Dengan berkembangnya teknologi, kini komputer dapat memberikan informasi tidak hanya dalam bentuk teks saja, melainkan juga bentuk *image*, suara, dan video. Perubahan cuaca ekstrim sering terjadi di Indonesia dapat mengganggu aktivitas sehari-hari. Sehingga menarik minat masyarakat akan kebutuhan informasi mengenai cuaca dan iklim. Oleh karena itu, perlu adanya perkiraan cuaca yang tepat. Tetapi, hingga kini pembuatan proses perkiraan cuaca dan iklim masih terbatas dari masukan data pemodelan. Ternyata didalam pemodelan tersebut masih terdapat kekurangan seperti penggunaan jumlah parameter, asumsi-asumsi matematis dan formulasi persamaan yang sangat rumit.

Dalam mengatasi hal tersebut, maka perlu mengembangkan kemampuan menganalisa dan menarik kesimpulan pada komputer. Kemampuan tersebut dikenal dengan deep learning yang memiliki cara kerja mirip jaringan saraf tiruan. Yang membedakan kalau jaringan syaraf tiruan hanya mengandalkan 1 layer, sedangkan deep learning lebih dari 1 layer semakin tinggi layer yang digunakan, learning komputer juga semakin lama.

Penelitian ini dilakukan dengan memakai algoritma CNN dari deep learning yang dirancang untuk mengolah data 2 dimensi, sehingga telah sukses mengidentifikasi citra cuaca dengan nilai tertinggi pada akurasi data latih sebesar 94,16% dan akurasi data testing sebesar 65,00% dengan menggunakan epoch sebanyak 150 epoch. Data citra yang digunakan masing-masing 90 untuk data latih dan 10 untuk data testing. Dapat disimpulkan bahwa training yang besar mampu mempengaruhi algoritma CNN, semakin besar data training maka semakin tinggi pula ketepatan yang didapatkan berdasarkan penelitian-penelitian terdahulu. Jumlah epoch yang dipakai dalam arsitektur CNN mempengaruhi hasil ketepatan, semakin banyak epoch maka semakin tinggi juga ketepatan yang didapatkan.

Kata Kunci - *Deep Learning*, CNN, VGG NET, Cuaca.

I. PENDAHULUAN

Dengan berkembangnya teknologi, kini komputer dapat melakukan banyak hal. Komputer dapat memberikan informasi yang tidak hanya dalam bentuk teks, tetapi juga dalam bentuk gambar, audio, dan video. Perubahan cuaca di Indonesia ada 2 yaitu musim penghujan dan musim kemarau 2 musim tersebut yang sering terjadi di Indonesia juga dapat mengganggu aktivitas sehari-hari. Sehingga kini masyarakat yang memiliki minat dan kebutuhan informasi mengenai cuaca dan iklim sangat tinggi. Menurut BMKG Indonesia menyatakan bahwa suhu udara di Indonesia pada bulan Agustus periode 1991-2020 adalah 26.46oC dan suhu udara di bulan Agustus 2022 adalah 26.81oC yang mana berdasarkan nilai tersebut menunjukkan anomali positif dengan nilai 0,35oC (BMKG). Anomali tersebut merupakan anomali tertinggi ke-5 sepanjang periode

pengamatan data sejak 1981. Dari data tersebut dapat disimpulkan bahwa tingkat akurasi dalam memprediksi anomali cuaca di Indonesia cukup rendah karena kondisi cuaca yang tidak menentu [2].

Prediksi cuaca dalam pelaksanaan proyek merupakan poin penting saat akan mengerjakan proyek. Sebagai contoh cuaca sangat mempengaruhi proyek pengerjaan tanah / pondasi. Pekerjaan ini dapat termasuk sebagai pekerjaan jangka pendek / juga bisa termasuk pekerjaan jangka panjang. Jika dijalankan di penutup musim penghujan atau pada masa pergantian, maka diharapkan pekerjaan ini dapat berjalan lebih lancar supaya dapat cepat terselesaikan. Pengerjaan tanah dan pondasi hampir tanpa adanya hambatan jika cuacanya mendukung yang dapat diselesaikan secara tepat waktu. Contoh lainnya yaitu pada proyek pengerjaan rumah. Jika ingin membuat tembok sebaiknya dilakukan sebelum tibanya musim hujan, karena kalau dilakukan pada musim penghujan, kualitas tembok menjadi kurang baik dan pengerjaannya akan berjalan lambat. Hal ini akan mengakibatkan tembok akan menjadi lembab karena curah hujan yang tinggi sehingga berpengaruh pada saat melakukan pengecatan. Oleh karena itu, agar kebutuhan masyarakat mengenai informasi cuaca dan iklim terpenuhi tentunya perlu adanya perkiraan cuaca yang tepat. Tetapi, sampai kini proses penciptaan perkiraan cuaca dan iklim masih cukup terbatas dari masukan data pemodelan. Dimana dalam pemodelan tersebut ternyata masih terdapat beberapa minus seperti pemakaian jumlah parameter, asumsi-asumsi matematis dan formulasi persamaan yang cenderung rumit [10].

Dalam mengatasi hal tersebut, pemanfaatan Algoritma pembelajaran untuk prediksi cuaca akan membantu mengatasi kurangnya akurasi dan efisiensi. Sebagai contoh dengan menggunakan Algoritma Deep Learning misalnya. Algoritma Deep Learning menggunakan Long Short-Term Memory (yang memiliki banyak lapisan tersembunyi) untuk mencapai hasil yang konsisten. Deep learning adalah bagian dari machine learning yang memiliki fungsi untuk mempelajari data yang telah tersedia melalui algoritma yang ada. Terdapat tiga pilihan yang dapat dilakukan dalam proses pembelajaran ini, yaitu secara terarah, semi terarah atau tak terarah. Yang pertama yaitu pembelajaran secara terarah mengizinkan algoritma untuk belajar berdasarkan data yang tersedia. Yang kedua pada pembelajaran semi terarah, algoritma belajar berdasarkan data yang telah tersedia walaupun tidak mumpuni. Kalau untuk pembelajaran yang tak terarah lebih mengandalkan algoritma untuk belajar tanpa ada masukan data satupun [6].

Contoh pendekatan yang telah sukses untuk mengklasifikasikan citra adalah engan memakai Jaringan Saraf

Tiruan (Artificial Neural Network/ANN). Contoh model ANN yang bisa diterapkan pada klasifikasi objek berupa citra adalah Convolutional Neural Network (CNN). CNN ialah operasi konvolusi yang nantinya akan menghubungkan beberapa lapisan pemrosesan dengan memakai beberapa elemen yang berjalan secara beriringan dan terinspirasi oleh sistem syaraf biologis [4].

CNN dirancang untuk mengolah data dua dimensi karena untuk merupakan pengembangan dari Multilayer Perceptron (MLP). Hal ini menjadikan CNN masuk dalam jenis Deep Neural Network dikarenakan banyak diaplikasikan pada data citra dan kedalaman jaringannya yang tinggi. Karena pada kasus klasifikasi citra, penggunaan MLP dinilai tidak cocok karena tak menyimpan informasi spesial dari data citra dan menganggap setiap piksel itu merupakan fitur yang independen sehingga menghasilkan hasil yang kurang bagus. CNN awal kali dikembangkan oleh Kuniko Fukushima, seorang peneliti dari NHK Broadcasting Science Research Laboratories, Kinuta, Setagaya, Tokyo, Jepang. Awalnya CNN bernama NeoCognitron. Lalu konsep CNN kemudian dimatangkan oleh Yann LeChun, seorang peneliti dari AT&T Bell Laboratories di Holmdel, New Jersey, Amerika Serikat. Kemudian model CNN yang bernama LeNet berhasil diterapkan oleh LeChun pada penelitiannya mengenai pengenalan angka dan tulisan tangan. Alex Krizhevsky di tahun 2012, melakukan penerapan CNN miliknya hingga berhasil menyabet juara pada kompetisi ImageNet Large Scale Visual Recognition Challenge 2012. Prestasi Alex pun menjadi ajang pembuktian kalau metode Deep Learning, khususnya CNN terbukti sukses mengalahkan metode lainnya dari Machine Learning seperti SVM pada kasus klasifikasi objek pada citra neuron yang saling berhubungan dan CNN memiliki arsitektur yang sama. CNN memiliki keunikan dari jaringan syaraf lain karena operasi konvolusional yang menerapkan filter pada setiap bagian dari input sebelumnya untuk mengekstraksi pola. Dengan menggunakan Convolutional Neural Network yang mampu meniru sistem pengenalan citra seperti visual cortex manusia sehingga memiliki kemampuan mengolah informasi citra tanpa mengekstraksi fitur [8].

Berdasarkan penelitian terdahulu yang pernah dilakukan oleh Randhy Sulistyio Budi et al (2021) yang berjudul "Klasifikasi Cuaca Menggunakan Metode Convolutional Neural Network (Cnn)" disimpulkan bahwa melakukan pemahaman kondisi cuaca menjadi hal yang penting di kehidupan manusia, sehingga perlu adanya sebuah sistem untuk membagi cuaca sesuai dengan kategorinya masing-masing agar informasi yang dihasilkan dapat digunakan untuk berbagai keperluan. Menggunakan Deep Learning dengan metode Convolutional Neural Network (CNN) yang berfungsi untuk mengolah data dua dimensi, yang dapat diimplementasikan menjadi data citra. Proses klasifikasi cuaca CNN terdiri dari tahap latihan dan uji coba untuk mengetahui hasil pembagiannya. Parameter kinerja yang diteliti berupa *accuracy*, *precision*, *recall*, dan F1-Score. Pada penelitian ini menghasilkan performa terbaik yaitu dengan penggunaan optimizer RMSprop, resolusi citra 224x224, learning rate

0,0001, dan epoch 50 dengan nilai performa sistem untuk *accuracy*, *precision*, *recall*, dan F1-score masing-masing sebesar 87%, 90%, 87%, dan 87%. Penelitian ini menggunakan arsitektur dari VGG16 yang digunakan karena lapisan konvolusinya yang sederhana, yaitu hanya menggunakan lapisan konvolusi. VGG16 atau yang bisa juga dikenali sebagai oxfordNet merupakan jaringan saraf convolutional. VGG16 digunakan untuk mengekstrak jalan dari citra penginderaan jauh dan CNN sebagai pengklasifikasi serta memberikan akurasi. VGG16 lebih utama dalam melayani tujuan klasifikasi dan melakukan deteksi pada objek [7].

Berdasarkan penjelasan di atas, maka peneliti akan melakukan penelitian menggunakan metode CNN yang akan digunakan untuk mengklasifikasikan jenis cuaca. Ada 4 class cuaca yang akan digunakan yaitu cloudy, rain, shine dan sunrise, dengan masing masing data set cloudy 243 data, rain 529 data, sunrise 276, shine 238 dengan total 1286 data training dan 250 data testing. Tujuan dari penelitian ini adalah untuk memprediksi cuaca serta untuk mengetahui tingkat akurasi dari hasil klasifikasi menggunakan metode deep learning dengan algoritma CNN dan K-Fold Validation.

Penulis juga telah melakukan uji coba dengan menggunakan citra cuaca sebanyak 100 epoch didapatkan hasil akurasi tertinggi sebesar 92,48% dengan loss terendah 0.29 dengan 1286 citra cuaca dapat diklasifikasi.

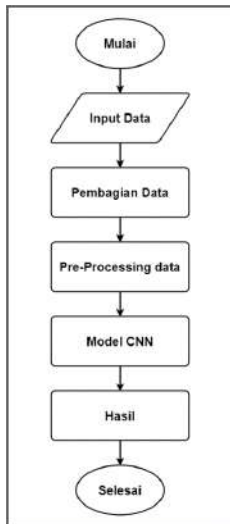
II. METODOLOGI PENELITIAN

Metode penelitian merupakan suatu langkah yang dilakukan oleh peneliti guna mengumpulkan informasi dan data serta untuk melakukan pengolahan pada data yang telah diperoleh. Teknik penelitian ini yaitu dengan menggunakan teknik studi literatur dan juga observasi. Studi literatur dilakukan untuk mencari tahu berbagai informasi terkait masalah, metode juga penelitian yang telah dilakukan dimasa lampau. Penelitian ini juga telah melakukan observasi dengan mengunjungi situs BMKG secara berkala serta beberapa website yang tertera pada Google.

Jenis penelitian ini merupakan penelitian eksperimen, yaitu metode penelitian yang memiliki tujuan untuk mengembangkan penerapan metode *deep learning* menggunakan Algoritma CNN untuk pengenalan cuaca. Algoritma CNN akan bertugas untuk melakukan pengolahan data dua dimensi, yang dapat diimplementasikan menjadi data citra. Metode merupakan suatu tata cara yang di rancang dan juga dipakai untuk menggapai suatu tujuan tertentu.

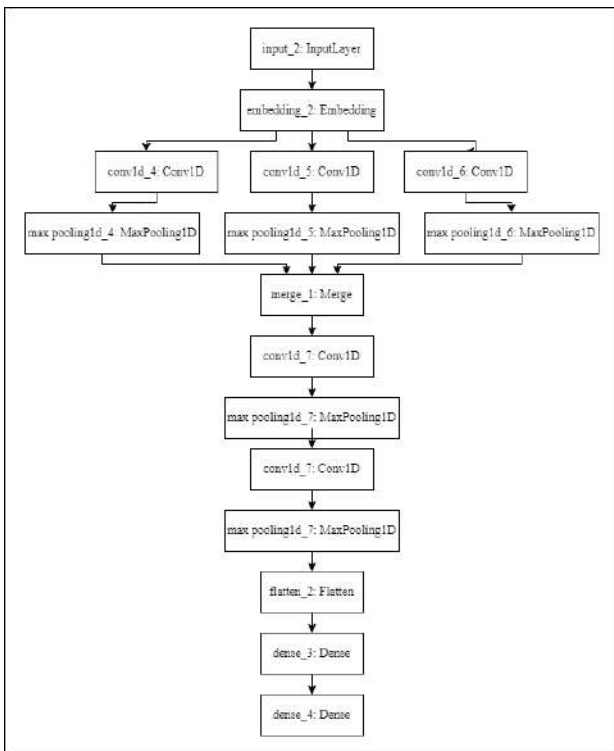
A. Analisis dan gambaran umum sistem

Metodologi penelitian digunakan agar alur dan hasil penelitian sesuai dengan tujuanc awal. Metodologi penelitian dapat digambarkan melalui langkah-langkah yang terlihat pada gambar 1. Mulai dari input dataset lalu membagi data training dan data validasi, setelah itu memproses data yang sudah disesuaikan tadi kedalam Model CNN.



Gambar 1. Langkah-langkah rancang bangun sistem

Sistem pengenalan cuaca yang dibuat pada penelitian ini memanfaatkan teknologi klasifikasi cuaca dengan metode deep learning. Proses klasifikasi cuaca dimulai dari input data berupa gambar cuaca yang diperoleh dari Website Kaggle. Setelah itu sistem akan melakukan pre-processing data dengan cara cropping dan pembersihan noise. Selanjutnya akan dilanjutkan dengan pembuatan training serta testing data. Kemudian dilanjutkan dengan implementasi algoritma CNN untuk memproses data gambar cuaca yang telah diinputkan. Lalu saat pemrosesan selesai akan ditampilkan kinerja model CNN secara grafis, sehingga didapatkan analisis gambar dari pemrosesan menggunakan algoritma CNN.



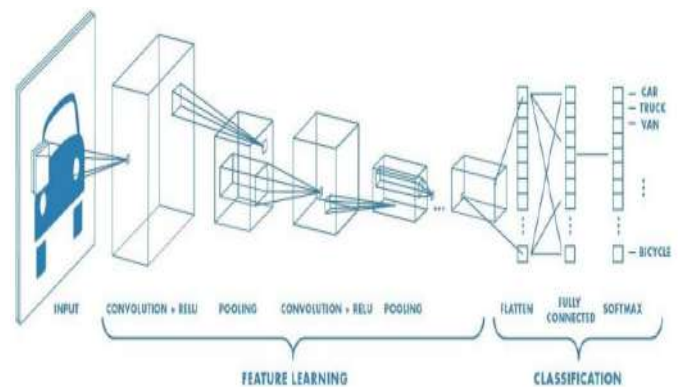
Gambar 2. Alur algoritma CNN

Gambar 2 adalah alur dari algoritma CNN yang akan diterapkan oleh peneliti.

B. Model Algoritma CNN

Algoritma CNN adalah salah satu algoritma yang menerapkan teknologi deep learning. CNN mampu untuk mengklasifikasi *image*, video, maupun mendeteksi obyek yang ada pada *image* atau bahkan wilayah yang ada didalam *image*. CNN terdiri dari layer yang mempunyai susunan neuron 3D yaitu, lebar, tinggi, dan kedalaman yang mana susunan lebar dan tinggi merupakan ukuran layar, dan kedalaman merupakan jumlah layar [12].

Selanjutnya akan diperlihatkan arsitektur dari CNN juga terbelah menjadi dua bagian besar yaitu Feature Extraction Layer dan Fully - Connected Layer (FC), seperti yang tertera pada gambar dibawah ini:



Gambar 3. Proses Convolutional Neural Network (CNN)

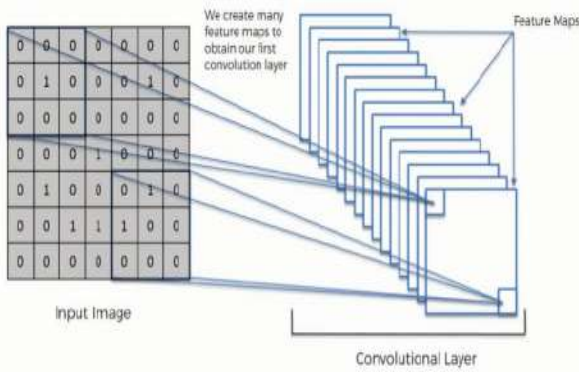
Jaringan saraf konvolusional. Kedengarannya seperti kombinasi aneh antara matematika dan biologi, tetapi jaringan ini telah menjadi beberapa inovasi paling berpengaruh di bidang visi komputer dan pemrosesan gambar. Jaringan saraf Convolutional adalah versi reguler dari multilayer perceptron (MLP). Mereka dikembangkan berdasarkan kerja neuron korteks visual hewan.

1. Lapisan Ekstraksi Fitur (Feature Extraction Layer)

Proses pada bagian ini adalah menjalankan “Encoding” dari sebuah *image* menjadi fitur yang berupa nomor-nomor yang merepresentasikan gambar tersebut. Lapisan Ekstraksi Fitur terdiri dari 2 bagian, Lapisan Konvolusional dan Lapisan Penggabungan.

a. Lapisan Konvolusional (Convolutional Layer)

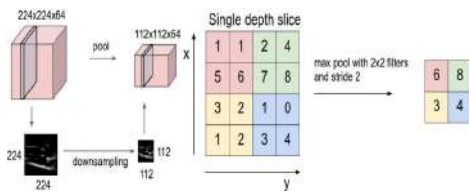
Lapisan Konvolusional merupakan lapisan dimana operasi konvolusi antara input dan nilai dari filter tersebut sehingga menghasilkan sebuah output atau biasa disebut sebagai Peta Aktivasi (Activation Map) atau Peta Fitur (Feature Map). Saringan (Filter) akan digeser ke seluruh bagian dari gambar. Proses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



Gbr 4. Lapisan Konvolusional

b. Lapisan Penggabungan

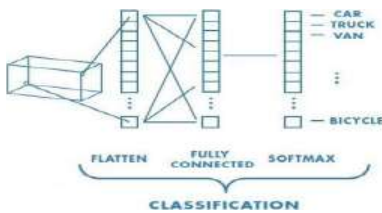
Lapisan penggabungan biasanya berada setelah lapisan konvolusional. Pada prinsipnya lapisan penggabungan terdiri dari sebuah filter dengan ukuran tertentu yang akan bergeser pada seluruh area Peta Fitur. Penggabungan yang biasa digunakan adalah Penggabungan nilai maksimum (*Max Pooling*). *Max pooling* membagi output dari convolution layer menjadi beberapa grid kecil lalu mengambil nilai maksimum dari setiap grid untuk menyusun *matrix* citra yang telah direduksi seperti yang terlihat pada gambar dibawah:



Gbr 5. Matriks Citra Gambar

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

2. Lapisan Terhubung Seluruhnya (Fully Connected Layer)



Gbr 6. Lapisan Terhubung

Lapisan Terhubung Seluruhnya adalah lapisan yang biasanya digunakan dalam penerapan Lapisan simpul (*neuron*) banyak yang terhubung (*Multilayer Perceptron*) dan memiliki tujuan untuk menjalakan transformasi pada tiap dimensi data supaya data dapat diklasifikasikan secara searah. Setiap simpul pada lapisan konvolusional perlu terlebih dahulu ditransformasi menjadi data satu dimensi sebelum dapat diinputkan ke dalam sebuah Lapisan Terhubung Seluruhnya. Karena hal tersebut mengakibatkan data kehilangan informasi spesialnya dan tidak reversibel,

Lapisan Terhubung Seluruhnya hanya dapat di implementasikan pada akhir jaringan. Algoritma ini memiliki beberapa langkah untuk merancang system klasifikasi cuaca [1]. Berikut ini merupakan langkah – langkah yang akan dilaksanakan dalam penelitian ini:

- a) Pembuatan model system
- b) Input data berupa gambar yang diperoleh dari google yang diperlukan untuk training dan validasi pada sistem.
- c) Melakukan preprocessing
- d) Memutuskan hyperparameter mana untuk model CNN. Hal ini meliputi jumlah convolutional layer, jumlah filter pada convolutional layer, ukuran *matrix* kernel dan filter pooling, jumlah neuron pada hidden layer dan fungsi aktivasi
- e) Mengkompilasi model CNN
- f) Melakukan training dan testing data
- g) Validasi model sesuai dengan dataset validasi
- h) Menampilkan kinerja model CNN secara grafis
- i) Penyelesaian dan evaluasi hasil klasifikasi

C. Machine Learning

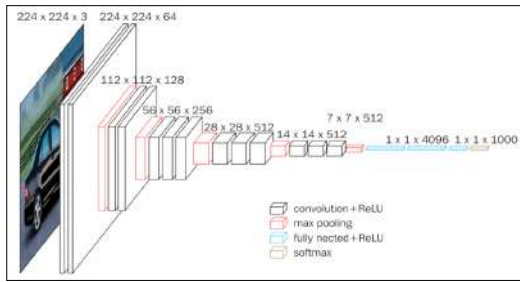
Machine Learning merupakan cabang dari *Artificial Intelligence* yang berfokus pada belajar melalui data yang mana pada pengembangan sistem yang mampu belajar secara mandiri tanpa harus diprogram berulang kali oleh manusia. *Machine Learning* membutuhkan data yang valid sebagai bahan untuk belajar untuk menghasilkan output yang optimal. *Machine Learning* pertama kali ditemukan oleh Adrien Marie Legendre, Thomas Bayes dan Andrey Markov pada tahun sekitar 1920 yang mana *machine learning* dapat diartikan sebagai metode komputasi yang menggunakan pengalaman untuk meningkatkan kinerja hingga membuat perkiraan. Selain itu, *machine learning* menerapkan teknik statistika untuk menghasilkan model otomatis dari sekumpulan data [3]. Beberapa metode Machine Learning sebagai berikut:

1. Supervised Machine Learning Algorithms
2. Unsupervised Machine Learning Algorithms
3. Semi-supervised Machine Learning Algorithms
4. Reinforcement Machine Learning Algorithms

D. Visual Geometry Group-16 Layer (VGG-16)

Visual Geometry Group-16 Layer (VGG-16) merupakan deep learning dengan arsitektur neural network berbasis convolutional layer. Ini telah banyak digunakan untuk model klasifikasi. VGG memiliki dua lapisan yang terhubung penuh diikuti dengan fungsi aktivasi softmax untuk lapisan keluaran. Semua lapisan ini memiliki bobot yang berbeda, dan setiap lapisan membentuk arsitektur jaringan sebanyak ini. VGG-16 telah menerapkan jaringan konvolusional untuk memperkecil bentuk input terlebih dahulu untuk mempelajari ringkasan input sebagai model klasifikasi [9]. Rectified linear unit (ReLU) ialah komponen dalam jaringan konvolusi untuk menyelesaikan lapisan tersembunyi fungsional untuk model pembelajaran. Lapisan yang terhubung sepenuhnya cenderung

mengidentifikasi jaringan konvensional. Lapisan keluaran menggunakan aktivasi softmax untuk klasifikasi multikelas.



Gbr 7. Arsitektur VGG16 (Sumber:neurohive.io)

E. Graphical User Interface (GUI)

GUI (Graphical User Interface) adalah antarmuka pengguna yang menggunakan elemen visual seperti jendela, tombol, menu, dan dialog untuk memfasilitasi interaksi antara pengguna dan perangkat lunak atau sistem komputer. Dalam GUI, pengguna dapat berinteraksi dengan perangkat lunak atau sistem dengan menggunakan input dari perangkat keras seperti mouse dan keyboard, dan menerima umpan balik atau output melalui tampilan grafis. Modul tkinter merupakan modul yang bekerja bersama dengan bahasa pemrograman Python untuk menyediakan kemudahan dalam pembuatan aplikasi dengan antarmuka. Dengan menggunakan modul tkinter, pengembang dapat membuat tampilan aplikasi yang intuitif dan menarik bagi pengguna melalui pengaturan jendela, penempatan tombol, dan tampilan grafis lainnya. Modul ini memungkinkan pengguna untuk berinteraksi dengan aplikasi melalui gambar dan ikon yang mempermudah pemahaman dan navigasi [11].

F. Proses Pengumpulan Data

Pengumpulan data dilakukan dengan cara mengumpulkan citra digital cuaca yang diperoleh dari website Kaggle dengan id author "pratik2901". Data latih digunakan algoritma CNN untuk membentuk sebuah model klasifikasi. Pengumpulan data dilakukan dengan cara mengumpulkan citra digital cuaca yang diperoleh dari website Kaggle dengan id author "pratik2901". Data latih dipakai algoritma CNN untuk membuat sebuah model klasifikasi, kemudian model ini dipakai untuk merepresentasikan pengetahuan yang akan dipakai untuk mengklasifikasikan data class yang ada. Model ini digunakan untuk merepresentasikan pengetahuan yang akan digunakan untuk mengklasifikasikan data class yang ada.

TABEL I
 DAFTAR CITRA DIGITAL AWAN

Jenis Cuaca	Jumlah
Cloudy	300
Rain	215
Sunrise	253
Shine	357
Total	1125

Pada tabel 1 diatas menjelaskan bahwa pengujian ini menggunakan 4 kelas cuaca untuk memenuhi kebutuhan data klasifikasi. Pada tiap kelas cuaca diambil 300 cloudy, 215 rain, 253 sunrise, 357 shine untuk data latihan, dan 250 citra untuk data uji coba.

G. Proses Pre-Processing Data

Agar diperoleh hasil akurasi yang maksimal maka perlu dilakukan preprocessing. Preprocessing bertujuan untuk menghapus latar belakang atau segmen bagian yang akan di analisis oleh sistem. Pada penelitian ini terdapat ada tiga proses yang dilakukan dalam pengolahan data awal, yaitu:

- 1) Cropping
 Cropping berfungsi meningkatkan kefokuskan image serta seragam. Cropping merupakan penghapusan area terluar dari image yang tidak diperlukan. Biasanya proses ini terdiri dari penghilangan beberapa area periferal image untuk menghapus sampah asing dari image, untuk meningkatkan pembingkaiannya dan untuk mengubah ratio aspek. Proses ini dilakukan dengan menggunakan perangkat lunak pengolah image secara manual [5].
- 2) Pembersihan Noise
 Noise adalah produk sampingan dalam pengambilan image yang tidak diharapkan dengan menghilangkan informasi yang diinginkan. Pembersihan Noise harus dilakukan untuk menjaga dan meningkatkan kualitas image sehingga nantinya akan menghasilkan nilai akurasi yang tinggi. Proses pembersihan noise ini dapat dilakukan secara manual juga dengan menggunakan bantuan perangkat lunak pengolah image.
- 3) Membuat data training dan testing
 Pada tahap preprocessing terdapat proses pembuatan data training dan testing. Total terdapat 100 citra pada setiap kategori. Pembagian data dilakukan secara sama rata pada masing – masing kategori. Penulis akan membuat tiga kali pengujian pembagian data mulai dari 90 data training dan 10 data validasi, lalu 80 data training dan 20 data validasi, dan 70 data training dan 30 data validasi.

H. Pelatihan Model

Pelatihan model menggunakan data latih dan data validasi. Dalam penelitian ini, peneliti menggunakan 100 epoch pertama dengan learning rate 3×10^{-3} dan 100 epoch kedua dengan learning rate dari 1×10^{-4} sampai 1×10^{-5} , serta batch size sebesar 16. Dari pelatihan model akan menghasilkan nilai akurasi (accuracy).

I. Uji Validitas Data

Proses uji validasi menggunakan pendekatan pembagian data menjadi subset pelatihan (train) dan subset validasi (val) berdasarkan direktori yang pilih melalui pemilihan folder. Subset pelatihan digunakan untuk melatih model, sedangkan subset validasi digunakan untuk validasi atau evaluasi model.

III. HASIL DAN PEMBAHASAN

Pada bagian bab ini akan dijelaskan mengenai dari hasil

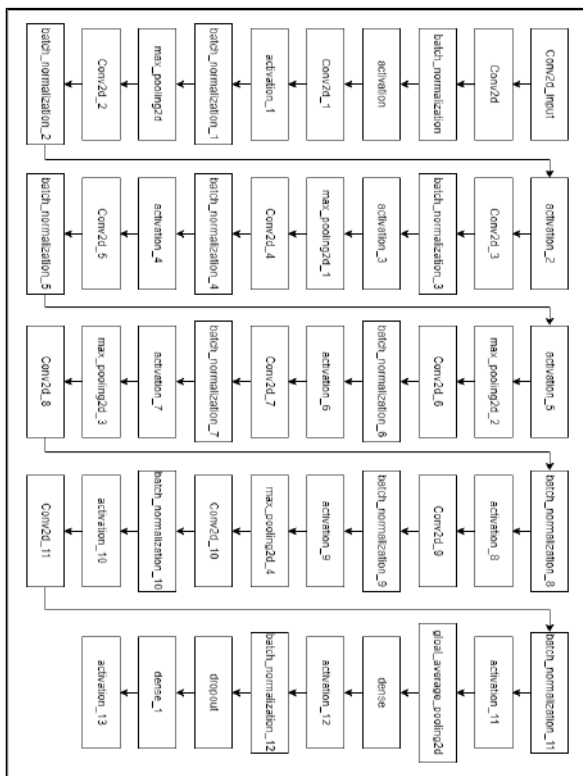
pengujian dari penerapan yang sebelumnya sudah diajukan dan disetujui dengan judul “Penerapan Metode *Deep Learning* Menggunakan Algoritma CNN Dengan Arsitektur VGG Net Untuk Pengenalan Cuaca”. Hasil penelitian dan pembahasan merupakan analisis data yang mencakup pembahasan berdasarkan rumusan masalah dan tujuan penelitian. Tahap ini mengindikasikan kesiapan sistem untuk diuji dalam situasi nyata, sehingga dapat mengevaluasi apakah hasil yang diperoleh sesuai dengan tujuan yang diharapkan.

A. Hasil Penelitian

Hasil akan dijelaskan berdasarkan pengujian yang telah dilakukan proses *training* data pada dataset yang sudah disiapkan menggunakan algoritma CNN. Berikut adalah hasil pembuatan aplikasi dan hasil pengujian yang telah dilakukan:

1. Penerapan Arsitektur CNN

Proses pelatihan ini digunakan teknik augmentasi data seperti pergeseran, pemotongan, dan pembalikan horizontal pada gambar. Selanjutnya, model CNN dibangun dengan beberapa lapisan konvolusi dan penggabungan. Setelah itu, model dilatih menggunakan generator data pelatihan dan validasi dalam beberapa epoch. Evaluasi dilakukan dengan memplot grafik akurasi pada data pelatihan dan validasi, serta menggunakan data pelatihan untuk evaluasi tambahan. Setelah proses pelatihan selesai, model disimpan untuk digunakan di masa mendatang.

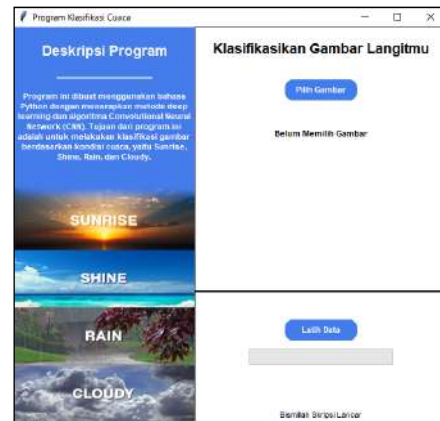


Gambar 8. Alur Algoritma CNN

Gambar 8 adalah arsitektur CNN yang digunakan untuk melatih dataset pada program Klasifikasi cuaca yang penulis buat.

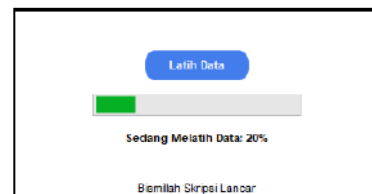
2. Tampilan Antarmuka Program

Tampilan program dibuat agar memudahkan interaksi user saat akan menggunakan fitur latih data atau prediksi gambar cuaca.



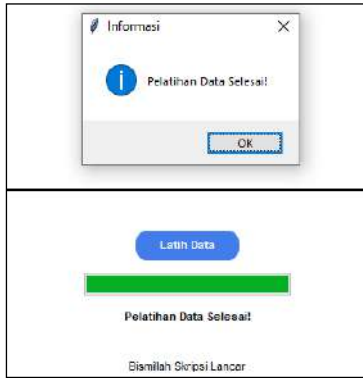
Gambar 9. Tampilan utama program

Gambar 9 adalah tampilan utama program, tampilan ini memiliki penjelasan program pada frame kiri. Pada frame kanan terdapat tombol Pilih Gambar dan tombol Latih Data. Tombol Pilih Gambar digunakan untuk input gambar ke program untuk diklasifikasikan cuacanya, sedangkan tombol Latih Data digunakan untuk melatih data lalu menghasilkan output berupa model CNN yang siap digunakan untuk prediksi gambar.



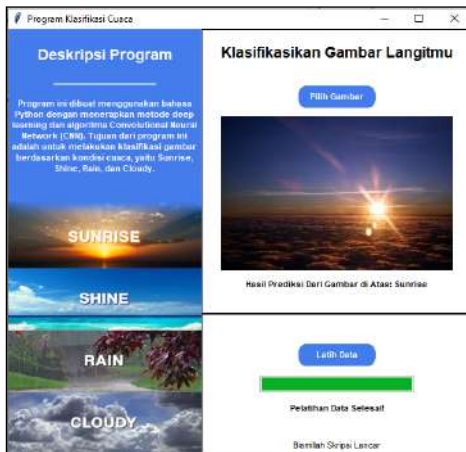
Gambar 10. Progress bar saat melatih data

User bisa memilih tombol pilih gambar jika sudah memiliki file hasil data training yang bernama “final_model.h5”. Jika belum mempunyai atau ingin melatih data baru maka pilih tombol latih data lalu memilih folder yang sudah berisi *dataset*. Gambar diatas merupakan tampilan *progress bar* yang akan muncul setelah memilih folder dataset.



Gambar 11. Progress bar dan notifikasi saat pelatihan selesai

Gambar diatas merupakan tampilan program klasifikasi cuaca saat proses latih data telah selesai. Progress bar akan menampilkan tulisan “Pelatihan data telah selesai”, dan juga akan memunculkan jendela notifikasi yang juga bertuliskan “Pelatihan Data selesai”.



Gambar 12. Tampilan saat gambar sudah diprediksi

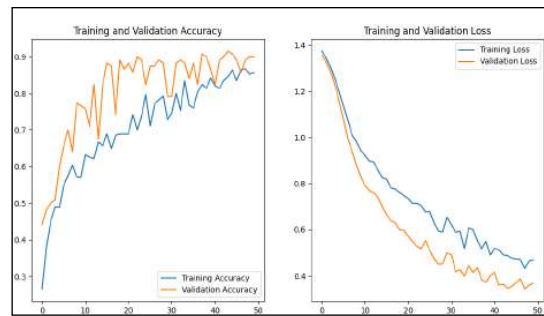
Gambar diatas adalah tampilan program klasifikasi cuaca saat sudah selesai melatih data dan juga sudah menginput gambar untuk diklasifikasi. Pada tampilan ini akan diperlihatkan hasil prediksi dari gambar yang telah diinput dengan tulisan “Hasil Prediksi Dari Gambar di Atas:” lalu hasil prediksinya akan berupa tulisan Cloudy atau Rain atau Sunrise atau Sunshine.

3. Hasil Pengujian Akurasi Dan Validasi

Dalam proses data training program ini, akurasi dihitung menggunakan matrik yang telah ditentukan saat mengkompilasi model. Nilai akurasi ini menghitung seberapa banyak prediksi yang benar dibandingkan dengan total jumlah data. Setelah model selesai dilatih, penulis akan memvalidasi model dengan cara mencoba testing menggunakan data gambar yang berbeda dari data latih yang sudah disiapkan. Dengan tujuan untuk menghitung jumlah gambar yang benar dan yang salah sebagai validasi akurasi untuk model. Penulis juga akan mencoba melakukan proses latih data dengan epoch atau iterasi yang

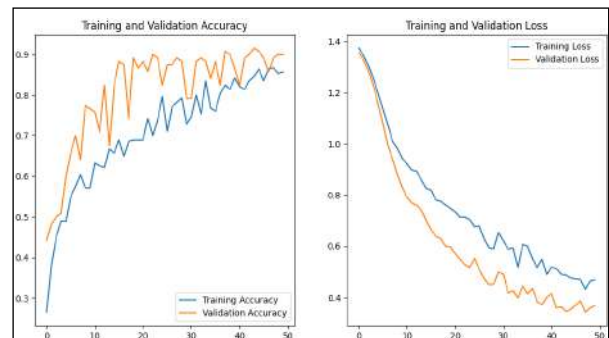
berbeda untuk melihat pengaruhnya terhadap hasil akurasi. Epoch yang akan digunakan untuk ujicoba adalah 50 Epoch, 100, dan 150.

a) Pengujian Dengan Menggunakan 50 Epoch



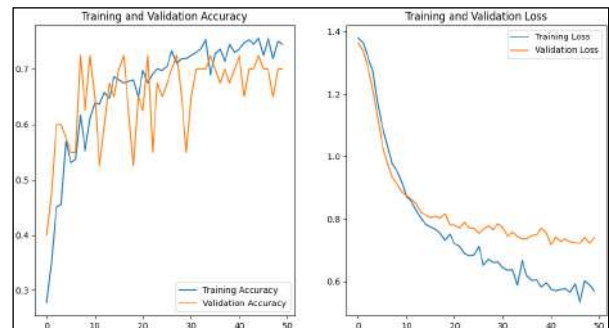
Gambar 13. Pengujian Pertama

Pada gambar diatas pengujian pertama dengan 50 epoch menggunakan 90 data latih dan 10 data validasi dihasilkan akurasi tertinggi yaitu 89,16%. Namun saat penulis mencoba untuk tesing gambar, hasil yang didapat adalah 22 gambar benar dan 18 gambar salah. Jadi yang didapatkan dari hasil testing hanyalah 55,00%..



Gambar 14. Pengujian Kedua

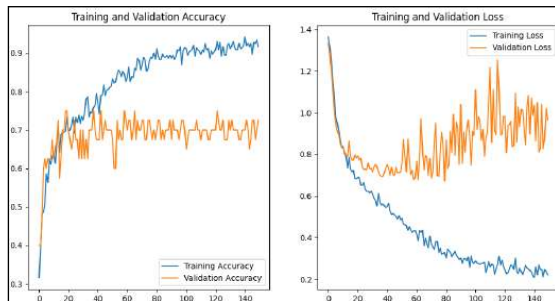
Pada gambar 14. pengujian kedua menggunakan 80 data latih dan 20 data validasi dihasilkan akurasi tertinggi yaitu 78,75%. Namun saat penulis mencoba untuk tesing gambar, hasil yang didapat adalah 45 gambar benar dan 35 gambar salah. Jadi yang didapatkan dari hasil testing hanyalah 56,25%.



Gambar 15. Pengujian Ketiga

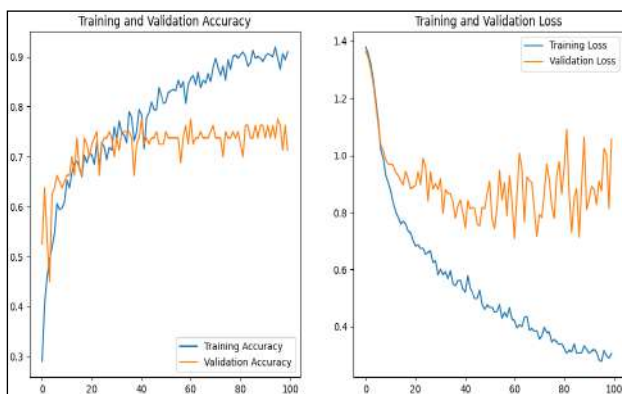
Pada gambar 15. pengujian ketiga 70 data latih dan 30 data validasi dihasilkan akurasi tertinggi yaitu 74,64%. Namun saat penulis mencoba untuk tesing gambar, hasil yang didapat adalah 64 gambar benar dan 56 gambar salah. Jadi yang didapatkan dari hasil testing hanyalah 53,33%.

b) Pengujian Dengan Menggunakan 100 Epoch



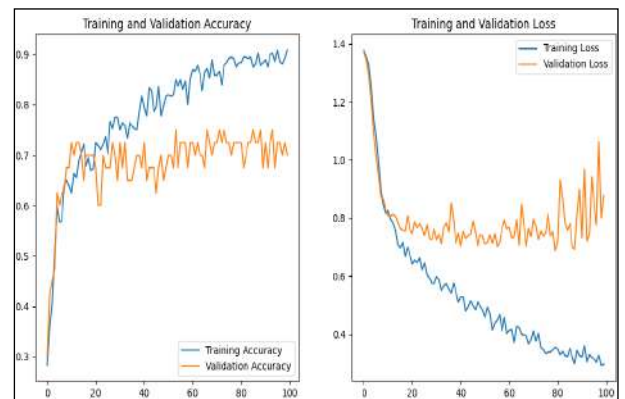
Gambar 16. Pengujian Keempat

Pada gambar 16. merupakan gambar pengujian keempat dengan sampel 100 epoch menggunakan 90 data latih dan 10 data validasi. Pengujian ini menghasilkan akurasi tertinggi yaitu 90,56%. Namun saat penulis mencoba untuk testing gambar, hasil yang didapatkan adalah 23 gambar benar dan 17 gambar salah. Jadi yang didapatkan dari hasil testing hanyalah 57,50%.



Gambar 17. Pengujian Kelima

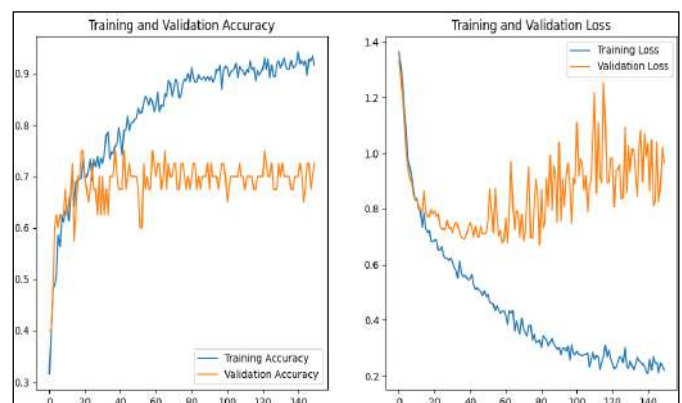
Pada gambar diatas merupakan gambar pengujian kelima dengan menggunakan sampel 100 Epoch terdiri dari 80 data latih dan 20 data validasi, Pengujian ini menghasilkan akurasi tertinggi yaitu 91,87%. Namun saat penulis mencoba untuk testing gambar, hasil yang didapatkan adalah 47 gambar benar dan 33 gambar salah. Jadi yang didapatkan dari hasil testing hanyalah 58,75%.



Gambar 18. Pengujian Keenam

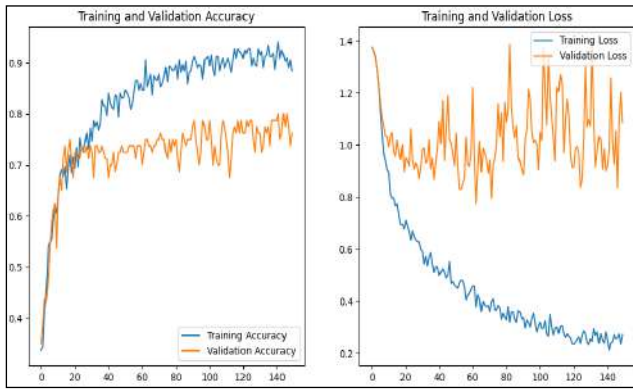
Pada gambar diatas merupakan gambar pengujian keenam dengan menggunakan sampel 100 Epoch yaitu 70 data latih dan 30 data validasi, pengujian ini menghasilkan akurasi tertinggi yaitu 89,28%. Namun saat penulis mencoba untuk tesing gambar, hasil yang didapatkan adalah 67 gambar benar dan 53 gambar salah. Jadi yang didapatkan dari hasil testing hanyalah 55,83%.

c) Pengujian Dengan Menggunakan 150 Epoch



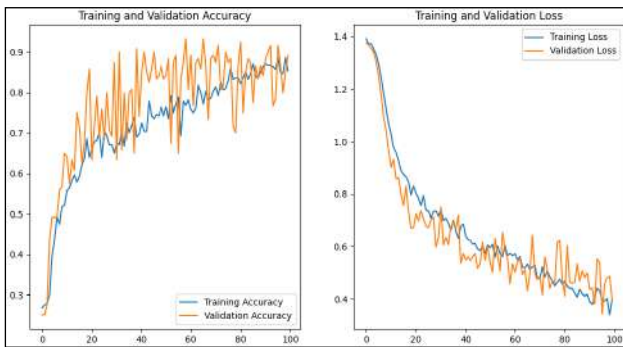
Gambar 19. Pengujian Ketujuh

Pada gambar 19. pengujian ketujuh dengan 150 epoch menggunakan 90 data latih dan 10 data validasi dihasilkan akurasi tertinggi yaitu 94,16 %. Namun saat penulis mencoba untuk tesing gambar, hasil yang didapat adalah 26 gambar benar dan 14 gambar salah. Jadi yang didapatkan dari hasil testing hanyalah 65,00%.



Gambar 20. Pengujian Kedelapan

Pada gambar 20, pengujian kedelapan dengan menggunakan 80 data latih dan 20 data validasi dihasilkan akurasi tertinggi yaitu 92,18%. Namun saat penulis mencoba untuk tesing gambar, hasil yang didapat adalah 49 gambar benar dan 31 gambar salah. Jadi yang didapatkan dari hasil testing hanyalah 61,25%.



Gambar 21. Pengujian Kesembilan

Pada gambar 21, pengujian kesembilan dengan menggunakan 70 data latih dan 30 data validasi dihasilkan akurasi tertinggi yaitu 72,50%. Namun saat penulis mencoba untuk tesing gambar, hasil yang didapat adalah 51 gambar benar dan 39 gambar salah. Jadi yang didapatkan dari hasil testing hanyalah 56,66%.

B. Pembahasan

Dari percobaan pertama sampai dengan percobaan kesembilan hasil dari akurasi akan dirangkum kedalam tabel dibawah ini.

TABEL III
NILAI AKURASI HASIL DATA LATIH

		Jumlah Data Latih / Data Validasi		
		70/30	80/20	90/10
Jumlah Epoch	50	74,64%	78,75%	89,16%
	100	89,28%	91,87%	90,56%
	150	89,52%	92,18%	94,16%

Pada tabel 2 dapat disimpulkan jumlah data latih dan jumlah epoch sangat mempengaruhi tingkat akurasi pada hasil model CNN yang dibuat. Semakin banyak jumlah data yang digunakan dan semakin banyak jumlah epoch yang dipakai maka akan semakin tinggi juga nilai akurasi yang didapat.

TABEL IIIII
NILAI AKURASI HASIL DATA TESTING

		Jumlah Data Latih / Data Validasi		
		70/30	80/20	90/10
Jumlah Epoch	50	53,33%	56,25%	55,00%
	100	55,83%	58,75%	57,50%
	150	56,60%	61,25%	65,00%

Pada tabel 3 ditampilkan nilai akurasi dari hasil data testing dengan menggunakan data yang berbeda. Dapat disimpulkan bahwa penambahan jumlah data dan jumlah epoch memang menambah nilai akurasi yang didapat, namun tidak bertambah sebanyak nilai akurasi hasil dari data latih. Penyebab pertambahan nilai akurasi data testing yang sedikit ini bisa disebabkan oleh kualitas data latih yang disediakan. Data latih yang berkualitas baik, dan mencakup variasi yang sesuai akan membantu model untuk belajar pola-pola yang lebih baik dan menggeneralisasi ke data yang belum pernah dilihat sebelumnya.

IV. KESIMPULAN

1. Penelitian dengan menerapkan metode deep learning menggunakan algoritma CNN berhasil memberikan hasil yang memuaskan dalam mengklasifikasikan cuaca. Adanya arsitektur VGGNet16 juga sangat cocok digunakan untuk melakukan pengolahan data yang berupa gambar. Namun, perbedaan pada arsitektur jaringan dan parameter seperti jumlah epoch, batch size, dan learning rate yang digunakan juga bisa mempengaruhi hasil dari model CNN.
2. Jumlah data dapat mempengaruhi hasil akurasi dari model CNN, karena semakin banyak data yang digunakan maka akan semakin baik kinerja model CNN yang dibuat. Jumlah epoch juga sangat mempengaruhi hasil model CNN yang dibuat, karena dengan meningkatkan jumlah epoch, model memiliki lebih banyak kesempatan belajar pola-pola yang ada di dalam data.

V. SARAN

Untuk memaksimalkan kinerja CNN dapat memakai data set di atas 2000 dan meningkatkan kualitas data yang digunakan, agar dapat melatih data pengelompokan lebih tepat lagi sehingga akan menghasilkan akurasi yang lebih tinggi. Selain itu untuk memaksimalkan program supaya dapat digunakan sebagai alat bantu klasifikasi cuaca secara masal dapat memakai arsitektur backend service berbasis RestApi. Dengan memakai arsitektur ini kedepannya dapat berguna

untuk berbagai client seperti web dan mobile. Dan sebagai pengganti k-fold cross validation dapat menggunakan percentage split agar metode untuk proses validasi dapat lebih bagus lagi.

UCAPAN TERIMA KASIH

Penulis mengucapkan banyak-banyak rasa syukur terhadap Tuhan Yang Maha Pengasih lagi Maha Penyayang atas segala pertolonganNya, sehingga peneliti sanggup menyelesaikan skripsi ini dengan baik. Terimakasih juga penulis sampaikan terhadap kedua orang tua dan kakak yang selalu memberikan motivasi beserta nasihat-nasihat, Dosen Pembimbing Skripsi yang selalu membimbing penulis serta menasihati, tidak lupa juga teman-teman seperjuangan yang selalu memberi dukungan terhadap penulis dalam melakukan penelitian.

REFERENSI

- [1] A. Sharma and Z. Saad Ismail. (2022). Weather Classification Model Performance: Using CNN, Keras-Tensor Flow. *ITM Web Conf.*, 42.
- [2] Badan Meteorologi, K. d. (2022). Retrieved from *PerubahanIklim*. Diambil kembali dari BMKG: <https://www.bmkg.go.id/iklim/?p=ekstrem-perubahan-iklim>
- [3] C. Imam, Sutrisno, A. S. Arief, H. Uswatun, and I. F. Yessica. (2020). Ai, Machine Learning & Deep Learning (Teori & Implementasi). 1-455.
- [4] I. Wulandari, H. Yasin, and T. Widiharih. (2020). Klasifikasi Citra Digital Bumbu Dan Rempah Dengan Algoritma Convolutional Neural Network (Cnn). *Jurnal Gaussian*, 9(3), 273-282.
- [5] J. Lu, L. Tan, and H. Jiang. (2021). Review on convolutional neural network (CNN) applied to plant leaf disease classification. *Agriculture*, 11(8), 1-18.
- [6] Keerthana P. J., Rajeswari P., Amirtharaj R., Pandiya Rajan G. (2022). "Weather Forecasting using Deep Learning Algorithm". *International Journal Of Special Education*. Vol. 37, No. 3.
- [7] Randhy Sulisty B., Raditiana Patmasari, Sofia Saidah. (2021). Klasifikasi Cuaca Menggunakan Metode Convolutional Neural Network (Cnn). *e-Proceeding of Engineering*, 8(5), 2355-9365.
- [8] Suartika, I. W. (2016). Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) Pada Caltech 101. *Jurnal Teknologi ITS*, 5(1)
- [9] Suci Aulia A. & Dadi Rahmat. (2022). Brain Tumor Identification Based on VGG-16 Architecture and CLAHE Method. *INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION*. Vol 6 No.1. 96-102.
- [10] Supriyadi, E. (2021). Prediksi Parameter Cuaca Menggunakan Deep Learning Long-Short Term Memory (Lstm). *Jurnal Meteorol. dan Geofisika*, 21(2).
- [11] Warnilah, A. I., Sutisna, H., Mulyana, A. J., Nuraeni, F. S., & Widiyanto, T. A. (2022). Program Aplikasi Pendeteksi Masker Dengan Menggunakan Algoritma Haarcascade. *Jurnal Sains dan Manajemen*. Vol 10 No.1. 89-98.
- [12] Zufar, M. (1998). *Introductory Computer Vision and Image Processing*. *Sens. Rev.*, 18(3), 2-4.