

LAPORAN AKHIR
PENELITIAN DASAR DAN TERAPAN



Klasifikasi Cuaca dengan Metode Convolutional Neural Network (CNN)

Oleh :

Ketua Peneliti : Ir. Rita Magdalena, M.T.
Anggota Peneliti : Ir. Jaspar Hasudungan, M.T.
: Raditiana Patmasari, S.T., M.T.
: Sean Alexander Suryaman

FAKULTAS TEKNIK ELEKTRO
UNIVERSITAS TELKOM
BANDUNG
AGUSTUS
2021

HALAMAN PENGESAHAN LAPORAN AKHIR

- | | |
|------------------------------|---|
| 1. Judul Penelitian | Klasifikasi Cuaca dengan Metode Convolutional Neural Network (CNN) |
| 2. Ketua Peneliti / Pengusul | Ir. Rita Magdalena, M.T |
| NIP/NIDN | 99640034 /0431086402 |
| Telp / Email | +628122361091/ ritamagdalena@telkomuniversity.ac.id |
| Jab. Fungsional / Struktural | Lektor |
| Jurusan / Fakultas | S1 Teknik Telekomunikasi / Fakultas Teknik Elektro |
| Bidang Keahlian | Pengolahan Sinyal Informasi |
| 3. Anggota Peneliti DOsen | 1. Ir. Jaspas Hasudungan, M.T.
2. Raditiana Patmasari, S.T., M.T. |
| Anggota Peneliti Mahasiswa | 1. Sean Alexander Suryaman |
| 4. Jadwal | 6 bulan |
| 5. Rencana Luaran | Publikasi Ilmiah |
| 6. Pembiayaan | Rp. 6.000.000,- |

Bandung, Agustus 2021

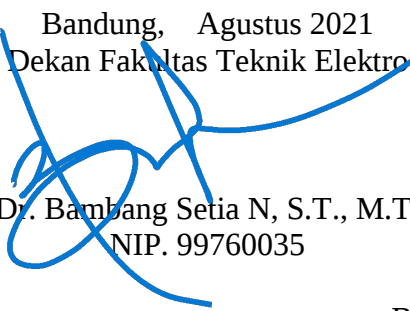
Ketua Peneliti



Ir. Rita Magdalena, M.T
NIP. 99640034

Menyetujui

Bandung, Agustus 2021
Dekan Fakultas Teknik Elektro



Dr. Bambang Setia N, S.T., M.T.
NIP. 99760035

Bandung, Agustus 2021
Ketua Kelompok Keilmuan,



Ir. Rita Magdalena, M.T
NIP. 99640034

Bandung, Agustus 2021
Direktur LPPM

Kemas Muslim Lhaksana, S.T., M.ISD., Ph.D.
NIP. 13820075

DAFTAR ISI

HALAMAN PENGESAHAN	2
DAFTAR ISI	3
RINGKASAN.....	5
BAB I PENDAHULUAN.....	6
BAB II TINJAUAN PUSTAKA	2
2.1. Cuaca	2
2.2 Convolutional Neural Network.....	3
2.3 Parameter Performansi Sistem	4
BAB II METODOLOGI PENELITIAN	2
3.1. Gambaran Umum.....	6
3.2. Dataset	6
3.3 Convolutional Layer	7
3.4 Max Pooling.....	7
3.5 Dropout.....	8
3.6 Fully Connected Layer	8
3.7 Activation Function	8
BAB IV BIAYA DAN JADWAL PELAKSANAAN.....	9
4.1 Anggaran Biaya.....	Error! Bookmark not defined.
4.2 Jadwal Penelitian	Error! Bookmark not defined.
DAFTAR PUSTAKA	12
LAMPIRAN 1. BIODATA PENELITI	13
LAMPIRAN 2. SURAT PERNYATAAN KETUA PENELITI.....	Error! Bookmark not defined.
LAMPIRAN 2.....	Error! Bookmark not defined.

RINGKASAN

Cuaca merupakan suatu fenomena alam yang sangat berdampak bagi manusia. Cuaca yang ekstrem dapat menyebabkan bencana alam yang sangat merugikan manusia. Informasi tentang kondisi cuaca sangat dibutuhkan oleh manusia. Informasi ini bermanfaat untuk menghindari hal - hal yang tidak diinginkan. Sistem klasifikasi saat ini mengandalkan serangkaian sensor mahal atau bantuan manusia. Oleh karena itu, dibutuhkan suatu sistem yang dapat mengatasi masalah tersebut.

Dalam penelitian ini, desain sistem untuk klasifikasi cuaca menggunakan metode *Convolutional Neural Network* (CNN) yang memiliki keunggulan untuk menghasilkan fitur dan karakteristik dari dataset sekunder yang digunakan. Dataset sekunder yang akan digunakan diambil dari *kaggle.com*. Model yang diusulkan CNN dalam penelitian ini terdiri dari 3 hidden layer yang menggunakan filter ukuran 3x3.

Model yang diusulkan akan digunakan untuk mengklasifikasikan cuaca menjadi cerah, berawan dan hujan. Parameter performansi system yang terbaik yaitu akurasi 86%, presisi 0.86, recall 0.86 dan f-1 score 0.86 dengan parameter terbaik yang digunakan adalah optimizer RMSProp, resolusi citra 224x224, *learning rate* 0,0001 dan epoch 50.

Kata Kunci: *Cuaca, CNN, dataset sekunder, akurasi, recall, precision, f-1 score*

BAB I

PENDAHULUAN

Semakin berkembangnya peradaban manusia, kemajuan teknologi pun semakin berkembang. Dengan kemajuan teknologi yang semakin pesat ini, pekerjaan-pekerjaan manusia semakin dimudahkan. Kemajuan teknologi di bidang Informasi yang sangat pesat belakangan ini, membantu kita dalam melakukan kegiatan sehari-hari. Salah satu informasi yang kita butuhkan adalah tentang cuaca. Keadaan cuaca yang tidak menentu belakangan ini membuat kita khawatir mengenai keselamatan. Cuaca yang ekstrim dapat menyebabkan bencana alam yang menyebabkan kerugian material dan menghilangkan nyawa manusia. Selain itu, informasi tentang kondisi cuaca juga sangat kita butuhkan ketika kita akan berpergian ke suatu tempat untuk menghindari hal-hal yang tidak diinginkan. Oleh karena itu, dibutuhkan suatu alat untuk memberikan informasi mengenai cuaca.

Citra atau gambar merupakan salah satu media yang sangat sering digunakan dalam teknologi modern. Contoh pengaplikasian citra ada pada sistem pendeteksi hujan es, sistem pendeteksi asap, debu vulkanik dan lain sebagainya. Alat klasifikasi cuaca yang ada saat ini menggunakan sensor kelembapan memiliki beberapa kekurangan yakni dari segi sistem serta biaya yang dikeluarkan. Karena itu pemanfaatan citra dalam klasifikasi cuaca dapat menjadi alternatif untuk sistem klasifikasi cuaca yang lebih efektif dan efisien.

Terdapat banyak penelitian yang memanfaatkan citra untuk sistem pendeteksi. Penelitian [1] mengidentifikasi kejadian hujan es menggunakan citra radar dan satelit cuaca menggunakan metode *Severe Hail Index* (SHI) menghasilkan bahwa hujan es dapat dideteksi menggunakan metode *Severe Hail Index* (SHI) akan tetapi untuk daerah yang dekat dengan radar, potensi hujan es tidak dapat dideteksi. Penelitian [2] mendeteksi asap akibat kebakaran hutan menggunakan RGB *false color* pada satelit Himawari 8 menghasilkan keluaran yang sama dengan penelitian pada satelit MODIS Terra-Aqua. Penelitian [3] mendeteksi sebaran debu vulkanik dapat dengan baik mendeteksi sebaran debu

vulkanik dan pergerakannya secara spasial dan *time-series*. Penelitian [4] mendeteksi debu vulkanik menggunakan metode Volcanic Ash Product (VOL)- EUMETSAT hanya dapat mendeteksi debu vulkanik secara temporal. Penelitian [5] Penerapan algoritma klasifikasi data mining C4.5 pada dataset cuaca wilayah Bekasi menghasilkan akurasi sebesar 88,89%. Melalui penelitian-penelitian tersebut penulis akan melakukan penelitian klasifikasi cuaca metode *Convolutional Neural Network* (CNN) yang memiliki keunggulan untuk menghasilkan fitur dan karakteristik dari dataset sekunder yang digunakan.

BAB II

TINJAUAN PUSTAKA

2.1. Cuaca

Cuaca adalah keadaan atmosfer di suatu tempat pada waktu tertentu yang berkaitan dengan suhu udara, sinar matahari, angin, hujan dan kondisi udara lainnya. Ilmu yang mempelajari cuaca disebut meteorology. Cuaca berbeda dengan iklim. Iklim adalah suhu rata-rata udara dalam waktu lama pada daerah yang sangat luas. Ilmu yang mempelajari iklim disebut klimatologi. Cuaca bisa panas atau dingin, basah atau kering, berangin atau tidak berangin. Cuaca disebabkan oleh perubahan udara di sekeliling bumi saat udara memanaskan atau mendingin.

Cuaca dibagi menjadi tiga kategori, yaitu cuaca cerah, berawan dan hujan. Cuaca cerah adalah cuaca yang menunjukkan langit dalam kondisi terang, sinar matahari memancar terang tetapi tidak begitu panas, terdapat awan yang berlapis-lapis tipis seperti bulu-bulu serat sutra halus. Cuaca berawan adalah cuaca yang menunjukkan bahwa di langit banyak terdapat awan. Awan merupakan kumpulan uap air yang terdapat di udara. Uap air ini berasal dari air sungai, air laut, air danau serta air kolam yang naik ke atas dan bergabung dengan udara karena pengaruh panas matahari. Cuaca hujan adalah kondisi ketika udara akan naik keatas dan membentuk awan. Makin keatas, suhu uap air menjadi makin rendah. Pada suhu tertentu, uap air akan mengembun menjadi titik-titik air. Titik-titik air akan berubah menjadi tetes-tetes air. Makin lama tetes-tetes air itu makin berat dan akhirnya jatuh ke bumi

Cuaca yang terjadi di pagi hari yang cerah, radiasi matahari dengan cepat memanaskan daratan bumi kemudian memanaskan udara. Sedangkan cuaca yang terjadi di malam hari, tidak ada radiasi matahari sehingga terjadi pelepasan panas dari bumi yang besar sehingga mendinginkan permukaan daratan dan udara[7]. Citra cuaca berawan, hujan, dan cerah ditunjukkan pada Gambar 2.1.

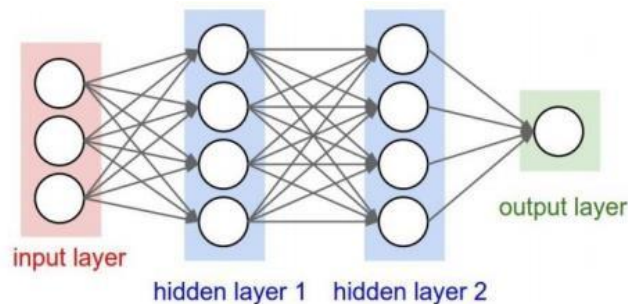


Gambar 2.1 *Dataset cuaca berawan (a), hujan (b), cerah (c)*

2.2 Convolutional Neural Network

Convolutional Neural Network (CNN) merupakan sebuah pengembangan dari Multilayer Perceptron (MLP) yang didesain untuk mengolah data dua dimensi dalam bentuk citra. CNN termasuk salah satu metode Deep Learning yang bisa digunakan untuk mendeteksi serta mengenali suatu objek pada sebuah citra digital. Deep Learning merupakan implementasi konsep dasar dari Machine Learning yang menerapkan algoritma Artificial Neural Network (ANN) dengan layer yang lebih banyak. Metode MLP kurang sesuai untuk digunakan dalam klasifikasi citra karena MLP tidak menyimpan informasi spasial dari data sebuah citra dan menganggap setiap piksel merupakan fitur yang independen sehingga menghasilkan hasil yang kurang baik.

Cara kerja CNN sama seperti pada MLP, namun pada CNN setiap neuron dipresentasikan dalam bentuk dua dimensi, berbeda dengan MLP yang setiap neuron hanya berbentuk satu dimensi.



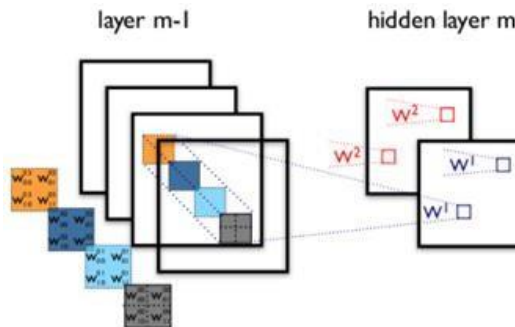
Pada Gambar 1 sebuah MLP mempunyai i layer (kotak merah dan biru) dengan masing-masing layer berisi j_i neuron (lingkaran putih). MLP menerima input data satu dimensi dan mempropagasikan data tersebut pada jaringan hingga menghasilkan output. Setiap hubungan antar neuron pada dua layer yang bersebelahan memiliki parameter bobot satu dimensi yang menentukan kualitas mode. Disetiap data inout pada layer dilakukan operasi linear dengan nilai bobot yang ada, kemudian hasil komputasi akan ditransformasi menggunakan operasi non linear yang disebut fingsi aktivasi.

Pada CNN, data yang dipropagasikan oleh jaringan adalah data dua dimensi, sehingga operasi linear dan parameter bobot pada CNN berbeda. Pada CNN operasi linear menggunakan konvolusi, sedangkan bobot tidak lagi satu dimensi saja, namun berbentuk empat dimensi yang merupakan kumpulan kernel konvolusi seperti pada Gambar 2.

Dimensi bobot pada CNN adalah:

$$\text{Neuron input} \times \text{neuron output} \times \text{tinggi} \times \text{lebar}$$

Karena sifat proses konvolusi, maka CNN hanya dapat digunakan pada data yang memiliki struktur dua dimensi seperti citra dan suara.



Gambar 2. Proses Konvolusi pada CNN [3]

2.3 Parameter Performansi Sistem

Parameter yang digunakan untuk mengukur kinerja sistem dalam mengklasifikasikan jenis kulit wajah meliputi akurasi, recall, presisi, dan skor F1. Confusion matrix digunakan untuk mempresentasikan hasil pengujian sistem dalam mengklasifikasikan jenis kulit wajah kering, normal, berminyak dan kombinasi. Berikut adalah persamaan untuk mengukur performansi sistem:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN}$$

$$\text{Recall} = \frac{TP}{TP+FN}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{F1 score} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}}$$

Dimana:

TP : True Positive

TN : True Negative

FN : False Negative

FP : False Positive

BAB III

TUJUAN DAN MANFAAT PENELITIAN

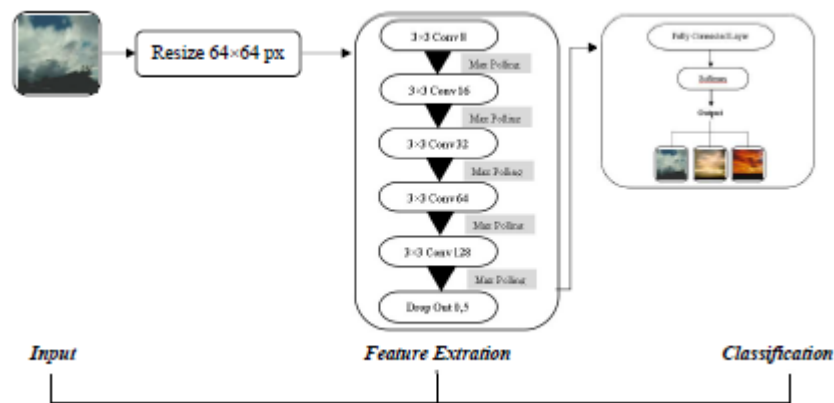
Tujuan dari penelitian ini adalah untuk mengetahui performansi dari algoritma yang diusung dalam melakukan klasifikasi terhadap cuaca serta mengetahui parameter – parameter terbaik pada algoritma yang diusung.

BAB IV

METODOLOGI PENELITIAN

3.1. Gambaran Umum

Dalam studi ini, sebuah sistem dirancang untuk mengklasifikasikan jenis cuaca yaitu berawan, cerah, panas dan hujan menggunakan metode CNN. Model CNN yang digunakan terdiri dari 3 hidden layer, di mana dalam tiga hidden layer ukuran filter 3x3 digunakan dengan output channel masing-masing 8, 16 dan 32, fully connected layer dan aktivasi softmax. Secara umum, Gambar 3.1 menunjukkan model CNN yang diusulkan dalam penelitian ini.



Gambar 3.3 Model Sistem CNN Usulan

Berdasarkan Gambar 3.3 model sistem CNN usulan tersebut akan digunakan pada proses klasifikasi cuaca. Pada proses input terdapat tiga kelas citra cuaca yang terdiri dari berawan, hujan, dan cerah, dengan data latih sejumlah 75% dan data uji 25%. Ukuran citra cuaca yang tidak sama rata akan disamakan ukurannya menjadi 64x64 pixel. Hasil scan data citra terdiri dari channel Grayscale.

3.2. Dataset

Dataset yang akan digunakan merupakan dataset sekunder berupa citra RGB yang diambil dari *Kaggle.com* dengan format citra JPG sejumlah 768 citra.

3.3 Convolutional Layer

Lapisan pertama dalam model CNN adalah convolutional layer, di mana operasi linier diproses melalui konvolusi yang memiliki empat dimensi bentuk bobot, yang merupakan kumpulan kernel konvolusi. Dimensi berat pada CNN adalah:

$$\text{Input neuron} \times \text{output neuron} \times \text{tinggi} \times \text{berat} \quad (1)$$

Karena sifat proses konvolusi, CNN hanya dapat digunakan pada data yang memiliki struktur dua dimensi seperti gambar. Tujuan konvolusi pada data gambar adalah untuk mengekstraksi fitur dari gambar input. Konvolusi akan menghasilkan transformasi linier dari data input sesuai dengan informasi khusus dalam data. Berat lapisan menentukan kernel konvolusi yang digunakan, sehingga kernel konvolusi dapat dilatih berdasarkan input pada CNN. Lapisan ekstraksi fitur terdiri dari lapisan konvolusional dan lapisan penggabungan. Proses konvolusi dilakukan untuk mengubah gambar untuk menghasilkan fitur terkonvolusi yang merupakan karakter unik dari gambar asli.

3.4 Max Pooling

Setelah layer konvolusi, maka ada layer pooling. Pooling adalah pengurangan ukuran matriks dengan menggunakan operasi pooling. Pada dasarnya layer Pooling terdiri dari filter dengan ukuran dan langkah tertentu yang akan bergiliran di seluruh area peta fitur. Bentuk lapisan penyatuan umumnya menggunakan filter 2x2 yang diterapkan dalam dua langkah dan beroperasi pada setiap irisan input.

Penggunaan layer pooling pada CNN adalah untuk mengurangi ukuran gambar sehingga dapat dengan mudah diganti dengan layer convolutional dengan langkah yang sama dengan layer pooling yang sesuai, dan untuk melihat gambar dalam gambar besar.

3.5 Dropout

Dropout adalah teknik yang digunakan untuk mengurangi overfitting di neural network dengan mencegah adaptasi kompleks pada data pelatihan. Dalam penelitian ini 0,5 dropout digunakan, yang berarti bahwa 50% dari unit akan dikeluarkan dari neural network.

3.6 Fully Connected Layer

Fully Connected Layer adalah lapisan di mana semua neuron aktivasi dari lapisan sebelumnya semuanya terhubung ke neuron di lapisan berikutnya seperti jaringan saraf biasa. Pada dasarnya, lapisan ini biasanya digunakan dalam MLP yang memiliki tujuan mengubah dimensi data untuk diklasifikasikan secara linear.

3.7 Activation Function

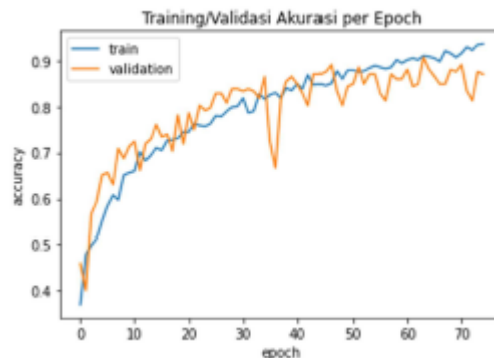
Fungsi aktivasi yang digunakan dalam model CNN yang diusulkan dalam penelitian ini adalah ReLU di lapisan tersembunyi dan softmax untuk lapisan output. Fungsi Softmax menghitung probabilitas setiap kelas target untuk semua kelas target yang mungkin dan akan membantu menentukan kelas target untuk input yang diberikan. Keuntungan utama menggunakan Softmax adalah rentang probabilitas keluaran dengan nilai dari 0 hingga 1, dan jumlah semua probabilitas akan sama dengan satu. Fungsi softmax dapat digunakan untuk model multi-klasifikasi.

BAB V

HASIL YANG DICAPAI

5.1. Hasil dan Pembahasan

Sistem klasifikasi cuaca yang dirancang pada penelitian ini dapat mengklasifikasikan jenis cuaca berawan, cerah dan hujan. Data yang digunakan pada penelitian ini merupakan dataset sekunder yang diambil pada sebuah *open web dataset* yaitu Kaggle.com . Dalam menguji performansi sistem telah Dalam menguji performansi sistem telah dilakukan training menggunakan model CNN arsitektur VGG16 yang diusulkan untuk 576 data latih dan 192 data uji. Model CNN yang diusulkan tersebut menggunakan RMSProp optimizer dengan learning rate sebesar 0,0001 serta menggunakan persamaan loss categorical crossentropy. Performansi sistem dalam mengklasifikasikan keempat kondisi kulit tersebut dievaluasi menggunakan performansi akurasi, precision, recall dan f-1 score. Berdasarkan hasil dari pelatihan model yang diusulkan dengan menggunakan nilai iterasi 50 kali (epoch 50) diperoleh hasil akurasi untuk setiap iterasi yang ditunjukkan pada Gambar 5.1



Gambar 5.1 Grafik Hasil Akurasi

Model CNN yang digunakan pada sistem klasifikasi cuaca yang telah dirancang pada penelitian ini tidak mengalami overfitting, artinya baik pada tahap pelatihan maupun pengujian model mampu mengenali setiap kondisi kulit, hal ini berdasarkan Gambar 5.1 yang menunjukkan kenaikan akurasi pada setiap iterasi serta hasil akurasi data training dan data uji yang diperoleh tidak jauh berbeda. Performansi terbaik sistem diperoleh akurasi sebesar 86 %.

BAB VI

RENCANA TAHAP BERIKUTNYA

Penelitian ini berhasil mendeteksi cuaca kedalam tiga kelas, yaitu cerah, hujan dan berawan. Untuk selanjutnya pengembangan akan dilakukan pada aplikasi Android untuk memudahkan user dalam melakukan pendeteksian jenis cuaca.

BAB VII

KESIMPULAN

Hasil penelitian yang telah dilakukan memberikan kesimpulan bahwa pengklasifikasian cuaca menggunakan Convolutional Neural Network (CNN) dengan arsitektur VGG16Net mampu mengklasifikasikan jenis cuaca berdasarkan kelasnya. Parameter terbaik didapatkan berdasarkan penggunaan optimizer RMSprop, dengan ukuran citra 224x224, Epoch 50, dan learning rate 0,0001 dengan nilai performansi sistem akurasi, presisi, recall, dan f1-score masing-masing sebesar 86%, 86%, 86%, dan 86%.

DAFTAR PUSTAKA

- [1] W. Suryanto and A. Luthfian, "PENGANTAR METEOROLOGI," Yogyakarta, UGM PRESS, 2019, p. 10.

- [2] S. Gischa, "Kompas.com," Kompas.com, 2021. [Online]. Available: <https://www.kompas.com/skola/read/2021/03/01/191900169/kondisi-cuaca-danpengaruhnya-terhadap-kegiatan-manusia?page=all> . [Accessed 22 March 2021].

- [3] S. Ernawati, "APLIKASI HOPFIELD NEURAL NETWORK UNTUK PRAKIRAAN CUACA," Jurnal Meteorologi dan Geofisika, vol. 10, no. 2, 2009.

- [4] I. G. A. Gunadi and A. Aprilyana Kusuma Dewi, "Klasifikasi Curah Hujan di Provinsi Bali Berdasarkan Metode Naïve Bayesian," Wahana Matematika dan Sains: Jurnal Matematika, Sains, dan Pembelajarannya, vol. 12, no. 1, 2018.

- [5] J. An, Y. Chen and H. Shin, "Weather Classification using Convolutional Neural Networks," in International SoC Design Conference (ISOCC), Daegu, 2018.

LAMPIRAN 1. PUBLIKASI

WEATHER CLASSIFICATION USING PRINCIPAL COMPONENT ANALYSIS (PCA) AND K-NEAREST NEIGHBOR (KNN) METHOD

Sean Alexander Suryaman¹, Rita Magdalena², Sofia Sa'idah³

^{1,2,3}Prodi S1 Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom, Bandung
Email: ¹seanalexanders@student.telkomuniversity.ac.id, ²ritamagdlena@telkomuniversity.ac.id,
³sofiansaidah@telkomuniversity.ac.id

(Naskah masuk: dd mmum yyyy, diterima untuk diterbitkan: dd mmum yyyy)

Abstract

Weather is a natural phenomenon that has an impact on humans. Extreme weather can cause natural disasters that are very detrimental to humans. Information about weather conditions is needed by humans. This information is very useful for living everyday life. Current classification systems rely on a series of expensive sensors or human assistance. This study classifies weather conditions using the Principal Component Analysis (PCA) and K-Nearest Neighbor (KNN) methods. This process utilizes Principal Component Analysis (PCA) to reduce data to be more effective. And using K-Nearest Neighbor (KNN) to classify the data. K-Nearest Neighbor (KNN) uses distance to classify data. The selected distance is the shortest distance that will show neighbors to produce output whether the weather is sunny, cloudy, foggy, rainy and sunrise. Based on the results of the study, the weather classification system was obtained using the Principal Component Analysis (PCA) and K-Nearest Neighbor (KNN) methods with an accuracy of 87.50%. The accuracy results were obtained when 450 test data and 1050 training data were used. The best parameters generated, namely the image size 256 x 256, the type of KNN is Cosine, the value of KNN is at $k = 9$, and the PCA percentage is 30%.

Keywords: *k-nearest neighbor (KNN) , principal component analysis (PCA), weather classification*

KLASIFIKASI CUACA MENGGUNAKAN METODE PRINCIPAL COMPONENT ANALYSIS (PCA) DAN K-NEAREST NEIGHBOR (KNN)

Abstrak

Cuaca merupakan suatu fenomena alam yang sangat berdampak bagi manusia. Cuaca yang ekstrem dapat menyebabkan bencana alam yang sangat merugikan manusia. Informasi tentang kondisi cuaca sangat dibutuhkan oleh manusia. Informasi ini sangat bermanfaat untuk menjalani kehidupan sehari – hari. Sistem klasifikasi saat ini mengandalkan serangkaian sensor mahal atau bantuan manusia. Penelitian ini mengklasifikasi kondisi cuaca menggunakan metode *Principal Component Analysis (PCA)* dan *K-Nearest Neighbor (KNN)*. Proses ini memanfaatkan *Principal Component Analysis (PCA)* untuk mereduksi data agar lebih efektif. Dan menggunakan *K-Nearest Neighbor (KNN)* untuk mengklasifikasi data. *K-Nearest Neighbor (KNN)* menggunakan jarak untuk mengklasifikasi data. Jarak yang dipilih merupakan jarak terpendek yang akan menunjukkan ketertarikan untuk menghasilkan keluaran apakah cuaca sedang cerah, berawan, berkabut, hujan dan matahari terbit. Berdasarkan hasil penelitian, diperoleh sistem klasifikasi cuaca dengan metode *Principal Component Analysis (PCA)* dan *K-Nearest Neighbor (KNN)* dengan akurasi sebesar 87,50%. Hasil akurasi tersebut diperoleh ketika digunakan 450 data uji dan 1050 data latih. Adapun parameter terbaik yang dihasilkan, yaitu ukuran citra 256 x 256, jenis KNN adalah *Cosine*, nilai KNN di $k = 9$, dan Persentase PCA 30%.

Kata kunci: *k-nearest neighbor(KNN), klasifikasi cuaca, principal component analysis(PCA)*

1. PENDAHULUAN

Semakin berkembang peradaban manusia, kemajuan teknologi pun semakin berkembang. Dengan kemajuan teknologi yang semakin pesat ini, pekerjaan-pekerjaan manusia semakin di mudahkan.

Kemajuan teknologi di bidang Informasi yang sangat pesat belakangan ini, membantu kita dalam melakukan kegiatan sehari-hari. Salah satu informasi yang kita butuhkan adalah tentang cuaca. Kondisi cuaca yang tidak menentu belakangan ini membuat kita khawatir mengenai keselamatan. Cuaca yang

ekstrem dapat menyebabkan bencana alam yang menyebabkan kerugian material dan menghilangkan nyawa manusia. Selain itu, informasi tentang kondisi cuaca juga bermanfaat dalam menjalani kehidupan. Oleh karena itu, dibutuhkan suatu alat untuk memberikan informasi mengenai cuaca.

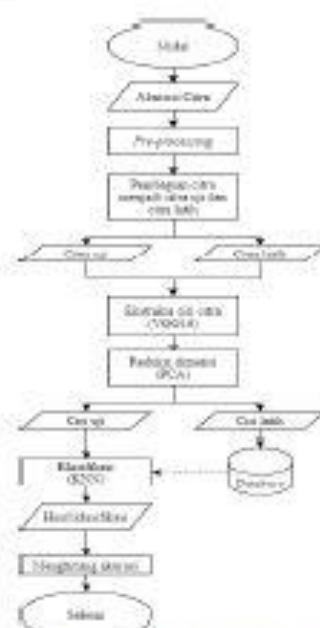
Citra atau gambar merupakan salah satu media digunakan dalam teknologi modern. Contoh pengaplikasian citra ada pada sistem pendeteksi hujan es, sistem pendeteksi asap, debu vulkanik dan lain sebagainya. Alat klasifikasi cuaca yang ada saat ini menggunakan sensor kelembapan memiliki beberapa kekurangan yakni dari segi sistem serta biaya yang dikeluarkan. Karena itu pemanfaatan citra dalam klasifikasi cuaca dapat menjadi alternatif untuk sistem klasifikasi cuaca yang lebih efektif dan efisien.

Terdapat banyak penelitian yang memanfaatkan citra untuk sistem pendeteksi. Penelitian [1] mengidentifikasi kejadian hujan es menggunakan citra radar dan satelit cuaca menggunakan metode *Severe Hall Index* (SHI) menghasilkan bahwa hujan es dapat di deteksi menggunakan metode *Severe Hall Index* (SHI) akan tetapi untuk daerah yang dekat dengan radar, potensi hujan es tidak dapat dideteksi. Penelitian [2] mendeteksi asap akibat kebakaran hutan menggunakan RGB *false color* pada satelit Himawari 8 menghasilkan ketepatan yang sama dengan penelitian pada satelit MODIS Terra-Aqua. Penelitian [3] mendeteksi sebaran debu vulkanik dapat dengan baik mendeteksi sebaran debu pergerakannya. Penelitian [4] mendeteksi debu vulkanik menggunakan metode *Volcanic Ash*

mendeteksi debu vulkanik secara temporal. Penelitian [5] *Deteksi sebaran debu vulkanik dengan metode PCA* pada citra satelit cuaca menghasilkan akurasi sebesar 88,89%. Melalui penelitian-penelitian tersebut penulis melakukan penelitian klasifikasi cuaca menggunakan ekstraksi ciri *Principal Component Analysis* (PCA) dan metode klasifikasi *K-Nearest Neighbor* (KNN) yang dapat diukur.

2. METODE PENELITIAN

Langkah pertama pada penelitian ini yaitu akuisisi citra yang merupakan proses pengumpulan citra cuaca yang akan digunakan dalam penelitian ini. Citra yang didapatkan tidak memiliki ukuran yang sama oleh karena itu dilakukan resize. Proses ini dinamakan *pre-processing*. Lalu langkah berikutnya adalah ekstraksi ciri. Pada penelitian ini ekstraksi ciri dilakukan dengan VGG16 dan PCA. Setelah citra memiliki ciri, maka citra akan dibagi menjadi citra uji dan citra latih. Citra latih akan menjadi database untuk proses klasifikasi pada citra uji, yang pada penelitian ini proses klasifikasi menggunakan *K-Nearest Neighbor*.



Gambar 1. Diagram Alir

3. HASIL DAN PEMBAHASAN

Bagian ini berisi hasil pengujian sistem terhadap beberapa skenario untuk mendapatkan akurasi terbaik. Adapun parameter yang ditinjau adalah ukuran citra, persentase PCA, jenis KNN, dan nilai k.

3.1. Skenario Pengujian

Pada proses pengujian akan dilakukan terhadap 1464 citra yang terbagi menjadi citra *cloudy*, citra *foggy*, citra *sunrise*. Dari 1464 citra tersebut, akan dibagi menjadi 70% citra latih dan 30% citra uji. Setelah itu dilakukan pencarian parameter-parameter dengan akurasi terbaik untuk diterapkan kedalam sistem.

1. Skenario 1: Pengujian ukuran citra

Skenario pertama ini akan dilakukan pengujian terhadap ukuran citra dengan cara merubah ukuran citra menjadi 64 x 64, 128 x 128, dan 256 x 256. Parameter lain yang digunakan adalah jenis KNN = *Costine*, nilai KNN di k = 3, dan Persentase PCA 10%. Dari hasil pengujian tersebut dapat diketahui bahwa ukuran citra yang paling baik digunakan adalah 256x256 dengan akurasi 81,36% ini disebabkan karena semakin besar sebuah citra maka jumlah informasi yang kita dapatkan akan semakin banyak.

2. Skenario 2: Pengujian persentase PCA

Pengujian kedua akan mencari persentase PCA terbaik antara 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%. Parameter lain yang digunakan adalah jenis KNN = *Costine*, ukuran citra = 256x256, dan nilai KNN di k = 3. Dari hasil pengujian tersebut dapat diketahui bahwa persentase PCA yang paling baik digunakan adalah 30% dengan akurasi sebesar 85,00%. Hal ini merandakan bahwa ada ciri yang tidak tepat pada saat proses ekstraksi ciri. Ciri ini

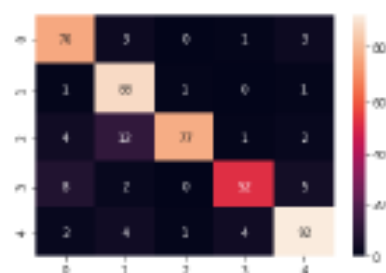
lah yang membuat persentase menurun ketika dipakai. Sehingga ciri yang optimal dipakai hanya 30% dari keseluruhan ciri.

3. Skenario 3: Pengujian jenis KNN
Skenario ketiga akan dilakukan pengujian terhadap jenis KNN yaitu Cityblok, Cosine dan Euclidean. Parameter lain yang digunakan adalah ukuran citra = 256 x 256, nilai KNN di $k = 3$, dan Persentase PCA 30%. Dari hasil pengujian tersebut dapat diketahui bahwa jenis KNN yang paling baik digunakan adalah Cosine dengan akurasi sebesar 85,00%.
4. Skenario 4: Pengujian nilai k
Skenario keempat ini, dilakukan pengujian terhadap nilai KNN 1, 3, 5, 7, dan 9. Parameter lain yang digunakan adalah jenis KNN = Cosine, ukuran citra = 256x256, dan Persentase PCA 30%. Dari hasil pengujian tersebut dapat diketahui bahwa nilai KNN yang paling baik digunakan pada $k = 9$ dengan akurasi sebesar 87,50%. Hal ini menunjukkan bahwa semakin banyak ketetanggaan maka semakin banyak perbandingan dalam menentukan klasifikasi.

3.2. Data Hasil Pengujian Sistem

Berdasarkan hasil pengujian didapatkan parameter terbaik dengan menggunakan ukuran citra 256 x 256, persentase PCA 30%, jenis KNN menggunakan cosine, nilai $k = 9$. Melalui parameter-parameter tersebut didapatkan hasil akurasi 87,50%. Akurasi tersebut didapatkan dengan melakukan pengujian terhadap 440 data uji. Gambar 5 menunjukkan confusion matrix dengan menggunakan parameter terbaik. Angka 0 menunjukkan citra cloudy, angka 1 menunjukkan citra foggy, angka 2 menunjukkan citra rainy, angka 3 menunjukkan citra shine, angka 4 menunjukkan citra sunrise. Dari confusion matrix pada gambar 5 dapat dikatakan bahwa:

1. Sistem mendeteksi 76 citra cloudy secara benar sedangkan 7 citra diprediksi salah.
2. Sistem mendeteksi 88 citra foggy secara benar sedangkan 3 citra diprediksi salah.
3. Sistem mendeteksi 77 citra rainy secara benar sedangkan 19 citra diprediksi salah.
4. Sistem mendeteksi 52 citra shine secara benar sedangkan 15 citra diprediksi salah.
5. Sistem mendeteksi 92 citra sunrise secara benar sedangkan 11 citra diprediksi salah.



Gambar 2. Confusion Matrix

4. KESIMPULAN

Cuaca adalah fenomena alam yang sangat berdampak bagi kehidupan manusia. Cuaca yang buruk dapat menyebabkan kehilangan harta benda maupun korban jiwa. Penelitian kali ini merancang sistem klasifikasi cuaca dengan metode *Principal Component Analysis* (PCA) dan *K-Nearest Neighbor* (KNN). Setelah penelitian dilakukan, didapatkan sistem klasifikasi cuaca dengan akurasi sebesar 87,50%. Adapun parameter – parameter terbaik dari sistem tersebut adalah :

1. Ukuran citra 256 x 256, hal ini dikarenakan semakin besar ukuran citra maka semakin banyak informasi yang terdapat dalam citra tersebut.
2. Jenis KNN adalah Cosine.
3. Nilai jumlah ketetanggaan di $k = 9$.
4. Persentase PCA di 30%. Hal ini dikarenakan tidak semua ciri merupakan ciri yang berpengaruh terhadap performansi sistem.

DAFTAR PUSTAKA

- [1] A. Kristianto, I. J. A. Saragih, G. Larasati and K. Akib, "Identifikasi Kejadian Hujan Es Menggunakan Citra Radar dan Satelit Cuaca," *PROSIDING PIT KE-5 Riset KEBENCANAAN IARI UNIVERSITAS ANDALAS*, pp. 349-362, 2018.
- [2] B. S. Pandjaitan and A. Panjaitan, "Pemanfaatan Data Satelit Cuaca Generasi Baru Himawari-8 Untuk Mendeteksi Asap Akibat Kebakaran Hutan dan Lahan di Wilayah Indonesia (Studi Kasus: Kebakaran Hutan dan Lahan di Pulau Sumatera dan Kalimantan Pada Bulan September 2015)," *SEMINAR NASIONAL PENGINDERAAN JAUH*, pp. 636-651, 2015.
- [3] P. Aditya, I. J. A. Saragih, M. P. Rosyady and A. Kristianto, "Deteksi Sebaran Debu Vulkanik Menggunakan Citra Satelit Himawari-8 (Studi Kasus: Gunung Ramug, Gunung Rinjani, and Dan Gunung Bromo)," *Seminar Nasional Penginderaan Jauh ke-5*, pp. 711-715, 2018.
- [4] M. Dwipayana, R. H. Jatmiko and M. Kamal, "Pemanfaatan Citra Satelit Himawari-8 Untuk Deteksi Debu Vulkanik Menggunakan Metode Volcanic Ash Product (VOL)-EUMETSAT (Studi Kasus: Erupsi Gunung Agung Tanggal 2 Juli 2018)," *Seminar Nasional Penginderaan Jauh ke-6 Tahun*, pp. 275-282, 2019.
- [5] A. Novandya, "Penerapan Algoritma Klasifikasi Data Mining C4.5 Pada Dataset Cuaca Wilayah Bekasi," *Kongres Nasional Ilmu Sosial & Teknologi (KNIST)*, pp. 368-372, 2017.
- [6] D. M. Wonohadidjojo, "Perbandingan Convolutional Neural Network pada Transfer

- Learning Method untuk Mengklasifikasi
Sel Darah Putih," *Ultimatrix*, vol. 13, pp. 51-
57, 2021.
- [7] D. Putra, *Sistem Biometrika: Konsep Dasar,
Teknik Analisis Citra dan Tahapan
Membangun Aplikasi Geometrika*,
Yogyakarta: ANDI, 2008.
- [8] A. A. Pratiwadi, I. K. Fithriyanti, N. Iriawan,
S. W. Purnama, M. Almahyar, W. Farisatuti
and T. Azmi, "UNet-VGG16 with transfer
learning for MRI-based brain tumor
segmentation," *TELKOMNIKA*, vol. 18, pp.
1310-1318, 2018.
- [9] A. Kadir and A. Susanto, *Teori dan Aplikasi
Pengolahan Citra*, Yogyakarta: Andi, 2013.
- [10] R. D. Jayanti, "Identifikasi Biometrik Rugas
Palatina Pada Individu Berdasarkan Bentuk
Dengan Metode Watershed dan KNN," 2017.
- [11] F. A. Haraswati, *Pengolahan Citra Digital:
Konsep dan Teori*, Yogyakarta: Andi
Publisher, 2013.
- [12] M. Harry and A. M. Purnomo, *Konsep
Pengolahan Citra Digital dan Ekstraksi Fitur*,
Yogyakarta: Graha Ilmu, 2010.
- [13] C. Dewi and M. Muslikh, "Perbandingan
Algoritma Backpropagation Neural Network dan
ANFIS Untuk Memprediksi Cuaca," *Journal
of Scientific Modeling & Computation*, pp. 7-
13, 2013.
- [14] M. B. Adinegara, "Morfologi Citra untuk
Analisa Kualitas Beras Menggunakan Metode
Klasifikasi K-Nearest Neighbor Berbasis
Android," 2018.
- [15] W. Q. Adhikari, "Deteksi Zat Narkotika
Berdasarkan Teksut Menggunakan Metode
Fractal Dimension dan Klasifikasi SVM,"
2019.

LAMPIRAN 2. LOGBOOK

**Buku Catatan Harian
(log book)**



**Program Penelitian Internal
Penelitian Dasar dan Terapan**

**Klasifikasi Cuaca dengan Metode Convolutional Neural Network
(CNN)**

Oleh :

Ketua Peneliti : Ir. Rita Magdalena, M.T.

Anggota Peneliti : Ir. Jaspar Hasudungan, M.T.

: Raditiana Patmasari, S.T., M.T.

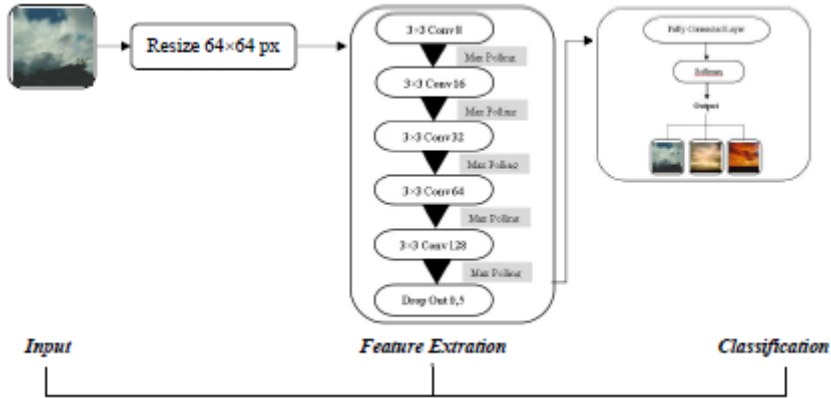
: Sean Alexander Suryaman

FAKULTAS TEKNIK ELEKTRO

Keterangan Penelitian

Judul Penelitian	: Klasifikasi Cuaca dengan Metode Convolutional Neural Network (CNN)
Institusi Peneliti	: Universitas Telkom
Kelompok Keahlian	: Pengolahan Sinyal Informasi
Fakultas	: Teknik Elektro
Bidang Fokus	: Pengolahan Sinyal Informasi
Tahun pelaksanaan	: 2021
Biaya	: Rp 6.000.000
Tujuan Penelitian	: Mendeteksi jenis cuaca apakah cuaca cerah, hujan atau berawan

Resume Catatan Kemajuan Penelitian

No	Tanggal	Aktivitas	Catatan Kemajuan *)
1	31 Maret 2021	Akuisisi Data Citra Cuaca	 (a) (b) (c)
2	1 – 30 April 2021	Perancangan Model CNN	 <i>Input</i> <i>Feature Extraction</i> <i>Classification</i>
3	1-31 Mei 2021	Penyempurnaan Model CNN	

4	1-15 Juni 2021	Pengujian data	 Training/Validasi Akurasi per Epoch The graph displays two data series: 'train' (blue line) and 'validation' (orange line). The x-axis represents the number of epochs from 0 to 75, and the y-axis represents accuracy from 0.4 to 0.9. Both curves show an upward trend, with the training accuracy reaching approximately 0.9 and the validation accuracy stabilizing around 0.85 after epoch 35.
5	21-31 Juli 2021	Drafting Publikasi	
6	Agustus 2021	Penyusunan laporan akhir	

