

Pengenalan Cuaca Indonesia Berdasarkan Citra Langit Menggunakan CNN Arsitektur MobileNetV2

Darryl Matthew Kurniawan¹, Chairisni Lubis²

^{1,2} Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara, Jln. Letjen S. Parman No. 1, Jakarta, 11440, Indonesia

E-mail: ¹darryl.535210011@stu.untar.ac.id, ²chairisnil@fti.untar.ac.id

Abstrak

Negara Indonesia memiliki cuaca yang bervariasi dan memiliki dampak yang akan berpengaruh pada kehidupan sehari-hari. Pengenalan cuaca dengan memanfaatkan citra langit merupakan salah satu solusi yang efektif untuk mendapatkan informasi cuaca berdasarkan dengan kondisi langit. Penelitian ini akan menggunakan salah satu arsitektur CNN yaitu MobileNetV2 untuk melakukan klasifikasi ke 4 kategori cuaca yaitu cerah, berawan, mendung, dan hujan, disertai dengan prediksi data numerik berupa suhu, kelembaban udara, kecepatan angin, sinar UV, dan tekanan udara. Dataset yang digunakan akan berupa citra langit dan data numerik yang diperoleh dengan pemantauan cuaca selama kurang lebih 2 bulan. Hasil eksperimen menunjukkan bahwa model yang dirancang dapat memperoleh akurasi validasi sebesar 78%, dengan nilai MSE untuk data numerik 0.6478, dan total loss validasi 1.547. Hasil ini menunjukkan bahwa model yang dirancang memiliki potensi untuk dapat melakukan pengenalan cuaca secara efektif di Indonesia dengan meningkatkan ukuran dataset dan optimasi lebih lanjut.

Kata kunci—Cuaca, Pengenalan, MobileNetV2

Abstract

Indonesia has a varied weather and has an impact that will affect daily life. Weather recognition by utilizing sky imagery is one of the effective solutions to obtain weather information based on sky conditions. This study will use one of the CNN architectures, namely MobileNetV2, to classify 4 weather categories, namely sunny, cloudy, overcast, and rainy, accompanied by predictions of numerical data in the form of temperature, air humidity, wind speed, UV rays, and air pressure. The dataset used will be in the form of sky imagery and numerical data obtained by monitoring the weather for approximately 2 months. The experimental results show that the designed model can obtain a validation accuracy of 78%, with an MSE value for numerical data of 0.6478, and a total validation loss of 1,547. These results indicate that the designed model has the potential to be able to perform effective weather recognition in Indonesia by increasing the dataset size and further optimization.

Keywords—Weather, Recognition, MobileNetV2

1. PENDAHULUAN

Negara Indonesia memiliki iklim tropis sehingga Indonesia hanya mempunyai 2 musim, yaitu musim kemarau dan musim hujan [1]. Dengan iklim tropis yang dimiliki negara Indonesia, membuat cuaca di Indonesia bervariasi dengan beberapa parameter yang saling terkait. Parameter-parameter tersebut seperti suhu, kelembaban, kecepatan angin, dan arah angin akan terpengaruh oleh faktor-faktor geografis dan juga jarak dan posisi dari matahari [2]. Pada lingkungan atau daerah-daerah kecil yang belum terjangkau oleh stasiun cuaca, pemantauan cuaca secara akurat masih menjadi tantangan. Hal ini merupakan hal penting karena cuaca dinamis yang ada pada Indonesia akan dapat memberikan dampak ke beberapa aktivitas seperti pertanian, perikanan, dan transportasi.

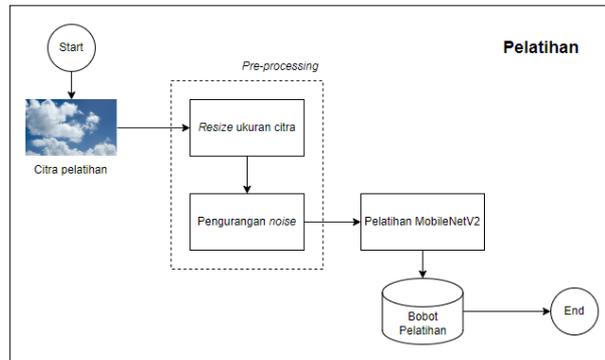
Penggunaan dari citra atau gambar langit untuk mengidentifikasi kondisi cuaca memungkinkan untuk citra langit dapat diolah menjadi sebuah informasi yang dapat memberikan manfaat mengenai kondisi dari cuaca. Metode yang efektif dan efisien untuk digunakan dalam pengenalan pada gambar yaitu adalah metode Convolutional Neural Network (CNN). Dengan kemampuan dari metode CNN dalam mengenali pola-pola visual yang terdapat pada gambar dengan baik, CNN banyak digunakan pada aplikasi pengenalan pola dan gambar seperti aplikasi pengenalan gestur, wajah, ataupun klasifikasi objek [3]. Salah satu arsitektur dari CNN yaitu MobileNet merupakan arsitektur yang dirancang dan dikembangkan secara khusus untuk melakukan pengolahan gambar pada perangkat dengan sumber daya yang terbatas [4]. Dengan komputasi yang ringan dan efisien, MobileNet merupakan arsitektur yang dapat digunakan untuk menangani kompleksitas dari citra atau gambar langit yang beragam dari segi pencahayaan, warna, sudut, dan bentuk awan yang bervariasi.

Dari cuaca Indonesia yang beragam, cuaca cerah, berawan, mendung, dan hujan dapat dikenali berdasarkan dengan citra langitnya. Pada cuaca cerah, berawan, dan mendung akan dipengaruhi oleh beberapa faktor seperti jenis awan, tutupan awan, dan juga sinar yang dipancarkan oleh matahari [5]. Sedangkan untuk cuaca hujan memiliki awan yang cenderung bersifat jenuh dan disertai dengan turunnya air dari langit [6].

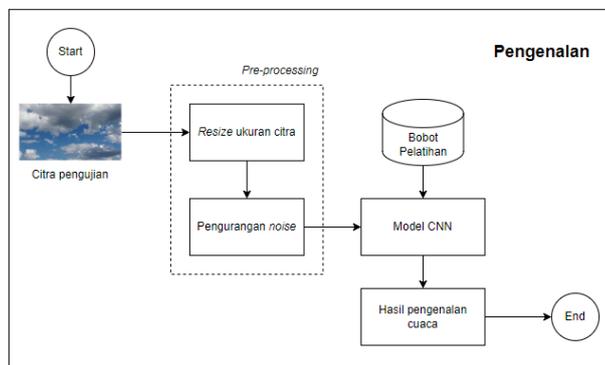
Pada penelitian ini bertujuan untuk melakukan pengenalan cuaca berdasarkan dengan citra langitnya. Dengan mengimplementasikan MobileNet, salah satu arsitektur dari CNN, diharapkan dapat memberikan hasil yang efisien dan akurat dalam melakukan klasifikasi cuaca cerah, berawan, mendung, dan hujan, disertai dengan informasi-informasi numerik tambahan yaitu suhu, kelembaban, kecepatan angin, sinar UV, dan tekanan udara yang diharapkan akan dapat memperkuat hasil dari pengenalan cuaca. Untuk data yang digunakan pada *dataset* merupakan gambar atau citra langit yang diperoleh dengan melakukan pemantauan cuaca selama kurang lebih 2 bulan yang dimulai dari Agustus 2024 sampai akhir Oktober 2024 disertai dengan informasi numerik untuk suhu, kelembaban, kecepatan angin, sinar UV, dan tekanan udara pada saat citra langit diambil.

2. METODE PENELITIAN

Sistem pengenalan cuaca berdasarkan citra langit akan memberikan hasil pengenalan dari cuaca yang sedang berlangsung berdasarkan dari citra atau gambar kondisi langit. Sistem akan berjalan dalam aplikasi web yang dimana dapat diakses melalui *browser* dari berbagai perangkat termasuk *mobile* ataupun *desktop*.



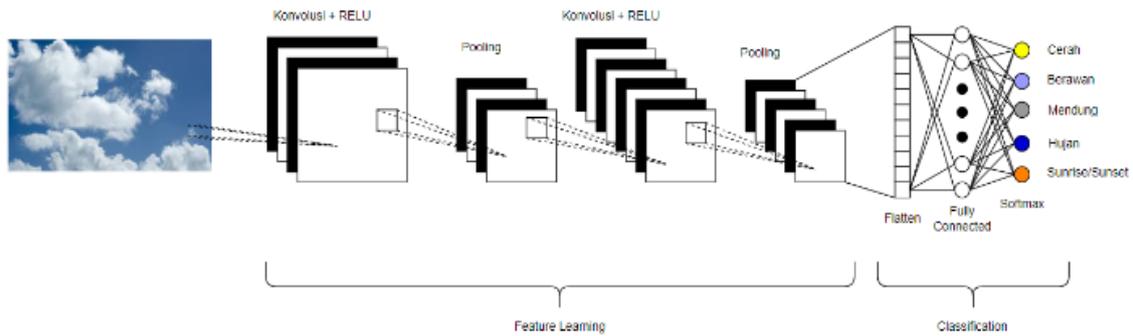
Gambar 1 Flowchart Tahap Pelatihan Model



Gambar 2 Flowchart Tahap Pengujian Model

2.1 Convolutional Neural Network

Convolutional Neural Network atau CNN merupakan metode yang biasanya digunakan dalam melakukan analisis pada citra dan pengenalan pola. CNN dikenal dengan hasilnya yang efektif dalam berbagai bidang seperti klasifikasi gambar, pendeteksian objek, atau pengenalan wajah [7]. Arsitektur dari CNN sendiri terbagi menjadi 2 tahap yaitu tahap pembelajaran fitur dan tahap klasifikasi.

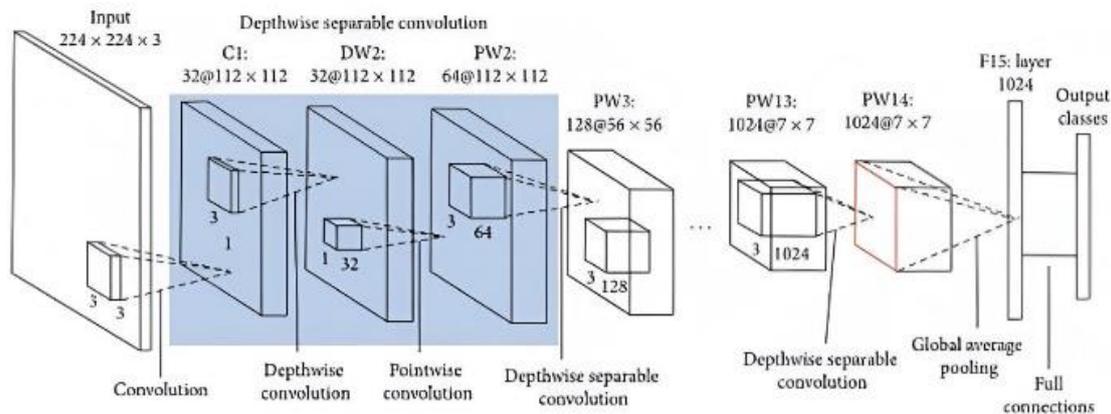


Gambar 3 Arsitektur CNN

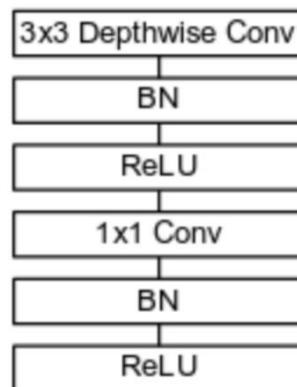
Pada tahap pertama, CNN akan mempelajari fitur-fitur ataupun pola visual yang ada pada citra. Pada tahap ini meliputi beberapa lapisan yaitu lapisan konvolusi, fungsi aktivasi, dan lapisan *pooling*, sedangkan pada tahap kedua yaitu tahap klasifikasi terdiri dari proses *flattening* untuk mengubah output dari lapisan sebelumnya menjadi vektor 1 dimensi, lapisan *fully-connected* dan fungsi *softmax* untuk mendapatkan nilai probabilitas pada setiap kelas. Kelas dengan nilai probabilitas yang paling tinggi nantinya akan diambil sebagai hasil dari klasifikasi atau pengenalan.

2.2 MobileNet

MobileNet merupakan arsitektur CNN yang dirancang khusus untuk pemrosesan gambar pada perangkat yang memiliki keterbatasan sumber daya. Perbedaan utama antara arsitektur MobileNet dengan arsitektur CNN yaitu pada lapisan konvolusinya, dimana MobileNet menerapkan konsep *depthwise separable convolution* yang memecah lapisan konvolusi menjadi 2 yaitu *depthwise convolution* dan *pointwise convolution* [8]. Pada *depthwise convolution* akan melakukan konvolusi menggunakan kernel 3x3 untuk setiap saluran, sedangkan pada *pointwise convolution* melakukan konvolusi menggunakan kernel 1x1 untuk menggabungkan output dari *depthwise convolution*. MobileNet juga menerapkan *batch normalization* dan ReLU dimana *batch normalization* berfungsi untuk menstabilkan dan mempercepat pelatihan, sedangkan ReLU merupakan fungsi aktivasi yang digunakan untuk menangani non-linearitas.



Gambar 4 Arsitektur MobileNet [9]



Gambar 5 Depthwise Separable Convolution [8]

2.3 MobileNetV2

MobileNetV2 merupakan penyempurnaan dari versi awalnya, dimana pada MobileNetV2 terdapat 2 inovasi tambahan yaitu *inverted residual blocks* dan *linear bottleneck module* [10]. Peningkatan ini bertujuan agar dapat menangkap informasi yang ada secara lebih kompleks dengan melakukan ekspansi pada saluran sebelum masuk ke proses *depthwise separable convolution*, dan saluran akan dikompresi kembali ke dimensi yang lebih kecil setelah melalui proses *depthwise separable convolution* menggunakan *inverted residual blocks* [10]. Setelah dimensi kembali diperkecil, fungsi aktivasi ReLU akan menjadi tidak efisien karena akan memungkinkan terjadinya kehilangan informasi. Dengan menggunakan *linear bottleneck module*, maka fungsi aktivasi ReLU tersebut akan digantikan dengan transformasi linier sehingga akan menghindari hilangnya informasi yang mungkin terjadi pada saat kompresi saluran [10].

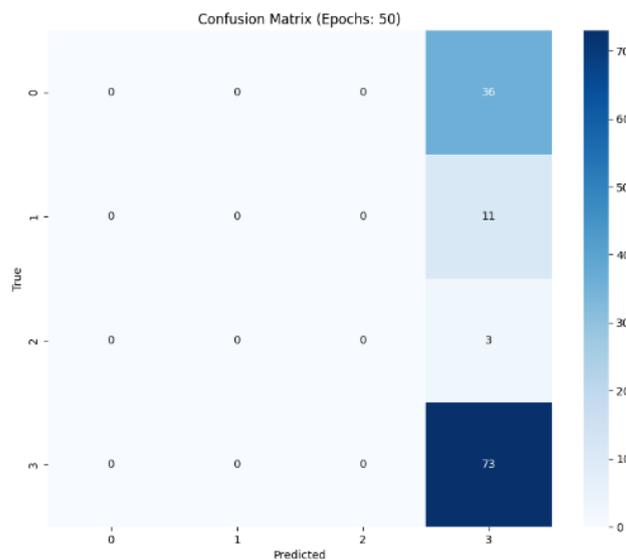
3. HASIL DAN PEMBAHASAN

Data citra yang diperoleh dengan pemantauana cuaca selama kurang lebih 2 bulan berjumlah sekitar 619 gambar pada 4 kelas cuaca. Namun terdapat ketidakseimbangan jumlah data per kelasnya, dimana kelas “Cerah” dan “Berawan” memiliki jumlah data yang jauh lebih banyak dibandingkan dengan data pada kelas “Mendung” dan “Hujan”. Hal ini disebabkan karena musim kemarau yang sedang berlangsung di Indonesia sehingga minimnya cuaca mendung atau hujan yang terjadi.

Tabel 1 Data Awal yang Diperoleh

Kategori Cuaca	Jumlah Data
Cerah	300
Berawan	207
Mendung	60
Hujan	32

Dari hasil pelatihan model menggunakan *dataset* awal yang tidak seimbang, memberikan nilai akurasi yang tinggi namun, bukan karena model berhasil mengenali fitur atau pola visual yang ada, melainkan model hanya memprediksi kelas dengan jumlah data yang mayoritas sehingga output yang diberikan akan selalu sama untuk setiap input-nya yaitu kelas “Cerah”.



Gambar 6 Confusion Matrix Menggunakan Dataset Tidak Seimbang

Untuk menghasilkan model yang optimal, maka memerlukan *dataset* yang seimbang untuk jumlah data pada setiap kelasnya agar model mampu mengenali dan mempelajari fitur-fitur yang ada untuk setiap kelas sehingga model dapat memberikan hasil output yang akurat sesuai dengan input yang dimasukkan. Untuk membuat *dataset* menjadi seimbang, pada kelas mayoritas akan dilakukan *undersampling* hingga menjadi 120 data, sedangkan pada kelas minoritas akan dilakukan augmentasi data dengan *rotate* dan *crop* sehingga jumlah data setiap kelas akan menjadi sama dan seimbang yaitu sebanyak 120 data per kelas.

Tabel 2 Data yang Digunakan Untuk *Dataset*

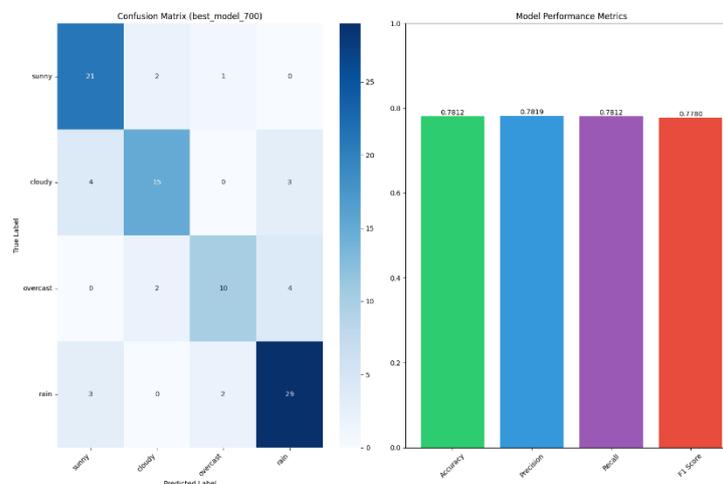
Kategori Cuaca	Jumlah Data
Cerah	120
Berawan	120
Mendung	120
Hujan	120

Pada pengujian model menggunakan *dataset* yang seimbang, akan dilakukan eksperimen menggunakan 4 nilai *epochs* yaitu 300, 500, 700, dan 1000 dan akan mengambil model dengan nilai akurasi validasi yang tertinggi. Model juga akan menggunakan *checkpoint* yang akan membandingkan nilai akurasi validasi per *epochs* dengan nilai akurasi validasi terbaik dan akan menyimpan model dengan nilai akurasi validasi yang lebih tinggi, sehingga pada akhir pelatihan akan menghasilkan 1 model dengan nilai akurasi validasi tertinggi per konfigurasi *epochs*.

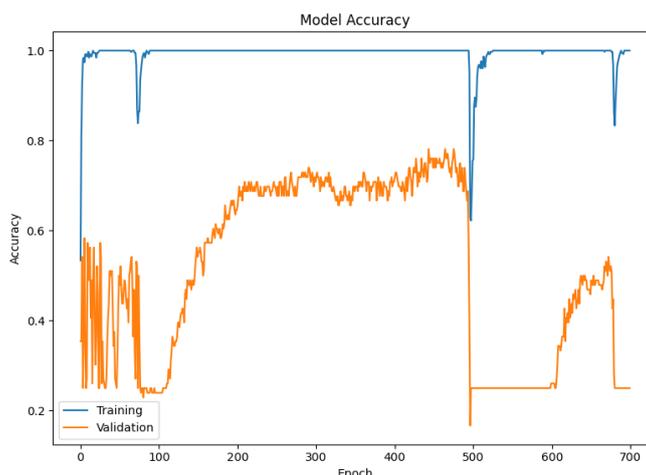
Tabel 3 Perbandingan Hasil Validasi per Konfigurasi *Epochs*

Epochs	Weather Accuracy	Numeric MSE	Loss
300	0.6875	0.6845	2.051
500	0.7708	0.6678	1.847
700	0.78125	0.6478	1.547
1000	0.7187	0.655	1.828

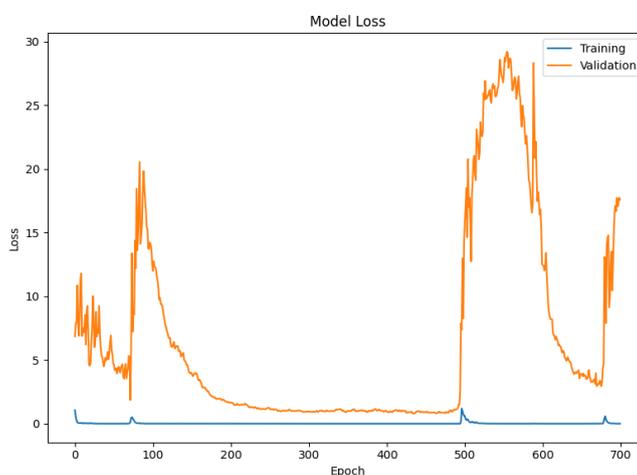
Dari hasil pelatihan yang telah dilakukan, didapat model dengan akurasi validasi terbaik yaitu pada model dengan *epochs* 700 dengan nilai akurasi validasi sebesar 78%, MSE validasi untuk nilai numerik sebesar 0.6478, dan nilai *loss* untuk validasi sebesar 1.547. Dari hasil yang diberikan, dapat dilihat bahwa untuk setiap peningkatan nilai *epochs* maka akurasi cuaca akan semakin naik, sedangkan untuk MSE nilai numerik dan jumlah *loss* akan semakin turun. Dengan nilai *epochs* yang lebih tinggi akan memberikan model kesempatan yang lebih untuk dapat mempelajari fitur atau pola yang ada. Namun, peningkatan nilai *epochs* tidak selalu memberikan hasil yang baik karena dapat memungkinkan bagi model untuk mengalami *overfitting* dimana model akan kesulitan dalam mengenali pola pada data citra baru karena terlalu menghafal data-data yang ada pada pelatihan.



Gambar 7 Confusion Matrix Model Dengan *Epochs* 700



Gambar 8 Grafik Akurasi Model Dengan Epochs 700



Gambar 9 Grafik Loss Model Dengan Epochs 700

Selanjutnya akan dilakukan pengujian pada model terbaik yang telah didapat yaitu model dengan 700 epochs. Pengujian akan dilakukan dengan 3 skenario yang berbeda. Pada skenario pertama akan menggunakan citra langit yang diperoleh dengan menggunakan kamera. Pada skenario pertama akan menggunakan sebanyak 20 citra per kelasnya dan akan membandingkan 6 parameter yaitu kategori cuaca dan 5 fitur numerik yaitu suhu, kelembaban, kecepatan angin, sinar UV, dan tekanan udara. Pada pengujian skenario kedua akan menggunakan citra langit yang diperoleh dari internet sehingga hanya akan membandingkan 1 parameter saja yaitu kategori cuaca dari citra tersebut. Pada skenario kedua ini juga akan menggunakan sebanyak 20 citra per kelasnya. Selanjutnya pada pengujian skenario ketiga akan menggunakan citra dari dataset yang digunakan yaitu sebanyak 120 citra per kelasnya dan akan membandingkan 6 parameter yaitu kategori cuaca dan 5 fitur numeriknya.

Tabel 4 Banyak Data yang Digunakan Pada Pengujian Model

	Skenario 1	Skenario 2	Skenario 3
Cerah	20	20	120
Berawan	20	20	120
Mendung	20	20	120
Hujan	20	20	120

Tabel 5 Perbandingan Akurasi Cuaca Hasil Pengujian Setiap Skenario

	Skenario 1	Skenario 2	Skenario 3
Akurasi	0.6875	0.5	0.9563
Precision	0.6775	0.6010	0.9566
Recall	0.6875	0.5	0.9563
F1 - Score	0.6772	0.4199	0.9563

Tabel 6 Perbandingan MSE Untuk Data Numerik Hasil Pengujian Skenario 1 dan 3

	Skenario 1	Skenario 3
Suhu	5.0195	0.6170
Kelembaban	106.7317	16.4949
Kecepatan angin	52.0801	6.0451
Sinar UV	1.6475	0.3131
Tekanan Udara	2.4625	0.5372

4. KESIMPULAN

Dari hasil yang telah diperoleh dari perancangan sistem pengenalan cuaca Indonesia menggunakan citra langit, dapat ditarik kesimpulan bahwa sistem telah berhasil untuk berjalan dengan sesuai rancangan, dimana akan menerima input berupa gambar atau citra langit yang akan digunakan untuk mendapatkan output berupa hasil klasifikasi dari 4 kategori cuaca yaitu cerah, berawan, mendung atau hujan, disertai dengan prediksi suhu, kelembaban udara, kecepatan angin, sinar UV, dan tekanan udara untuk mendukung hasil dari pengenalan cuaca. Dari hasil pelatihan model dengan menggunakan 4 konfigurasi *epochs*, didapat model terbaik yaitu dengan *epochs* 700 dimana akurasi validasi mencapai di 78% dengan MSE validasi untuk data numerik 0.6478, dan total *loss* validasi 1.547. Namun, model ini masih mengalami kesulitan dalam menggeneralisasi data baru seperti pada hasil pengujian yang diperoleh pada skenario pertama dan kedua dimana pada skenario pertama menggunakan citra kamera mendapat akurasi sebesar 68% dan pada skenario kedua menggunakan citra internet mendapat akurasi sebesar 50%. Sedangkan pada skenario ketiga yang menggunakan citra *dataset* mendapat akurasi sebesar 95% yang menunjukkan bahwa model telah mempelajari fitur dan pola yang ada dari *dataset* yang digunakan pada pelatihan.

Adapun saran yang dapat diberikan yaitu akurasi dari model baik untuk akurasi validasi maupun akurasi pengujian akan dapat menjadi lebih akurat dimana jika ukuran *dataset* diperbesar dengan menggunakan gambar atau citra langit yang lebih bervariasi. Pengambilan citra langit juga dapat dilakukan dalam periode waktu yang lebih panjang dan dilakukan di beberapa lokasi sehingga kondisi langit akan dapat lebih beragam sehingga sistem akan dapat melakukan pengenalan cuaca secara akurat tidak hanya pada satu daerah tertentu saja. Dengan ukuran *dataset* yang lebih besar, model akan dapat lebih mengenali fitur atau pola yang ada pada citra sehingga hasil klasifikasi dan prediksi akan lebih akurat.

DAFTAR PUSTAKA

- [1] Weather, Climate & Seasons, <https://www.indonesia.travel/gb/en/general-information/climate.html>, diakses tanggal 20 Agustus 2024.
- [2] Guide, A., 2007, *Understanding Weather and Climate*
- [3] Sharma, N., Jain, V., dan Mishra, A., 2018, An Analysis of Convolutional Neural Networks for Image Classification, *Procedia Computer Science*, Vol. 132, Hal. 377-384.
- [4] Vallabhaneni, R., Vaddadi, S. A., Pillai, S. E. V. S., Addula, S. R., dan Ananthan, B., 2024, MobileNet Based Secured Compliance Through Open Web Application Security Projects in Cloud System, *Indonesian Journal of Electrical Engineering and Computer Science*, Vol. 35, No. 3, Hal. 1661-1669.
- [5] Darula, S., dan Kittler, R., 2015, Classification of Daylight Conditions in Cloud Cover Situations, *Light & Engineering*, Vol. 23, No. 1, Hal. 4-14.
- [6] Rutledge, K., McDaniel, M., Teng, S., Hall, H., Ramroop, T., Sprout, E., Hunt, J., Boudreau, D., dan Costa, H., 2023, Rain, <https://education.nationalgeographic.org/resource/rain/>, diakses tanggal 20 Agustus 2024.
- [7] Indolia, S., Goswami, A. K., Mishra, S. P., dan Asopa, P., 2018, Conceptual Understanding of Convolutional Neural Network—A Deep Learning Approach, *Procedia Computer Science*, Vol. 132, Hal. 679-688.
- [8] Arfan, T. H., Hayaty, M., dan Hadinegoro, A., 2021, Classification of Brain Tumours Types Based on MRI Images Using MobileNet, *Proceeding of 2nd International Conference on Innovative and Creative Information Technology (ICITech)*, Hal. 69-73.
- [9] Hastomo, W., dkk., 2021, Convolution Neural Network Arsitektur Mobilenet-V2 Untuk Mendeteksi Tumor Otak, *Prosiding Seminar SeNTIK*, Vol. 5, No. 1, Hal. 17–21.
- [10] Li, Z., Liu, F., Yang, W., Peng, S., dan Zhou, J., 2021, A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects, *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 33, No. 12, Hal. 6999-7019.