



INSYST

Journal of Intelligent System and Computation

p-ISSN: 2621-9220

e-ISSN: 2722-1962

Volume 5 Nomor 2, Oktober 2023



Published By **Lembaga Penelitian dan Pengabdian Masyarakat (LPPM)**
Institut Sains dan Teknologi Terpadu Surabaya (ISTTS)
formerly **Sekolah Tinggi Teknik Surabaya (STTS)**



Managed By
Departement of Informatics
Institut Sains dan Teknologi Terpadu Surabaya (ISTTS)

INSYST

Journal of Intelligent System and Computation

Volume 05 Nomor 02 Oktober 2023

Editor in Chief:

Dr. Yosi Kristian, S.Kom, M.Kom.

Institut Sains dan Teknologi Terpadu Surabaya, Indonesia

Managing Editor:

Dr. Esther Irawati Setiawan, S.Kom., M.Kom.

Institut Sains dan Teknologi Terpadu Surabaya, Indonesia

Hendrawan Armanto, S.Kom., M.Kom.

Institut Sains dan Teknologi Terpadu Surabaya, Indonesia

Editorial Board:

Dr. Ir. Endang Setyati, M.T.

Institut Sains dan Teknologi Terpadu Surabaya, Indonesia

Ir. Edwin Pramana, M.App.Sc, Ph.D

Institut Sains dan Teknologi Terpadu Surabaya, Indonesia

Prof. Dr. Ir. Mauridhi Hery Purnomo, M.T.

Institut Sepuluh November, Indonesia

Hindriyanto Dwi Purnomo, Ph.D.

Universitas Kristen Satya Wacana, Salatiga, Indonesia

Reddy Alexandro H., S.Kom., M.Kom.

Institut Sains dan Teknologi Terpadu Surabaya, Indonesia

Dr. Diana Purwitasari, S.Kom., M.Sc.

Institut Sepuluh November, Indonesia

Dr. Joan Santoso, S.Kom., M.Kom.

Institut Sains dan Teknologi Terpadu Surabaya, Indonesia

INSYST

Journal of Intelligent System and Computation

Volume 05 Nomor 02 Oktober 2023

Reviewer:

Dr. I Ketut Eddy Purnama, ST., MT.
Institut Sepuluh November, Indonesia

Dr. Anang Kukuh Adisusilo, ST, MT.
Universitas Wijaya Kusuma, Surabaya, Indonesia

Teguh Wahyono, S.Kom., M.Cs.
Universitas Kristen Satya Wacana, Salatiga, Indonesia

Prof. Dr. Benny Tjahjono, M.Sc.
Coventry University, United Kingdom

Dr. Ir. Gunawan, M.Kom.
Institut Sains dan Teknologi Terpadu Surabaya, Indonesia

Dr. Umi Laili Yuhana S.Kom., M.Sc.
Institut Sepuluh November, Indonesia

Dr. Tita Karlita, S.Kom., M.Kom.
Politeknik Elektronika Negeri Surabaya, Indonesia

Dr. Ir. Rika Rokhana, M.T.
Politeknik Elektronika Negeri Surabaya, Indonesia

Dr. I Made Gede Sunarya, S.Kom., M.Cs.
Universitas Pendidikan Ganesha, Indonesia

Dr. Yuni Yamasari, S.Kom., M.Kom.
Universitas Negeri Surabaya, Indonesia

Dr. Adri Gabriel Sooai, S.T., M.T.
Universitas Katolik Widya Mandira, Indonesia

Dr. Lukman Zaman PCSW, M.Kom.
Institut Sains dan Teknologi Terpadu Surabaya, Indonesia

INSYST

Journal of Intelligent System and Computation

Volume 05 Nomor 02 Oktober 2023

Reviewer:

Windra Swastika, Ph.D

Universitas Ma Chung, Indonesia

Romy Budhi Widodo, Dr.Eng.

Universitas Ma Chung, Indonesia

Theresia Ratih Dewi Saputri, Ph.D.

Ciputra University, Indonesia

Richard Evan Sutanto, Ph.D.

Ciputra University, Indonesia

Prof. Dr. Yuhandri, S.Kom., M.Kom.

Universitas Putra Indonesia YPTK Padang, Indonesia

INSYST

Journal of Intelligent System and Computation

Volume 05 Nomor 02 Oktober 2023

Daftar Isi

Model Deteksi Kebakaran Hutan dan Lahan Menggunakan Transfer Learning DenseNet201 Rifqi Akmal Saputra, Faisal Dharma Adhinata	65
Klasifikasi Kualitas dan Prediksi Kondisi Air Tanah di DKI Jakarta Menggunakan Algoritma Naïve Bayes Arsya Fathiarahma; Nina Sulistiyowati, Taufik Ridwan, Apriade Voutama	73
Pengendalian Suhu dan Kelembaban Budidaya Jamur Kuping dengan Kendali PID Penalaan Ziegler-Nichols Setya Ardhi, Tjwanda Putra Gunawan, Suhatati Tjandra	83
KLASIFIKASI MICRO-EXPRESSION MENGGUNAKAN K-NEAREST NEIGHBORS MENGGUNAKAN FITUR CAS DAN HOG Nikko Riestian Putra Wardoyo, Joan Santoso, Esther Irawati Setiawan	96
Implementasi Metode Fuzzy Analytica Hierarchy Process (FAHP) Dalam Penentuan Bobot Seleksi Mahasiswa Program Pendidikan Dokter Spesialis Ni Nyoman Murni, Gede Rasben Dantes, I Made Candiasa	104
Algoritma Random Forest pada Prediksi Status Kredit Usaha Rakyat untuk Mengurangi Nonperforming Loan Rate Wulansari Wulansari, Diana Purwitasari	109
Sistem Deteksi dan Klasifikasi Truk Air Menggunakan YOLO v5 dan EfficientNet-B4 Ardian Kurniawan, Daniel Martomanggolo Wonohadidjojo	115

Model Deteksi Kebakaran Hutan dan Lahan Menggunakan Transfer Learning DenseNet201

Rifqi A. Saputra¹ dan Faisal D. Adhinata^{2,*}

¹Program Studi Rekayasa Perangkat Lunak, Fakultas Informatika, Institut Teknologi Telkom Purwokerto, Purwokerto, Indonesia

Corresponding author: Faisal Dharma Adhinata (e-mail: faisal@ittelkom-pwt.ac.id).

ABSTRACT Forest and land fires in Indonesia are frequent occurrences and cause significant health, ecological, and social losses. Human and natural factors play a role in triggering this fire. However, the management of forest and land fires still faces obstacles in accurately predicting the location of hot spots, so optimal control is difficult. This research aims to create a model capable of detecting forest and land fires using a transfer learning approach, incorporating the DenseNet201 architecture to improve accuracy. The dataset used in this study comes from the Fire Forest Dataset on the Kaggle website. The feature extraction process is carried out using the DenseNet201 architecture, and the resulting model is tested using the confusion matrix method to classify images into two classes, namely fire and non-fire classes. The test results using a data test of 380 data show an accuracy rate of 99% in identifying images of forest and land fires. This research makes an important contribution to the development of forest and land fire detection technology. The use of a transfer learning approach with the DenseNet201 architecture has the potential to improve fire detection accuracy. It is hoped that this research will provide a basis for developing intelligent systems that are more sophisticated and effective in overcoming the problem of forest and land fires, as well as protecting the environment and public health in Indonesia.

KEYWORDS Confusion Matrix, DenseNet201, Forest and Land Fire, Transfer Learning

ABSTRAK Kebakaran hutan dan lahan di Indonesia merupakan peristiwa yang sering terjadi dan menimbulkan kerugian yang signifikan dalam bidang kesehatan, ekologi, dan sosial. Faktor manusia dan alam berperan dalam memicu terjadinya kebakaran ini. Namun, penanganan kebakaran hutan dan lahan masih menghadapi kendala dalam memprediksi lokasi titik panas secara akurat, sehingga pengendalian yang optimal sulit dilakukan. Oleh karena itu, diperlukan pengembangan sistem cerdas untuk mendeteksi kebakaran hutan dan lahan dengan lebih efektif. Penelitian ini bertujuan untuk menciptakan sebuah model yang mampu mendeteksi kebakaran hutan dan lahan dengan menggunakan pendekatan *transfer learning*, dengan memanfaatkan arsitektur *DenseNet201* guna meningkatkan akurasi deteksi. *Dataset* yang digunakan dalam penelitian ini berasal dari *Fire Forest Dataset* pada situs *Kaggle*. Proses ekstraksi fitur dilakukan menggunakan arsitektur DenseNet201, dan model yang dihasilkan diuji dengan menggunakan metode *confusion matrix* untuk mengklasifikasikan gambar menjadi dua kelas, yaitu kelas api dan non-api. Melalui pelatihan menggunakan arsitektur *DenseNet201*, diperoleh model yang efektif dalam mendeteksi kebakaran hutan dan lahan. Hasil pengujian dengan menggunakan data uji sebanyak 380 data menunjukkan tingkat akurasi sebesar 99% dalam mengenali gambar kebakaran hutan dan lahan. Penelitian ini memberikan kontribusi penting dalam pengembangan teknologi deteksi kebakaran hutan dan lahan. Penggunaan pendekatan *transfer learning* dengan arsitektur *DenseNet201* memiliki potensi untuk meningkatkan akurasi deteksi kebakaran yang lebih baik. Diharapkan penelitian ini dapat memberikan landasan bagi pengembangan sistem cerdas yang lebih canggih dan efektif dalam mengatasi masalah kebakaran hutan dan lahan, serta melindungi lingkungan dan kesehatan masyarakat di Indonesia.

KATA KUNCI *Confusion Matrix, DenseNet201, Kebakaran Hutan dan Lahan, Transfer Learning*

I. PENDAHULUAN

Hutan memainkan peran penting dalam menjaga keseimbangan ekologi bumi. Menurut *Food and Agriculture Organization* (FAO), luas hutan di bumi adalah 4,06 miliar hektar (Ha) [1]. Menurut laporan Kementerian Lingkungan Hidup dan Kehutanan, pada tahun 2020 luas hutan di Indonesia mencapai 95,6 juta hektar (Ha), setara dengan 50,9% dari total luas daratan di Indonesia [2]. Luasnya lahan hutan akan menimbulkan masalah yang dapat terjadi yaitu bencana kebakaran hutan dan lahan. Dampak kebakaran ini dapat mempengaruhi beberapa sektor, seperti kesehatan, ekologi, ekonomi, dan masyarakat [3].

Laporan Kementerian Lingkungan Hidup dan Kehutanan tahun 2021 menugambarkannya luas kebakaran hutan dan lahan di Indonesia mencapai 354.528 hektare (Ha). Dibandingkan tahun 2020, angka tersebut meningkat 19,4% dengan luas 296.942 hektare (Ha) [4]. Penyebab utama kebakaran hutan dan lahan di Indonesia adalah perilaku manusia, dengan persentase 99% dan 1% disebabkan oleh faktor alam [5]. Upaya penanggulangan kebakaran hutan dan lahan di Indonesia belum maksimal karena sulit dideteksi [6], dan kurangnya kesadaran masyarakat dan konsolidasi penanganan kebakaran hutan dan lahan yang belum dilakukan secara menyeluruh telah menyebabkan kebakaran hutan dan lahan. sering terjadi di Indonesia.

Berdasarkan permasalahan yang telah dijelaskan sebelumnya, maka diperlukan suatu model yang dapat mendeteksi kebakaran hutan dan lahan dengan perkembangan teknologi kecerdasan buatan yang semakin pesat yang dapat menyelesaikan masalah atau bekerja seperti manusia [7]. Kecerdasan buatan dapat diterapkan dalam penanggulangan bencana [8], seperti mendeteksi kebakaran hutan dan lahan menggunakan *image object recognition* [9]. Penelitian sebelumnya telah dilakukan dengan menggunakan salah satu metode *transfer learning* yaitu *DenseNet* [10][11] sudah memperoleh hasil yang cukup baik dengan akurasi masing-masing 98,16% dan 92% namun masih terdapat kendala pada ketidakseimbangan data kebakaran hutan yang digunakan sehingga dapat menyebabkan masalah seperti bias kelas dan akurasi yang rendah. Model cenderung menghasilkan hasil yang lebih baik pada kelas mayoritas dan memiliki kinerja yang buruk dalam mendeteksi kelas minoritas.

Transfer learning adalah teknik yang menerapkan model yang telah dilatih sebelumnya dengan menggunakan *dataset* yang nantinya dapat digunakan untuk menyelesaikan masalah model yang memiliki kesamaan [12]. Selain itu, teknik pembelajaran transfer dapat dimodifikasi, dan parameternya diubah untuk beradaptasi dengan *dataset* baru. Dalam penelitian ini, kami mengusulkan salah satu jenis pembelajaran transfer, yaitu arsitektur *DenseNet201* [13], jaringan saraf *convolutional* yang terdiri dari 201 lapisan dalam yang dapat memuat model pra pelatihan menggunakan gambar dari *dataset* CIFAR 100 [14]. Dengan kedalaman lapisan dan model pra pelatihan penelitian ini juga menggunakan *dataset* yang telah dipersiapkan agar data

yang digunakan seimbang antara data kebakaran dan non-kebakaran guna memaksimalkan proses pelatihan menggunakan metode *DenseNet201* yang menghasilkan model yang dapat mendeteksi kebakaran hutan dan lahan dengan akurasi tinggi.

II. TINJAUAN PUSTAKA

Penelitian yang dilakukan oleh [10] pada tahun 2020 tentang deteksi kebakaran hutan dengan menggunakan metode *DenseNet* dan proses augmentasi data berbasis *cycleGAN*. Penelitian ini memiliki tujuan untuk mengatasi masalah ketidakseimbangan data dalam klasifikasi gambar kebakaran hutan yang menyebabkan *overfitting* yang berdampak pada penurunan kinerja model. Untuk mengatasi permasalahan tersebut digunakan proses augmentasi data berbasis *cycleGAN* dengan menghasilkan gambar-gambar kebakaran sintesis yang menyerupai gambar kebakaran sebenarnya. Model yang akan dibangun pada penelitian ini menggunakan *densely connected convolutional networks* (*DenseNet*) yang memiliki kelebihan mengurangi *vanishing gradient problem*, memperkuat *feature propagation*, mengutamakan penggunaan ulang fitur, dan mengurangi jumlah parameter sehingga dapat mengurangi waktu pelatihan [15]. Data yang digunakan pada penelitian ini terdiri dari 4959 gambar non-kebakaran yang diperoleh dari 8 *database* dengan kategori pemandangan dan *database* tempat wisata Korea. 1395 gambar kebakaran hutan yang diperoleh dengan *crawling* data pada *website* dan tangkapan gambar video *drone*. Penelitian ini juga melakukan perbandingan hasil dengan metode *transfer learning* lainnya yaitu *VGG-16* dan *ResNet-50*. Hasil yang didapatkan dari hasil pengujian dengan menggunakan data hasil augmentasi menggunakan *cycleGAN* diperoleh metode *DenseNet* memiliki hasil yang paling baik dengan akurasi 98,27% dan *F1-Score* 98,16. Hasil tersebut menunjukkan bahwa metode *DenseNet* mampu memberikan akurasi deteksi kebakaran hutan yang tinggi.

Penelitian dilakukan oleh [11] pada tahun 2022 tentang deteksi kebakaran hutan dengan menggunakan metode *DenseNet* dengan tujuan untuk menghindari peringatan palsu yang berasal dari objek yang menyerupai api dan objek api yang berukuran kecil. Penelitian ini memiliki beberapa tahapan proses yang dilakukan seperti pengumpulan *dataset* kemudian dilakukan tahapan *preprocessing dataset* dengan melakukan *cropped* dan *resize* menjadi 224x224 piksel untuk mendapatkan fitur yang diinginkan dilanjutkan dengan proses augmentasi data untuk menghasilkan variasi gambar serta memperluas kumpulan data agar tidak terjadi *overfitting* pada tahap pelatihan. *Dataset* sudah diproses sebelumnya, dan fitur-fiturnya diekstrak. Kemudian model tersebut dilatih dengan *dataset* yang diklasifikasikan berdasarkan apakah objek tersebut termasuk api dan non api. Data yang digunakan berisi gambar api, objek seperti api, dan gambar non-api yang diperoleh dalam berbagai kondisi iklim, berbagai jenis vegetasi, dan juga gambar yang memiliki jarak dan kecerahan yang berbeda-beda api. Selain itu, beberapa gambar memiliki objek dwiwarna, cahaya,

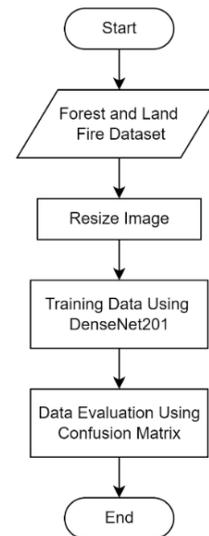
matahari terbit, dan matahari terbenam yang diperhitungkan untuk mendiversifikasi dan meningkatkan akurasi model dengan total data 1760 yang terdiri dari kelas api dan non-api. Hasil pelatihan yang diperoleh dari model *DenseNet* 92% sedangkan akurasi validasi adalah 74%. Model ini memberikan akurasi yang lebih tinggi jika dibandingkan dengan algoritma *deep learning* lainnya seperti YOLO V3 dengan akurasi 81.9%, YOLO V5 dengan akurasi 88.2%, dan K-Means dengan akurasi 90.5. Hasil tersebut menunjukkan model *DenseNet* dapat digunakan pada kamera pengintai dan *drone* yang terletak di alam liar dan dapat digunakan untuk memprediksi kebakaran. kebakaran dan wilayahnya.

Berdasarkan penelitian yang telah dijelaskan sebelumnya hasil yang diperoleh menggunakan metode *DenseNet* sudah baik, namun masih terdapat kendala pada ketidakseimbangan data kebakaran hutan yang digunakan sehingga dapat menyebabkan masalah seperti bias kelas dan akurasi yang rendah. Model cenderung menghasilkan hasil yang lebih baik pada kelas mayoritas dan memiliki kinerja yang buruk dalam mendeteksi kelas minoritas. Dalam hal ini, ini berarti model dapat lebih baik dalam mendeteksi non-kebakaran, tetapi kurang baik dalam mengenali gambar kebakaran. Oleh karena itu pada penelitian ini menggunakan *dataset* yang telah dipersiapkan agar data yang digunakan seimbang antara data kebakaran dan non-kebakaran guna memaksimalkan proses pelatihan menggunakan metode *DenseNet201*.

III. METODOLOGI

Penelitian mengenai deteksi kebakaran hutan dan lahan dengan menggunakan algoritma *DenseNet201* diawali dengan pencarian *dataset* hutan dan lahan melalui situs *Kaggle* yang akan digunakan dalam proses pelatihan di mana data yang diperoleh terdiri dari dua kelas yaitu *fire* dan *no-fire*. Selanjutnya data yang telah didapatkan akan melalui tahap *preprocessing* dengan melakukan *resize* pada gambar menjadi 224 x 224 piksel. Data yang telah dilakukan proses *preprocessing* akan masuk ke dalam tahap pelatihan dengan skema percobaan komposisi data pelatihan dan validasi yaitu 60:40, 70:30, 80:20, dan 90:10 menggunakan algoritma *DenseNet201* alasan metode ini digunakan dikarenakan memiliki kelebihan antara lain memiliki 201 lapisan di mana dapat membantu model untuk mengekstraksi fitur-fitur penting yang diperlukan untuk mendeteksi kebakaran hutan dan lahan dengan lebih baik kemudian memiliki konektivitas yang padat yang berguna untuk membantu model dalam mengenali dan menghubungkan pola-pola penting yang berkaitan dengan kebakaran hutan dan lahan dan dapat meminimalkan masalah gradien yang hilang di mana dalam konteks deteksi kebakaran hutan dan lahan dapat membantu model untuk belajar dengan lebih baik dari data pelatihan yang kompleks dan mengatasi masalah gradien yang hilang. Hasil yang diperoleh dari proses pelatihan memperoleh empat model terbaik dari masing-masing skema komposisi data dengan format ekstensi h5 kemudian empat model terbaik

tersebut akan dilakukan proses evaluasi model dengan menggunakan *confusion matrix* untuk mencari model paling baik dari keempat skema komposisi data yang di coba. Gambar 1 menunjukkan *flowchart* deteksi kebakaran hutan dan lahan.



GAMBAR 1. Flowchart Penelitian

A. DATASET

Dataset yang digunakan pada penelitian ini bersumber dari situs *Kaggle* [16]. Kumpulan data tersebut digunakan untuk mengatasi permasalahan penanganan kebakaran hutan dan lahan. Data tersebut memiliki format jpg dan memiliki ukuran data 250x250 piksel. Data tersebut terdiri dari dua kelas yaitu *fire* dan *non-fire* pada objek hutan dan lahan [17]. Data tersebut memiliki total 1900 data gambar, terdiri dari 1580 data gambar pada direktori pelatihan dan 380 data gambar yang digunakan untuk pengujian dengan jumlah sampel data yang sama pada masing-masing kelas dengan rincian pada direktori pelatihan kelas *fire* dan *non-fire* memiliki jumlah sampel data masing-masing sebanyak 790 data gambar. Direktori pengujian kelas *fire* dan *non-fire* memiliki jumlah sampel data masing-masing sebanyak 190 data gambar tahap ini diperlukan agar terjadi keseimbangan jumlah data agar memaksimalkan proses pelatihan menggunakan metode *DenseNet201*. Gambar 2 adalah ilustrasi dari *Forest Fire Dataset*.



GAMBAR 2. Ilustrasi Gambar Pada Dataset

B. PREPROCESSING DATASET

Tahap *preprocessing* dilakukan untuk menyiapkan data sebelum proses pelatihan menggunakan algoritma DenseNet201. Pada tahap ini, data gambar hutan dan lahan yang sebelumnya memiliki ukuran 250x250 piksel kemudian dilakukan proses *resize* menjadi 224x224 piksel menggunakan fungsi yang disediakan oleh *tensorflow* yaitu `tensorflow.keras.preprocessing.image_dataset_from_directory()` untuk ditambahkan parameter `image_size` [18]. Ukuran 224x224 piksel dipilih karena proses ekstraksi fitur menggunakan algoritma *DenseNet201* memerlukan konversi ke ukuran tersebut dikarenakan memiliki kompatibilitas dengan model *pre-trained* yang digunakan selain itu dapat meningkatkan efisiensi komputasi dan informasi visual pada ukuran 224x224 pixel sudah cukup untuk mempertahankan fitur-fitur penting yang ada pada gambar yang akan digunakan dalam proses pelatihan [19]. Gambar 3 menunjukkan contoh hasil *preprocessing* gambar ke ukuran yang seragam yaitu 224x224 piksel.



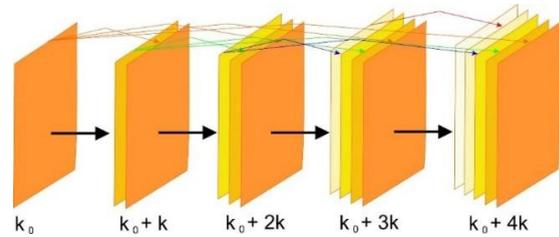
GAMBAR 3. Contoh Gambar Setelah *Preprocessing*

C. PELATIHAN MODEL MENGGUNAKAN DENSENET201

Densely Connected Convolutional Networks (DenseNet) merupakan model arsitektur *Deep Learning* yang menerapkan konsep *transfer learning* [15]. Cara kerja *DenseNet* adalah dengan menghubungkan setiap layer beserta *feature map* ke semua layer berikutnya sehingga layer selanjutnya menerima *input* peta fitur dari semua layer sebelumnya [20][21]. Keuntungan dari *DenseNet* ringan pada gradien, dan penerapan fitur, mendorong penggunaan kembali fitur, dan fungsionalitas mengurangi jumlah parameter [22]. Arsitektur *DenseNet201*, yang merupakan *convolutional neural network* yang terdiri dari 201 *deep layer* yang dapat memuat model pra pelatihan menggunakan gambar dari *dataset CIFAR 100* terdiri dari 60000 gambar berwarna yang memiliki ukuran 32x32 piksel dan terdiri dari 100 kelas yang menyebabkan jaringan untuk mempelajari representasi fitur yang luas untuk berbagai jenis gambar.

Lapisan yang digunakan adalah konvolusi dengan filter 3x3, aktivasi *ReLU*, dan normalisasi *batch*. Setiap lapisan blok padat memiliki matriks masukan untuk piksel gambar. Proses berikut adalah normalisasi *batch*, yang mencegah *overfitting* pada tahap pelatihan. Kemudian aktivasi *ReLU* berfungsi untuk mengubah nilai *y* menjadi 0 jika nilai *y* negatif. Jika nilai *y* positif, maka akan memiliki nilai yang sama. Filter 3x3 pada matriks konvolusi gambar yang telah lolos operasi aktivasi *ReLU* akan dikalikan dengan matriks konvolusi dengan filter

3x3. *Output* yang dihasilkan adalah nilai matriks yang telah diproses sebelumnya [23]. Gambar 4 adalah *arsitektur DenseNet201*.



GAMBAR 4. Arsitektur *DenseNet201*

D. EVALUASI DATA MENGGUNAKAN CONFUSION MATRIX

Pada tahap evaluasi model, metode *confusion matrix* dapat digunakan untuk mengukur performansi atau kinerja suatu model pada kasus klasifikasi [24]. Model dapat diukur performansi atau kinerjanya dengan menggunakan beberapa ketentuan dalam *confusion matrix*. Empat nilai merupakan hasil dari tahapan klasifikasi dalam matriks kebingungan sebagai berikut *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)*, dan *False Negative (FN)* [25]. Uji coba pada model *DenseNet201* dilakukan dengan menggunakan sebanyak 380 data uji yang nantinya akan mengukur akurasi model. Gambar 5 adalah tabel *confusion matrix* dua kelas.

		Actual Values	
		1 (Positive)	0 (Negative)
Predicted Values	1 (Positive)	TP (True Positive)	FP (False Positive)
	0 (Negative)	FN (False Negative)	TN (True Negative)

GAMBAR 5. Contoh Tabel *Confusion Matrix* Dua Kelas

IV. HASIL DAN PEMBAHASAN

Bagian Model deteksi kebakaran hutan dan lahan dijalankan pada komputer dengan CPU Intel Core i3-5005U, RAM DDR3 8 GB, dan sistem operasi Windows 10 Pro 64-bit. Penelitian ini dilakukan untuk menemukan model menggunakan *DenseNet201*. Kemudian, model tersebut diuji dengan data testing untuk mendapatkan model dengan akurasi terbaik.

A. HASIL PENGUMPULAN DATA

Dalam penelitian ini menggunakan *dataset* kebakaran hutan dan lahan dengan total data yang digunakan adalah 1900 gambar yang terdiri dari 1520 gambar untuk pelatihan dan validasi, dan untuk pengujian terdiri dari 380 gambar. Total

data gambar untuk data pelatihan adalah 1520, terdiri dari 760 gambar hutan dan lahan dengan kebakaran dan 760 gambar hutan dan lahan tanpa kebakaran. Kemudian untuk pengujian data, menggunakan 380 gambar yang terdiri dari 190 gambar hutan dan lahan dengan kebakaran dan 190 gambar hutan dan lahan tanpa kebakaran. Peneliti mengubah nama api dan bukan api sebelumnya menjadi label di mana ada api dan tidak ada api. Penelitian ini juga menggunakan skema komposisi data yaitu 60:40, 70:30, 80:20, dan 90:10 untuk data pelatihan dan validasi. *Dataset* terdiri dari dua kelas yaitu *fire* dan *no fire*. Tabel 1 menunjukkan skema data yang digunakan dalam penelitian ini. Tabel 1 berisi data pelatihan, validasi, dan pengujian data gambar sementara Tabel 2 adalah pembagian data dari masing-masing skema yang digunakan.

TABEL I
JUMLAH DATA SETIAP KATEGORI

Data	Fire	No-Fire	Total
Training & Validation	760	760	1520
Testing	90	90	380

TABEL II
PEMBAGIAN SKEMA DATA

Scheme Data Training & Validation (%)	Data Training & Validation	Data Testing
60:40	912:608	308
70:30	1064:456	308
80:20	1216:304	308
90:10	1368:152	308

B. HASIL PELATIHAN

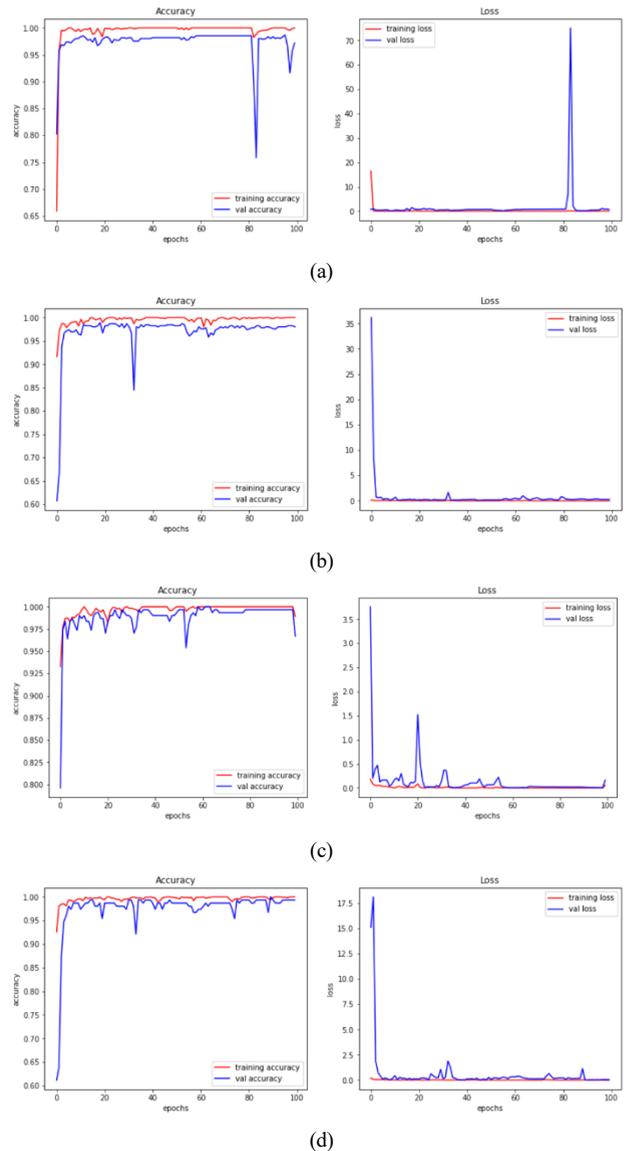
Pada tahap pelatihan ini menggunakan *DenseNet201* dengan pendekatan untuk membekukan setengah dari jaringan pra pelatihan karena jumlah *dataset* yang digunakan kecil, dan perbedaan domain dari *dataset* asli dan baru berbeda. Pada tahap pelatihan ini menggunakan iterasi sebanyak 100 *epoch*, jumlah distribusi sampel data menggunakan 32 *batch size*. Proses pelatihan ini menggunakan *ModelCheckpoint* untuk menyimpan model atau bobot pada interval tertentu sehingga model atau bobot dapat dimuat nanti untuk melanjutkan pelatihan dari keadaan tersimpan. Proses pelatihan menggunakan *google colab* dengan runtime GPU Google sehingga proses pelatihan model *DenseNet201* dapat dilakukan lebih cepat. Tabel 3 menunjukkan hasil pelatihan menggunakan *DenseNet201*.

TABEL III
HASIL PELATIHAN

Scheme Data Training: Validation (%)	Validation Accuracy	Validation Loss	Time (minutes)
60:40	97%	0.6995	27
70:30	98%	0.2702	30
80:20	97%	0.1623	29
90:10	99%	0.0588	30

Berdasarkan tabel di atas, hasil pada skenario data 90:10 memiliki hasil terbaik, dengan tingkat akurasi validasi sebesar

99% dan nilai validasi *loss* sebesar 0,0588 dengan waktu pelatihan 30 menit. Skenario data 70:30 memiliki akurasi validasi sebesar 98%, kemudian skenario 60:40 dan 80:20 mendapatkan akurasi yang sama yaitu 97%. Berikut adalah Gambar 6 grafik hasil pelatihan dari keempat jenis skenario data yang digunakan.



GAMBAR 6. Akurasi dan loss grafik (a) skema 60:40, (b) skema 70:30, (c) skema 80:20, dan (d) skema 90:10

C. HASIL CONFUSION MATRIX

Pada pengujian model ini menggunakan metode *confusion matrix* untuk mengukur tingkat akurasi model untuk data pengujian yang diberikan sebanyak 380 data gambar yang terdiri dari 190 gambar berlabel kebakaran dan 190 gambar berlabel tidak terbakar. Pengujian ini dilakukan terhadap masing-masing model terbaik yang telah disimpan dari skenario data 60:40, 70:30, 80:20, dan 90:10 terhadap model yang telah dilatih sebelumnya. Tabel IV menunjukkan hasil

pengujian model terbaik dari masing-masing skema data dengan menggunakan metode *confusion matrix*.

TABEL IV
HASIL PENGUJIAN MENGGUNAKAN CONFUSION MATRIX

		Prediction	
		No-Fire	Fire
Actual	No-Fire	190	0
	Fire	10	180

(a)

		Prediction	
		No-Fire	Fire
Actual	No-Fire	185	5
	Fire	2	188

(b)

		Prediction	
		No-Fire	Fire
Actual	No-Fire	185	5
	Fire	7	183

(c)

		Prediction	
		No-Fire	Fire
Actual	No-Fire	188	2
	Fire	0	190

(d)

Tabel 4 merupakan hasil pengujian menggunakan confusion matrix, a) skema 60:40, b) skema 70:30, c) skema 80:10, dan d) skema 90:10. Berdasarkan Tabel 4 hasil *confusion matrix* dari masing-masing skema data 60:40, 70:30, 80:20, dan 90:10, diperoleh hasil perhitungan nilai *precision*, *recall*, *f1-score*, dan *accuracy* dari setiap skema data dengan menggunakan fungsi *classification_report()* pada *library scikit learn* dapat dilihat pada tabel di bawah ini. Berdasarkan hasil tersebut, model skema data 90:10 adalah yang terbaik, dengan tingkat akurasi 99%. Kemudian model skema data 70:30 menghasilkan tingkat akurasi sebesar 98%. Model skema data 60:40 dan 80:20 memiliki tingkat akurasi yang sama yaitu 97%. Tabel 5 menampilkan hasil *precision*, *recall*, *f1-score*, dan *accuracy* untuk masing-masing model berdasarkan skema data yang diuji.

TABEL V
HASIL PENGUJIAN MODEL

Scheme Data (%)	Class	Precision	Recall	F1-Score	Accuracy
60:40	0 (No-Fire)	95%	100%	97%	97%
	1 (Fire)	100%	95%	97%	
0:30	0 (No-Fire)	99%	97%	98%	98%
	1 (Fire)	97%	99%	98%	
80:20	0 (No-Fire)	96%	97%	97%	97%
	1 (Fire)	97%	96%	97%	
90:10	0 (No-Fire)	100%	99%	99%	99%
	1 (Fire)	99%	100%	99%	

D. PEMBAHASAN

Berdasarkan hasil pelatihan model untuk menemukan model yang memiliki akurasi terbaik dalam mendeteksi kebakaran hutan dan lahan, maka peneliti melakukan perbandingan hasil pengujian dalam penelitian tersebut dengan metode dan hasil penelitian lainnya seperti pada Tabel 6. Penelitian yang dilakukan oleh [10] mencoba membandingkan metode *DenseNet* dengan *VGG-16* dan *ResNet-50* dari hasil pengujian dilakukan metode *VGG-16* memiliki nilai akurasi 93.75% dan *ResNet-50* memiliki nilai akurasi 96.92%. Penelitian selanjutnya yang dilakukan oleh [26] dengan menggunakan *TeutongNet* dan *ResNet-50V2* pada percobaan menggunakan *TeutongNet* memperoleh nilai akurasi 98.68% dan *ResNet-50V2* memperoleh nilai akurasi 98.68%. Penelitian ini model *DenseNet201* pada skema data 90:10 mendapatkan akurasi tertinggi yaitu 99% didapatkan dari pengujian menggunakan 380 data pengujian. Berdasarkan hasil tersebut metode *DenseNet201* yang menggunakan *dataset* yang telah dipersiapkan agar data yang digunakan seimbang antara data kebakaran dan non-kebakaran guna memaksimalkan proses pelatihan menggunakan metode *DenseNet201* sehingga memperoleh tingkat akurasi yang lebih baik dibandingkan metode *transfer learning* lainnya dan penelitian sebelumnya yang sama terkait deteksi kebakaran hutan dan lahan dengan menggunakan *DenseNet* [10][11]. Model *DenseNet201* dari penelitian ini dapat diterapkan pada kamera pengawas yang ditempatkan di area yang rawan terjadi kebakaran hutan dan lahan sehingga dapat digunakan untuk melakukan deteksi dini kebakaran hutan dan lahan untuk mendukung upaya pemantauan dan deteksi kebakaran hutan di Indonesia.

TABEL VI
HASIL PERBANDINGAN DENGAN PENELITIAN TERKAIT

Classification Model	Accuracy
VGG-16	93.75%
ResNet-50	96.92%
ResNet-50V2	94.74%
TeutongNet	98.68%
Proposed Method (DenseNet201)	99%

Penelitian ini juga melakukan eksperimen menggunakan data gambar acak gambar hutan di Indonesia yang mengalami kebakaran maupun tidak yang diperoleh dari hasil *crawling* data melalui internet. Hasil eksperimen dan klasifikasi menggunakan model *DenseNet201* dari skema data 90:10 yang memiliki akurasi terbaik sebesar 99% dapat dilihat pada Gambar 7. Hasil percobaan menunjukkan 15 gambar uji yang dipilih secara acak dan memiliki dua keluaran yaitu kelas gambar aktual dan klasifikasi kelas. Hasil eksperimen pengujian menggunakan model *DenseNet201* dari skema eksperimen 90:10 menghasilkan 15 dari 15 gambar yang terklasifikasi benar sesuai kelasnya walaupun data latih yang digunakan menggunakan *dataset* hutan dan lahan luar negeri namun Model mampu mempelajari pola dan fitur yang ada dalam gambar untuk mengenali objek atau kejadian tertentu.



GAMBAR 7. Contoh Hasil Klasifikasi Dengan Menggunakan Model Dari Skema Data 90:10S

Oleh karena itu, meskipun pohon-pohon di Indonesia memiliki perbedaan dengan pohon-pohon dalam *dataset* pelatihan, model dapat belajar untuk mengenali pola dan fitur yang khas dari kebakaran, seperti warna-warna cerah dan pola api, yang biasanya terlihat serupa di berbagai lokasi.

V. KESIMPULAN

Eksperimen menggunakan data gambar acak dari direktori pengujian. Hasil eksperimen dan klasifikasi menggunakan model *DenseNet201* dari skema data 90:10 yang memiliki akurasi terbaik sebesar 99% dapat dilihat pada Gambar 7 hasil eksperimen menunjukkan 15 gambar uji yang dipilih secara acak dan memiliki dua *output* yaitu kelas gambar aktual dan kelas klasifikasi. Hasil eksperimen pengujian menggunakan model *DenseNet201* dari skema eksperimen 90:10 menghasilkan 15 dari 15 gambar yang terklasifikasi benar sesuai kelasnya. Diharapkan model yang diperoleh dari penelitian ini dapat membantu mendeteksi kebakaran hutan dan lahan.

Pada penelitian ini hanya terbatas pada proses pembuatan model yang dapat melakukan deteksi kebakaran hutan dan lahan dengan menggunakan metode *DenseNet201*. Diharapkan pada penelitian yang akan datang dapat dilakukan pengembangan atau pengujian model dengan

dataset yang lebih besar, eksplorasi dengan metode *transfer learning* yang lainnya, atau melakukan integrasi model ke dalam sistem aplikasi agar dapat digunakan untuk meningkatkan deteksi kebakaran.

PERAN PENULIS

Rifqi Akmal Saputra: Analisis Formal, Investigasi, Perangkat Lunak, Penulisan Penyusunan Draf Asli;

Faisal Dharma Adhinata: Konseptualisasi, Sumber Daya, Validasi, Visualisasi, Review Penulisan & Penyuntingan;

COPYRIGHT



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

DAFTAR PUSTAKA

- [1] “Deforestasi Global Turun, Luas Hutan Alam Naik,” Aug. 2020. <https://www.forestdigest.com/detail/689/deforestasi-global-turun-luas-hutan-alam-naik> (accessed Apr. 12, 2022).
- [2] N. Anugrah, “PPID | Kementerian Lingkungan Hidup dan Kehutanan | Capaian TORA dan Perhutanan Sosial di Tahun 2021,” Dec. 2021. <http://ppid.menlhk.go.id/berita/siaran-pers/6330/capaian-tora-dan-perhutanan-sosial-di-tahun-2021> (accessed Apr. 12, 2022).

- [3] M. Ahmad, P. Reza, and A. Z. S. Rahma, "Peran Polisi Kehutanan Untuk Mewujudkan Sustainable Development Goals Forest Fire Control Through Strengthening the Role of the Forest Police To Realize Sustainable Development Goals," *Hukum Lingkungan, Tata Ruang dan Agraria*, vol. 1, no. 1, pp. 23–44, 2021.
- [4] V. A. Diihni, "Luas Kebakaran Hutan dan Lahan RI Bertambah 19% pada 2021 | Databoks," Jan. 2022. <https://databoks.katadata.co.id/datapublish/2022/01/11/luas-kebakaran-hutan-dan-lahan-ri-bertambah-19-pada-2021> (accessed Apr. 13, 2022).
- [5] B. Editorial, "99% Penyebab Kebakaran Hutan dan Lahan Adalah Ulah Manusia - BNPB," *BADAN NASIONAL PENANGGULANGAN BENCANA*, Mar. 04, 2019. <https://bnpb.go.id/berita/99-penyebab-kebakaran-hutan-dan-lahan-adalah-ulah-manusia> (accessed Apr. 14, 2022).
- [6] Z. F. Abror, "Klasifikasi Citra Kebakaran Dan Non Kebakaran Menggunakan Convolutional Neural Network," *Jurnal Ilmiah Teknologi dan Rekayasa*, vol. 24, no. 2, pp. 102–113, 2019, doi: 10.35760/tr.2019.v24i2.2389.
- [7] R. Kusumawati, "Kecerdasan Buatan Manusia (Artificial Intelligence); Teknologi Impian Masa Depan," *ULUL ALBAB Jurnal Studi Islam*, vol. 9, no. 2, pp. 257–274, 2018, doi: 10.18860/ua.v9i2.6218.
- [8] W. Sun, P. Bocchini, and B. D. Davison, *Applications of artificial intelligence for disaster management*, vol. 103, no. 3. Springer Netherlands, 2020. doi: 10.1007/s11069-020-04124-3.
- [9] A. Sepriando, H. Hartono, and R. H. Jatmiko, "Deteksi Kebakaran Hutan Dan Lahan Menggunakan Citra Satelit Himawari-8 Di Kalimantan Tengah," *Jurnal Sains & Teknologi Modifikasi Cuaca*, vol. 20, no. 2, pp. 79–89, 2020, doi: 10.29122/jstmc.v20i2.3884.
- [10] M. Park, D. Q. Tran, D. Jung, and S. Park, "Wildfire-detection method using densenet and cyclegan data augmentation-based remote camera imagery," *Remote Sensing*, vol. 12, no. 22, pp. 1–16, 2020, doi: 10.3390/rs12223715.
- [11] P. Sridhar, N. Rexna Devi, S. Samyuktha, A. Sanjeev, and C. Srinivasan, "Wildfire Detection and Avoidance of false Alarm Using Densenet," *2022 13th International Conference on Computing Communication and Networking Technologies, ICCCNT 2022*, pp. 1–4, 2022, doi: 10.1109/ICCCNT54827.2022.9984493.
- [12] F. Zhuang *et al.*, "A Comprehensive Survey on Transfer Learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2021, doi: 10.1109/JPROC.2020.3004555.
- [13] F. D. Adhinata, D. P. Rakhmadani, M. Wibowo, and A. Jayadi, "A Deep Learning Using DenseNet201 to Detect Masked or Non-masked Face," *JUITA: Jurnal Informatika*, vol. 9, no. 1, p. 115, 2021, doi: 10.30595/juita.v9i1.9624.
- [14] "CIFAR-10 and CIFAR-100 datasets." <https://www.cs.toronto.edu/~kriz/cifar.html> (accessed Dec. 08, 2022).
- [15] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 2261–2269, 2017, doi: 10.1109/CVPR.2017.243.
- [16] "Forest Fire Dataset | Kaggle." <https://www.kaggle.com/datasets/alik05/forest-fire-dataset> (accessed Nov. 16, 2022).
- [17] A. Khan, B. Hassan, S. Khan, R. Ahmed, and A. Abuassba, "DeepFire: A Novel Dataset and Deep Transfer Learning Benchmark for Forest Fire Detection," *Mobile Information Systems*, vol. 2022, 2022, doi: 10.1155/2022/5358359.
- [18] "Module: tf.keras.preprocessing | TensorFlow v2.11.0." https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing (accessed Dec. 17, 2022).
- [19] "tf.keras.applications.densenet.DenseNet201 | TensorFlow v2.11.0." https://www.tensorflow.org/api_docs/python/tf/keras/application/s/densenet/DenseNet201 (accessed Dec. 17, 2022).
- [20] Darmatasia, "Analisis Perbandingan Performa Model Deep Learning untuk Mendeteksi Penggunaan Masker," *Jurnal IT*, vol. 11, no. 2, pp. 101–107, 2020.
- [21] "Introduction to DenseNets (Dense CNN) - Analytics Vidhya." <https://www.analyticsvidhya.com/blog/2022/03/introduction-to-densenets-dense-cnn/> (accessed Dec. 21, 2022).
- [22] J. Pardede and D. A. L. Putra, "Implementasi DenseNet Untuk Mengidentifikasi Kanker Kulit Melanoma," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 6, no. 3, pp. 425–433, 2020, doi: 10.28932/jutisi.v6i3.2814.
- [23] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *Journal of Biomolecular Structure and Dynamics*, vol. 39, no. 15, pp. 5682–5689, 2021, doi: 10.1080/07391102.2020.1788642.
- [24] G. Wicaksono, S. Andryana, and B. -, "Aplikasi Pendeteksi Penyakit Pada Daun Tanaman Apel Dengan Metode Convolutional Neural Network," *JOINTECS (Journal of Information Technology and Computer Science)*, vol. 5, no. 1, p. 9, 2020, doi: 10.31328/jointecs.v5i1.1221.
- [25] "Confusion Matrix untuk Evaluasi Model pada Supervised Learning | by Kuncahyo Setyo Nugroho | Medium." <https://ksnugroho.medium.com/confusion-matrix-untuk-evaluasi-model-pada-unsupervised-machine-learning-bc4b1ae9ae3f> (accessed Dec. 21, 2022).
- [26] G. M. Idroes, A. Maulana, R. Suhendra, A. Lala, and T. Karma, "Leuser Journal of Environmental Studies TeutongNet : A Fine-tuned Deep Learning Model for Improved Forest Fire Detection," vol. 1, no. 1, pp. 1–8, 2023.

Klasifikasi Kualitas dan Prediksi Kondisi Air Tanah di DKI Jakarta Menggunakan Algoritma Naïve Bayes

Arsya Fathiarahma¹, Nina Sulistiyowati¹, Taufik Ridwan¹, dan Apriade Voutama¹

¹Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang, Karawang, Indonesia

Corresponding author: Arsyafathiarahma (e-mail: arsyafathiarahma19005@student.unsika.ac.id)

ABSTRACT Groundwater remains an important source for meeting the daily clean water needs of most DKI Jakarta residents, with 32% of the population relying on this source. High population density and associated environmental issues, such as the lack of adequate water catchment and sanitation infrastructure, have triggered concerns over groundwater quality in the region. In particular, the contamination of groundwater by bacteria such as *E. coli* presents serious risks to public health such as stomach cramps, the presence of blood during diarrhea, and vomiting. This research aims to address these issues by analyzing groundwater quality in different areas of DKI Jakarta. This approach involves visual mapping using QGIS to illustrate groundwater pollution, thus facilitating a deeper understanding of groundwater conditions. The total data of 1068 used in this study was obtained from two different sources and analyzed using the Naïve Bayes Gaussian classification algorithm. The evaluation technique involved the use of cross-validation and percentage split in 10 scenarios. The results showed that the highest accuracy rate of 84.36% as well as precision values of 0.8566, recall 0.8436, and F1-score 0.8436 were obtained using the percentage split technique with a ratio of 80:20. The findings can increase public awareness of groundwater quality in DKI Jakarta and contribute to more effective planning of water resources protection efforts.

KEYWORDS Classification, Effective, Groundwater, Naïve Bayes Gaussian

ABSTRAK Air tanah tetap menjadi sumber penting untuk pemenuhan kebutuhan air bersih sehari-hari bagi sebagian besar masyarakat DKI Jakarta, dengan 32% penduduknya mengandalkan sumber ini. Kepadatan penduduk yang tinggi dan masalah lingkungan yang terkait, seperti kurangnya infrastruktur resapan air dan sanitasi yang memadai, telah memicu keprihatinan akan kualitas air tanah di wilayah ini. Terutama, terkontaminasinya air tanah oleh bakteri seperti *E. coli* menghadirkan risiko serius bagi kesehatan masyarakat seperti perut kram, adanya darah saat diare, hingga muntah-muntah. Penelitian ini bertujuan untuk mengatasi masalah tersebut dengan menganalisis kualitas air tanah di berbagai wilayah DKI Jakarta. Pendekatan ini melibatkan pemetaan visual menggunakan QGIS untuk mengilustrasikan pencemaran air tanah, sehingga memfasilitasi pemahaman yang lebih mendalam mengenai kondisi air tanah. Total data sejumlah 1068 yang digunakan dalam penelitian ini diperoleh dari dua sumber yang berbeda dan dianalisis menggunakan algoritma klasifikasi *Naïve Bayes Gaussian*. Teknik evaluasi melibatkan penggunaan *cross-validation* dan *percentage split* dalam 10 skenario. Hasil penelitian menunjukkan tingkat akurasi tertinggi sebesar 84.36% serta nilai *precision* sejumlah 0.8566, *recall* 0.8436, dan *F1-score* 0.8436 diperoleh menggunakan teknik *percentage split* dengan rasio 80:20. Temuan ini dapat meningkatkan kesadaran masyarakat tentang kualitas air tanah di DKI Jakarta dan berkontribusi pada perencanaan upaya perlindungan sumber daya air yang lebih efektif.

KATA KUNCI Air tanah, Efektif, Klasifikasi, *Naïve Bayes Gaussian*

I. PENDAHULUAN

Air adalah suatu senyawa kimia yang sangat dibutuhkan untuk keberlangsungan hidup di bumi. Sebagian besar wilayah permukaan bumi, sebanyak 72% ditutupi oleh air. Pernyataan tersebut dikuatkan oleh penelitian yang dilakukan oleh agensi ilmiah pemerintah Amerika Serikat, USGS (United States Geological Survey). Menurut USGS, jumlah volume air di bumi sekitar 332.500.000 mil kubik atau 1.4 juta km³ dengan sumber air yang terdiri dari air laut, air permukaan, air atmosfer, air sungai, dan air tanah[1]

Air disebut sebagai sumber kehidupan dilandasi dengan banyaknya kebutuhan yang dapat terpenuhi diantaranya adalah proses fotosintesis yang dilakukan oleh tumbuhan serta sumber minum dan tempat tinggal untuk hewan. Selain itu, bergantungnya kepada air dirasakan oleh manusia untuk keperluan sehari-hari seperti mandi, memasak, mencuci dan lain-lain[2]. Standar kelayakan kebutuhan air bersih berjumlah 49,5 liter per kapita per hari dan untuk kebutuhan tubuh manusia diperlukan sekiranya 2.5 liter/hari, jumlah tersebut tidaklah kecil, sehingga perlu adanya pemilihan sumber air yang memadai. Jenis air tanah dangkal atau air sumur adalah salah satu sumber air untuk memenuhi kebutuhan tersebut[3].

DKI Jakarta merupakan salah satu daerah yang sebanyak 32% masyarakat masih memanfaatkan air tanah untuk memenuhi kebutuhan sehari-hari. Menurut ahli Hidrologi dari Universitas Brawijaya, Dr Gunawan Wibisono, perlu adanya sistem pengelolaan air yang baik dan teratur serta pengawasan dan penelusuran sumber air tanah yang akan dikonsumsi. Buruknya sanitasi di DKI Jakarta, membuat air limbah seperti septic tank mudah terkontaminasi dengan bakteri, salah satunya bakteri *e.coli*. Bakteri *Escherichia coli* (*E. coli*) adalah salah satu jenis bakteri yang menyebabkan infeksi usus pada manusia dengan menghasilkan racun yang disebut verotoksin sebagai penyebab kerusakan dinding usus kecil dan menyebabkan gejala seperti perut kram, adanya darah saat diare, hingga muntah-muntah[4]

Data mining merupakan sebuah proses mengekstrak informasi dan wawasan yang berguna dari data yang besar dan kompleks. Proses ini menggunakan teknik-teknik analisis data yang canggih untuk mengidentifikasi pola, hubungan, dan tendensi yang mungkin tidak terlihat secara langsung dari data tersebut. Guna menganalisa atau memperkirakan data hasil ukur selanjutnya mengenai kualitas air tanah di DKI Jakarta, dapat dilakukan proses penambahan data dari data sebelumnya. Tahapan pengolahan data melibatkan beberapa langkah, seperti pengumpulan data awal, pembersihan data dari *outlier* atau data yang tidak valid, transformasi data ke format yang sesuai, dan pemilihan atribut atau fitur yang relevan untuk analisis.

Selain dilakukannya penambahan data, untuk meningkatkan kesadaran masyarakat terhadap kualitas air tanah di DKI Jakarta, penelitian ini bertujuan memberikan informasi yang dapat disebarluaskan kepada masyarakat

dengan visualisasi data kualitas air tanah yang diintegrasikan dalam bentuk peta interaktif, sehingga informasi tersebut dapat lebih mudah dipahami dan diakses oleh masyarakat umum. Dengan adanya pemaparan mengenai kualitas air tanah serta kandungan kimia atau bakteri yang berakibat fatal bagi tubuh, diharapkan masyarakat mencegah dampak negatif tersebut pada kesehatan, menjadi lebih sadar akan pentingnya menjaga kualitas air tanah dan dapat mengambil tindakan yang sesuai untuk melindungi sumber daya air yang begitu berharga ini.

Penelitian terdahulu yang dilakukan oleh M. Ja'far Sodik pada tahun 2019 mengatakan bahwa klasifikasi kualitas udara di DKI Jakarta yang dilakukan dengan metode *Naive Bayes* mendapatkan nilai *accuracy* sebesar 91.862 % [5] Selain itu, penelitian lain dilakukan oleh Dani Purwanto dalam kasus klasifikasi *naive Bayes* untuk memprediksi kualitas udara berdasarkan Indeks Standar Pencemar Udara (ISPU)[6] yang memiliki *accuracy* sebesar 93%. Kemudian, penelitian yang dilakukan oleh Aditya Rizky Sanjaya di tahun 2019 dengan judul klasifikasi jenis mobil paling diminati di Indonesia dengan menggunakan algoritma *naive Bayes* mendapatkan *accuracy* 96,7[7]. Namun, perlu diperhatikan bahwa penelitian sebelumnya belum adanya fokus pada kualitas air tanah dan penggunaan algoritma *Naive Bayes Gaussian* untuk tujuan tersebut. Oleh karena itu, tidak dapat dilakukan perbandingan rinci dengan penelitian sebelumnya. Kombinasi antara kualitas air tanah dan algoritma *Naive Bayes Gaussian* ini merupakan hal yang baru dan inovatif, yang diharapkan dapat memberikan wawasan berharga mengenai kualitas air tanah di DKI Jakarta.

Berdasarkan dari beberapa penelitian sebelumnya, metode *naive Bayes* merupakan metode yang memiliki *accuracy* cukup tinggi. Kelebihan penggunaan algoritma *Naive Bayes Gaussian* adalah kemampuannya untuk mengatasi data yang memiliki sifat kontinu [8], seperti konsentrasi bahan kimia dalam air tanah. Algoritma ini juga dikenal karena sederhananya, yang memungkinkan pemrosesan cepat dan efisien. Selain itu, *Naive Bayes* biasanya memerlukan jumlah data pelatihan yang lebih kecil dibandingkan dengan beberapa algoritma lainnya, yang sangat bermanfaat karena data kualitas air tanah di DKI Jakarta terbatas dalam jumlahnya [9]. Dengan dilakukannya beberapa skema pada tahap pemodelan, diharapkan dapat menghasilkan model dengan akurasi terbaik dalam klasifikasi kualitas air tanah di DKI Jakarta

II. TINJAUAN PUSTAKA

A. DATA MINING

Data mining merupakan suatu proses yang memanfaatkan teknik dan algoritma komputasi untuk mengolah data secara masif dalam jumlah yang sangat besar dan bervariasi[10]. *Data mining* bermula dari pertemuan antara berbagai disiplin ilmu seperti *machine learning*, *pattern recognition*, *artificial intelligence*, dan sistem basis data, yang semuanya memiliki

tujuan untuk menangani data yang sangat besar, heterogen, dan tersebar [11]

Data mining sangat berguna dalam menyelesaikan masalah yang dihadapi oleh perusahaan, organisasi, atau individu yang mengumpulkan data secara masif dan ingin mengetahui pola atau tren dari data tersebut[12]. Dengan menggunakan data mining, kita dapat menemukan pola-pola yang tidak terlihat secara visual dari data yang ada, sehingga dapat memberikan informasi yang berguna bagi pengambilan keputusan atau perencanaan strategis[13]. *Data mining* juga dapat digunakan untuk mengoptimalkan sistem informasi, meningkatkan efisiensi proses bisnis, dan meningkatkan keuntungan[14]

B. NAÏVE BAYES

Naive bayes merupakan salah satu metode populer dalam pengklasifikasian probabilitas sederhana untuk keperluan data mining karena dengan struktur yang cukup sederhana, metode ini dapat mudah diimplemntasikan serta tingginya efektifitas yang dihasilkan[15]. Metode yang dikemukakan oleh ilmuwan Inggris Thomas Bayes ini dikenal sebagai teorema bayes yaitu memprediksi peluang kedepannya, dengan menggunakan pengalaman sebelumnya sebagai bahan perhitungan[16]. Arti naif dari metode ini adalah karena membuat asumsi kemunculan fitur tidak saling bergantung dan arti tersebut merupakan ciri dari *naive bayes*, yaitu independent[17]. Proses klasifikasi pada metode *naive bayes* memerlukan sebuah data set yang terdiri dari *data training* atau dapat disebut data latih dan *data testing* atau dapat disebut data uji yang kemudian akan diproses dan membentuk pengetahuan berupa nilai probabilitas pada setiap data[18]. Metode *naive bayes* bertujuan untuk melakukan klasifikasi data pada kelas tertentu dengan kelebihan dapat menangani data baik yang bersifat diskrit maupun kontinu[8]. Pada atribut data yang bersifat kontinu, metode yang digunakan yaitu *naive bayes gaussian*, dimana Persamaan (1) sebagai metode *naive bayes gaussian*.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

C. CONFUSION MATRIX MULTICLASS

Confusion matrix multiclass merupakan perkembangan dari *matriks confusion binary*, di mana sebelumnya terdapat FN (*False Negative*), FP (*False Positive*), dan TN (*True Negative*). Namun, pada *matriks confusion multiclass*, hanya terdapat TP karena untuk menentukan FN adalah dari jumlah baris keseluruhan per variabel, sedangkan untuk menentukan FP adalah dari jumlah kolom keseluruhan per variabel, dan TN adalah kasus-kasus di mana prediksi kita tidak ada dan nilai aktualnya salah.

Confusion matriks multiclass adalah tabel yang umumnya digunakan untuk menjelaskan kinerja model klasifikasi pada dataset pengujian tertentu, di mana nilai sebenarnya sudah diketahui[19]. Contoh *confusion matriks* dapat dilihat pada Gambar 1

		Predicted				
		Label 1	Label 2	...	Label n	
Actual	Label 1	TP	E	...	E	Total Baris 1
	Label 2	E	TP	...	E	Total Baris 2

	Label n	E	E	...	TP	Total Baris n
		Total Kolom 1	Total Kolom 2	...	Total Kolom n	

GAMBAR 1. *Confusion matrix multiclass*

Hasil evaluasi *confusion matrix multiclass* berupa nilai *accuracy*, *precision* dan *recall*. *Accuracy* adalah jumlah proporsi prediksi benar dengan membandingkan dengan total seluruh data. Adapun rumus perhitungan yang digunakan untuk menghitung *accuracy* dapat dilakukan dengan Persamaan (2).

$$\frac{TP(\text{label } 1) + \dots + TP(\text{label } n)}{\text{total}(\text{baris } 1) + \dots + \text{total}(\text{baris } n)} \tag{2}$$

Recall adalah jumlah proporsi prediksi benar dengan membandingkan dengan total data positif. Adapun rumus perhitungan yang digunakan untuk menghitung *recall* dapat dilakukan dengan Persamaan (3).

$$\frac{TP(\text{label } i)}{\text{Total}(\text{baris } i)} \tag{3}$$

Precision adalah jumlah proporsi prediksi benar dengan membandingkan dengan yang diprediksi positif. Adapun rumus perhitungan yang digunakan untuk menghitung *precision* dapat dilakukan dengan Persamaan (4).

$$\frac{TP(\text{label } i)}{\text{Total}(\text{kolom } i)} \tag{4}$$

F1-score adalah rata-rata harmoni *precision* dan *recall* dengan menggambarkan keseluruhan kinerja model yang memberikan keseimbangan antara *precision* dan *recall*. Adapun rumus perhitungan yang digunakan untuk menghitung *F1-score* dapat dilakukan dengan Persamaan (5).

$$2 \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \tag{5}$$

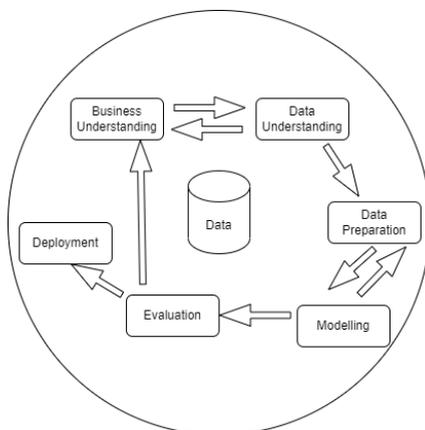
D. PERCENTAGE SPLIT

Percentage split adalah salah satu opsi pengujian yang tersedia dalam data mining. Hasil klasifikasi *percentage split* diuji dengan menggunakan bagian dari dataset sebagai *data training*. Dalam *percentage split*, data dibagi secara acak menjadi dua kelompok terpisah. Kelompok pertama digunakan untuk mengambil pengetahuan dari *data training*, dan pengetahuan yang diambil kemudian diuji pada kelompok kedua yang disebut sebagai data uji. Pembagian data acak dapat dilakukan dengan beberapa rasio yang disesuaikan dengan kebutuhan penelitian[20]

E. CROSS VALIDATION

K-Fold Cross validation merupakan metode statistik yang digunakan untuk evaluasi kinerja model klasifikasi dengan membagi data menjadi 2 bagian yaitu data uji dengan guna untuk proses pembelajaran untuk membentuk model klasifikasi dan data latih dengan guna melakukan validasi terhadap model tersebut. *K-Fold Cross validation*, data latih dibagi menjadi K bagian yang disebut "*fold*". Setiap *fold* kemudian digunakan secara bergantian sebagai data uji dan sisanya sebagai data latih. Proses ini dilakukan K kali, dengan setiap *fold* digunakan sekali sebagai data uji. Setelah *K-Fold Cross validation* selesai dilakukan, kita dapat menghitung rata-rata dari K hasil evaluasi yang diperoleh untuk mendapatkan estimasi kinerja model secara keseluruhan. Dengan menggunakan *K-Fold Cross validation*, kita dapat memperoleh estimasi yang lebih akurat dari kinerja model klasifikasi kita karena data latih yang lebih banyak digunakan untuk validasi[21]

III. METODOLOGI PENELITIAN



GAMBAR 2. Metodologi CRISP-DM

Metodologi yang dilakukan pada penelitian ini dipaparkan pada Gambar 2, yaitu metode CRISP-DM dengan enam proses, *bussines understanding*, *data understanding*, *data preparation*, *modelling*, *evaluation* dan *deployment*.

A. BUSINESS UNDERSTANDING

Tahap ini akan dilakukannya pemahaman penelitian secara menyeluruh pada laporan akhir dari pemantauan kualitas air di DKI Jakarta sehingga mendapatkan tujuan yang akan dicapai pada penelitian ini. Kegiatan tersebut dilakukan terhadap data, serta dokumen yang telah disediakan Dinas Lingkungan Hidup DKI Jakarta dan dapat diunduh secara bebas oleh publik. Setelah dilakukannya pemahaman terhadap data dan dokumen yang ada, maka selanjutnya menyiapkan rencana dan strategi untuk mencapai tujuan tersebut.

B. DATA UNDERSTANDING

Tahap ini akan dilakukannya pengumpulan data yang didapat dari Dinas Lingkungan Hidup DKI Jakarta dan Open Data DKI Jakarta. Data tersebut dieksplor secara menyeluruh, kemudian dipelajari agar dapat menemukan wawasan awal atau informasi secara spesifik yang mempunyai karakteristik, sehingga akan didapatkannya kualitas dan kelengkapan data

C. DATA PREPARATION

Tahap ini akan dilakukannya suatu tahap intensif karena perlu adanya analisa data secara rinci agar hasil yang diinginkan dapat dicapai. Pengolahan data atau persiapan data dilakukan sebelum proses klasifikasi dilakukan dapat disebut preprocessing. Beberapa preprocessing yang akan dilakukan diantaranya pemilihan data dan atribut yang akan digunakan pada penelitian untuk proses klasifikasi, selanjutnya penggabungan data dari berbagai sumber yang kemudian dipadukan satu sama lain, melakukan format data agar data dapat diolah oleh tools yang digunakan, pemilihan feature dan label, dan preprocessing terakhir adalah normalisasi data

D. MODELLING

Tahap ini akan dilakukannya training dan testing terhadap data yang sudah dikumpul dan diolah. Digunakannya algoritma *naïve bayes* dengan dua teknik yang berbeda, *cross validation* dan *percentage split* serta bantuan tools google colab yang di dalamnya terdapat beberapa library python menjadikan adanya perbedaan hasil sebagai bahan perbandingan agar dapat memilih skenario dengan hasil terbaik

E. EVALUATION

Tahap ini dilakukannya interpretasi dari hasil pemodelan data mining. Berdasarkan tujuan yang telah dijelaskan pada tahap bussines understanding, maka pada tahap ini akan dilakukan Analisa hasil klasifikasi data prediksi kualitas air tanah. Evaluasi dilakukan dengan pengimplementasian *confusion matrix* dengan 2 teknik data mining, *cross validation* serta *percentage split* sehingga akan mendapatkan hasil yang telah direncanakan.

F. DEPLOYMENT

Tahap ini akan dilakukannya penulisan laporan hasil kegiatan yang sudah dilakukan dan tinjauan mengenai kualitas air tanah di DKI Jakarta pada tahun 2019 dengan 2 periode dan 2021 dengan 2 periode divisualisasikan dengan menggunakan bantuan tools QGIS. Pada laporan akhir didapatkan mengenai hasil analisis dari awal sampai hasil evaluasi yang mencakup *accuracy*, *precision*, *recall*, dan *f1-score*.

IV. HASIL DAN PEMBAHASAN

A. BUSINESS UNDERSTANDING

Pemahaman bisnis berdasarkan penelitian kualitas air tanah di DKI Jakarta pada tahun 2019 dan 2021 menghasilkan 3 poin utama diantaranya *determine business objectives, asses the situation, dan determine data mining goal.*

1) DETERMINE BUSINESS OBEJECTIVES

Tujuan bisnis dari penelitian ini berdasarkan data dan laporan akhir kualitas air tanah di DKI Jakarta yaitu mendapatkan hasil mengenai mutu air tanah di DKI Jakarta pada tahun 2019 dan 2021. Selain itu, peneletian ini mempunyai tujuan lain diantaranya dapat mengevaluasi hasil *accuracy* dari kualitas kondisi air tanah serta performa terkait *confusion matrix*. Parameter suksesnya prediksi kualitas kondisi air tanah dilihat dari beberapa atribut diantaranya kelurahan, kecamatan, nilai pH, deterjen, mangan, total coliform, e.coli, nilai indeks pencemaran dan status mutu.

2) ASESS THE SITUATION

Kondisi data awal yang diperoleh mempunyai bentuk yang berbeda karena didapatkan dari 2 sumber. Pada tahun 2019, sumber data diambil dari open data DKI Jakarta dan untuk tahun 2021, data diambil dari Dinas Lingkungan Hidup DKI Jakarta pada dokumen yang berjudul "Laporan Akhir Pemantauan Kualitas Air Tanah Tahun 2021". Data awal 2019 mempunyai 20 paramterer sebagai bahan perhitungan indeks pencemaran dalam bentuk csv. Namun, pada tahun 2021, data diperoleh dari grafik dengan 5 parameter di dalamnya, diantaranya, pH, mangan, deterjen, total coliform, dan e.coli. Selain grafik, terdapat data berbentuk tabel yang menjelaskan mengenai nilai indeks pencemaran serta status mutu pada setiap kelurahan. Terdapat beberapa data yang kurang relevan untuk dijadikan bahan penelitian.

3) DETERMINE DATA MINING GOAL

Penelitian dilakukan guna mengukur kinerja algoritma *naive bayes* untuk klasifikasi kualitas kondisi air tanah di DKI Jakarta menggunakan data 2019 dan 2021 dengan beberapa skenario yang berbeda. Dari hasil klasifikasi skenario tersebut, akan diketahui hasil algoritma *naive bayes* dengan kinerjanya. Selain itu, penelitian ini mempunyai tujuan untuk menemukan informasi baru terkait kualitas kondisi air tanah di DKI Jakarta dari data yang sudah diolah sehingga dapat menjadikan dasar atau bahan acuan untuk dilakukannya penelitian selanjutnya

B. DATA UNDERSTANDING

Pemahaman data dilakukan dengan menggunakan dua sumber data mengenai kualitas air tanah di DKI Jakarta, yaitu open data Jakarta dan website Dinas Lingkungan Hidup DKI Jakarta. Data yang digunakan dalam penelitian ini adalah data tahun 2019 dan 2021. Data untuk tahun 2019 diambil dari open data DKI Jakarta dalam bentuk file csv, sedangkan data untuk tahun 2021 diambil dari laporan pemantauan kualitas air

tanah yang diterbitkan oleh Dinas Lingkungan Hidup DKI Jakarta.

Alasan pemilihan kedua tahun tersebut adalah jarak waktu yang tidak terlalu jauh dari dilakukannya penelitian ini. Tahun 2020 tidak digunakan karena tidak dilakukannya pemantauan kualitas air tanah. Pada tahun 2020, hanya dilakukan penyusunan analisis kualitas air tanah dalam kurun waktu 5 tahun terakhir.

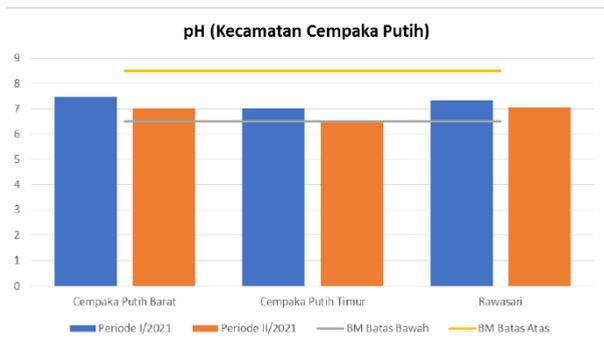
Sumber pertama adalah situs web open data DKI Jakarta pada tahun 2019 yang dikumpulkan secara terpisah antara periode 1 dan periode 2 dalam bentuk csv. Jumlah data kualitas air tanah yang terkumpul pada masing-masing periode adalah 267. Jumlah tersebut dihasilkan dari 5 Kota Administrasi di DKI Jakarta yaitu Jakarta Pusat, Jakarta Selatan, Jakarta Timur, Jakarta Utara dan Jakarta Barat. Parameter yang digunakan atas penilaian status mutu di data tahun 2019 diantaranya warna, kekeruhan, air raksa, besi, fluoride, cadmium, kesadahan, krom heksavalen, mangan, nitrat, nitrit, seng, sulfat, timah hitam, deterjen, zat organic, total coliform, dan e.coli. Gambar 3 merupakan dataset di tahun 2019

1	kecamatan,kelurahan,lintang_selatan,bujur_timur,parameter,hasil_analisa
2	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Warna,21
3	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Kekeruhan,0,98
4	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Air Raksa,"<0,3"
5	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Besi (Fe),"<0,8"
6	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Fluorida,"0,43"
7	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Cadmium (Cd),0
8	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Kesadahan,"55,15"
9	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Krom Heksavalen,"<0,3"
10	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Mangan (Mn),"<0,5"
11	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Nitrat,"0,22"
12	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Nitrit,"0,6"
13	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Seng (Zn),"0,1"
14	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Sulfat,"40,85"
15	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Timah Hitam (Pb),0
16	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Detergen,"0,6"
17	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Zat Organik,"1,82"
18	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Total Coliform,100
19	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,E.Coli,100
20	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Indeks Pencemaran,"1,786,711,641"
21	Gambir,Cideng,"06o10'05,1" " " " 106o48'33,5"" ,Kategori,Tercemar Ringan

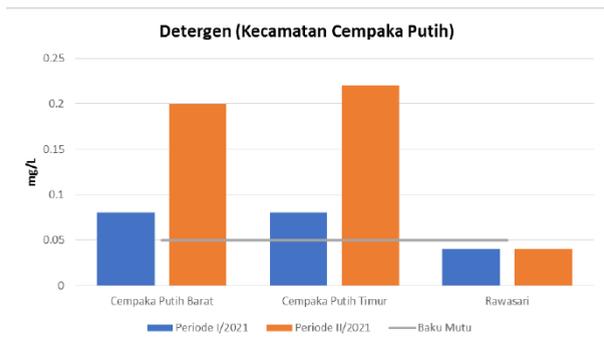
GAMBAR 3. Dataset 2019

Untuk sumber kedua yaitu didapatkan dari laporan kualitas air tanah di DKI Jakarta yang diterbitkan oleh Dinas Lingkungan Hidup DKI Jakarta dengan jumlah data dan Kota Adminstrasi yang digunakan sama dengan data di tahun 2019. Perbedaan dari kedua data tersbut terlak pada bentuk data yang didapatkan dan parameter yang digunakan guna klasifikasi status mutu air tanah. Bentuk data yang didapatkan dari laporan kualitas air tanah di DKI Jakarta yaitu grafik dan tabel dari 5 parameter yang mendominasi dan belum memenuhi -- baku mutu. Kelima parameter tersebut adalah pH, deterjen, mangan, e.coli, dan total coliform. Dengan contoh Kecamatan Cempaka Putih yang dibagi menjadi 3 kelurahan, Cempaka Putih Barat, Cempaka Putih Timur, dan Rawasari, Gambar 4 merupakan salah satu contoh bentuk data mengenai parameter pH, Gambar 5 merupakan salah satu contoh bentuk data mengenai parameter deterjen, Gambar 6 merupakan salah satu

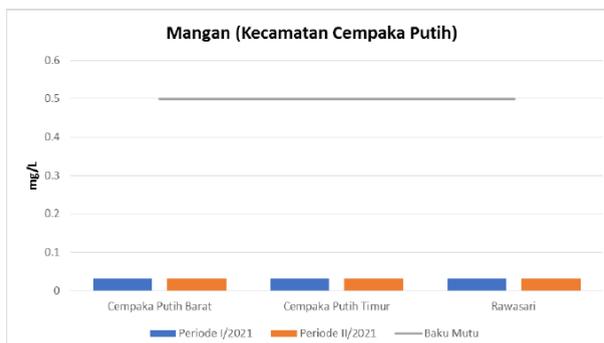
contoh bentuk data mengenai parameter mangan, Gambar 7 merupakan salah satu contoh bentuk data mengenai parameter total coliform, dan Gambar 8 merupakan salah satu contoh bentuk data mengenai parameter e.coli.



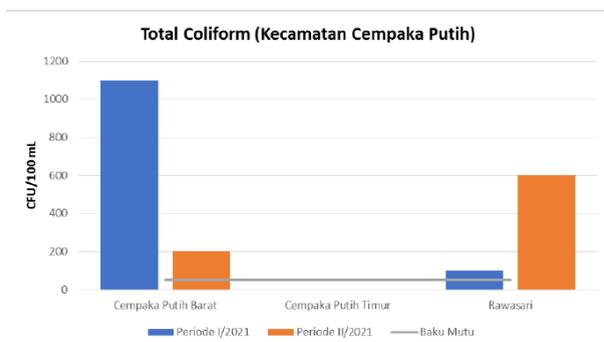
GAMBAR 4. Grafik pH Kecamatan Cempaka Putih



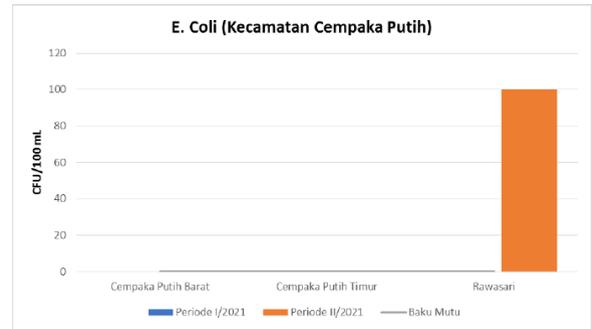
GAMBAR 5. Grafik Deterjen Kecamatan Cempaka Putih



GAMBAR 6. Grafik Mangan Kecamatan Cempaka Putih



GAMBAR 7. Grafik Total Coliform Kecamatan Cempaka Putih



GAMBAR 8. Grafik E.Coli Kecamatan Cempaka Putih

Selain paparan grafik dengan menggambarkan berapa nilai jumlah kandungan parameter pada tiap kelurahan pada 2 periode, terdapat paparan lanjutan mengenai hasil perhitungan kandungan 5 parameter yang menghasilkan nilai indeks pencemaran serta status mutu yang diperoleh. Tabel I adalah paparan dari hasil perhitungan grafik dari 5 parameter tersebut.

TABEL I
NILAI IP, STATUS MUTU, PARAMETER KRITIS KECAMATAN CEMPAKA PUTIH

Periode I				
No Titik	Kelurahan	Nilai IP	Status Mutu	Parameter Kritis
26	Cempaka Putih Barat	5.48	Cemar Sedang	Detergen Total Coliform
27	Cempaka Putih Timur	1.15	Cemar Ringan	Detergen
28	Rawasari	1.79	Cemar Ringan	Total Coliform
Periode II				
No Titik	Kelurahan	Nilai IP	Status Mutu	Parameter Kritis
26	Cempaka Putih Barat	2.87	Cemar Ringan	Detergen Total Coliform
27	Cempaka Putih Timur	2.99	Cemar Ringan	Detergen
28	Rawasari	4.54	Cemar Ringan	Total Coliform

Status mutu yang ditetapkan pada setiap kelurahan dinilai berdasarkan besarnya nilai IP atau indeks pencemaran. Terdapat kriteria yang telah dibuat oleh Dinas Lingkungan Hidup DKI Jakarta, yaitu, jika IP berada dalam rentang antara 0 hingga 1,0, maka kondisinya dapat diklasifikasikan sebagai "Baik," jika IP berada dalam rentang antara 1,0 hingga 5,0, maka kondisinya adalah "Cemar Ringan," ketika IP berada dalam rentang antara 5,0 hingga 10, kondisinya adalah "Cemar Sedang," terakhir, jika IP melebihi angka 10, maka kondisinya dapat diklasifikasikan sebagai "Cemar Berat".

Adanya perbedaan bentuk perolehan data, perlu dilakukannya perubahan atas data 2021 agar dapat mempermudah pembacaan, mempermudah melakukan sinkronisasi dengan data 2019, dan mempermudah

TABEL II
 PERUBAHAN DATASET 2021

Kota	Kecamatan	Kelurahan	PH	Detergen	Mangan	Total Coliform	E_Coli	Indeks	Mutu
Jakarta Pusat	Gambir	Cideng	7,2	0,6	0,3	100	100	1,78	Cemar Ringan
Jakarta Pusat	Gambir	Duri Pulo	7,6	2,94	1,51	0	0	0,72	Baik
Jakarta Pusat	Gambir	Gambir	7,9	0,14	0,6	100	0	2,74	Cemar Ringan
Jakarta Pusat	Gambir	Kebon Kelapa	7,5	0,6	0,3	0	0	0,99	Baik
Jakarta Pusat	Gambir	Petojo Selatan	7,7	0,6	0,3	0	0	1,31	Cemar Ringan
Jakarta Pusat	Gambir	Petojo Utara	6,9	0,8	0,3	100	0	1,79	Cemar Ringan
Jakarta Pusat	Sawah Besar	Gunung Sahari Utara	7,5	0,2	0,29	0	0	0,44	Baik
Jakarta Pusat	Sawah Besar	Karang Anyar Utara	7,7	0,5	0,3	0	0	0,71	Baik
Jakarta Pusat	Sawah Besar	Kartini	7,9	0,4	0,3	0	0	0,58	Baik
Jakarta Pusat	Sawah Besar	Mangga Dua Selatan	7,3	0,18	0,35	100	0	2,7	Cemar Ringan
Jakarta Pusat	Cempaka Putih	Cempaka Putih Barat	7,3	0,9	0,8	0	0	1,62	Cemar Ringan
Jakarta Pusat	Cempaka Putih	Cempaka Putih Timur	7,3	0,5	0,3	0	0	0,71	Baik
Jakarta Pusat	Cempaka Putih	Rawasari	6,7	0,5	0,35	600	0	4,54	Cemar Ringan

preprocessing yang akan dilakukan pada penelitian ini. Tabel II merupakan hasil perubahannya dengan format dokumen di kecamatan Cempaka Putih serta beberapa kecamatan lainnya

C. DATA PREPARATION

Preprocessing dilakukan untuk menciptakan dataset klasifikasi kualitas air tanah di DKI Jakarta dapat digunakan dalam pemodelan naïve bayes. Beberapa tahapan ini merupakan cara agar mendapatkan hasil data yang bersih, terorganisir, dan lebih siap untuk dianalisa dalam penelitian ini.

1) PEMILIHAN ATRIBUT DAN LABEL

Berdasarkan *dataset* 2019, terdapat 6 atribut seperti yang dipaparkan pada Gambar 3. Sedangkan untuk *dataset* 2021, terdapat 10 atribut seperti yang dipaparkan pada Tabel II. Namun pada penelitian ini, atribut yang digunakan akan berdasarkan pada *dataset* di tahun 2021 karena data tersebut merupakan data terbaru yang diterbitkan oleh Dinas Lingkungan Hidup Jakarta. Oleh karena itu, Atribut yang akan digunakan berjumlah 10 diantaranya kota, kecamatan, kelurahan, pH, deterjen, mangan, total coliform, e.coli, indeks, dan status mutu. Berdasarkan atribut yang ada pada data, label yang akan digunakan pada penelitian ini adalah status mutu yang dibagi menjadi 4, yaitu Baik, Cemar Ringan, Cemar Sedang dan Cemar Berat.

2) NORMALISASI

Perolehan dataset dengan range nilai pada e.coli dan coliform yang terlampaui jauh dibutuhkannya normalisasi untuk meningkatkan *accuracy* model. Dihasilkan normalisasi dengan nilai pada setiap atribut numerik dari rentang 0 - 1.

D. MODELLING

Pemodelan dilakukan menggunakan algoritma *naïve bayes* dengan 2 teknik, yaitu *cross validation* serta *percentage split*.

Teknik *cross validation* dilakukan dengan 6 skenario yang terdiri dari nilai *fold* 5 sampai dengan *fold* 10.

Teknik *percentage split* dilakukan dengan 4 skenario yang terdiri dengan perbandingan rasio 80 % : 20%, 70 % : 30%, 60 % : 50%, dan 50 % : 50%. Rasio pemisahan terdiri *data training* dan *data testing*

 TABEL III
 ACCURACY DAN ERROR RATE

Cross Validation		
Scenario	Accuracy	Error Rate
Skema 1	83,89%	16,10%
Skema 2	83,33%	16,66%
Skema 3	83,33%	16,66%
Skema 4	83,80%	16,19%
Skema 5	83,33%	16,66%
Skema 6	83,80%	16,19%
Percentage split		
Scenario	Accuracy	Error Rate
Skema 1	84,36%	15,63%
Skema 2	84,26%	15,73%
Skema 3	83,61%	16,38%
Skema 4	83,23%	16,76%

 TABEL IV
 PRECISION, RECALL, DAN F1-SCORE

Cross Validation			
Scenario	Precision	Recall	F1-Score
Skema 1	0,8478	0,839	0,8389
Skema 2	0,843	0,8333	0,8333
Skema 3	0,8431	0,8333	0,8332
Skema 4	0,8488	0,838	0,8382
Skema 5	0,844	0,8333	0,8333
Skema 6	0,8499	0,838	0,8379
Percentage split			
Scenario	Precision	Recall	F1-Score
Skema 1	0,8566	0,8436	0,8436
Skema 2	0,8556	0,8427	0,8426
Skema 3	0,8517	0,8361	0,8361
Skema 4	0,8478	0,8324	0,8322

E. EVALUATION

Berdasarkan 10 skema yang sudah dipaparkan pada tahap *modelling*, dengan 6 skema *cross validation* dan 4 skema *percentage split*, Tabel III adalah hasil *accuracy*, *error rate*, untuk Tabel IV adalah hasil *precision*, *recall*, dan *f1-score* dari 10 skema yang dilakukan.

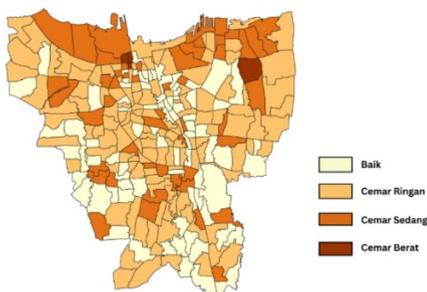
Pengujian *cross validation* pada skema 1 dengan folds 5 menghasilkan akurasi tertinggi yaitu 83.89 % dengan *error rate* 16.10% dan juga menempati nilai tertinggi pada *precision* dengan 0.8478, *recall* 0.8390, dan *F1-score* 0.8389, seperti yang dipaparkan pada Tabel III dan Tabel IV.

Pengujian *percentage split* pada skema 1 dengan rasio 80:20 menghasilkan nilai *accuracy* tertinggi yaitu 84.36% dengan *error rate* 15.63%. dan juga menempati nilai tertinggi pada *precision* 0.8566, *recall* 0.8436, dan *F1-score* 0.8436.

F. DEPLOYMENT

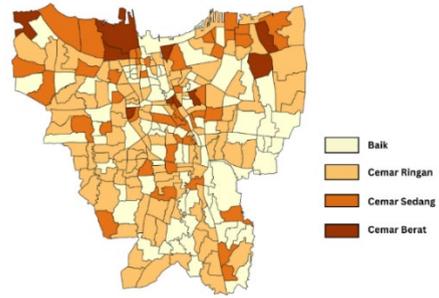
Deployment pada penelitian klasifikasi kualitas dan prediksi kondisi air tanah di DKI Jakarta dengan menggunakan algoritma naïve bayes, menghasilkan laporan penelitian. Laporan akhir tersebut berisi paparan tentang penggunaan teknik *modelling cross validation* dan *percentage split* pada data kualitas air tanah 2019 dan 2021 yang telah melalui tahap *preprocessing* sehingga diperoleh data yang bersih dan terorganisir. Selanjutnya, dilakukan perbandingan *accuracy* serta kinerja teknik algoritma terbaik yang dapat dijadikan dasar atau bahan acuan untuk penelitian selanjutnya.

Selain pembuatan laporan akhir, terdapat paparan mengenai pemantauan kualitas air tanah di DKI Jakarta dengan 2 periode di tahun 2019 serta 2 periode di tahun 2021.



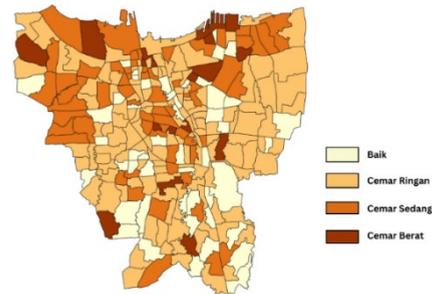
GAMBAR 9. Pemetaan Kualitas Air Tanah Tahun 2019 Periode 1

Berdasarkan data yang didapatkan dari Open Data DKI Jakarta dan visualisasi pada Gambar 9, untuk kualitas air tanah di DKI Jakarta pada tahun 2019 periode 1 terdapat 4 wilayah dengan status mutu air tanah cemar berat, diantaranya yaitu Kota Administrasi Jakarta Utara pada kelurahan Sukapura, dan Kota Administrasi Jakarta Barat pada kelurahan Pekojan, Roamalaka, dan Tambora. Selain adanya kelurahan yang memiliki status mutu air cemar berat, pada Kota Administrasi Jakarta Utara mendominasi memiliki status mutu air cemar sedang.



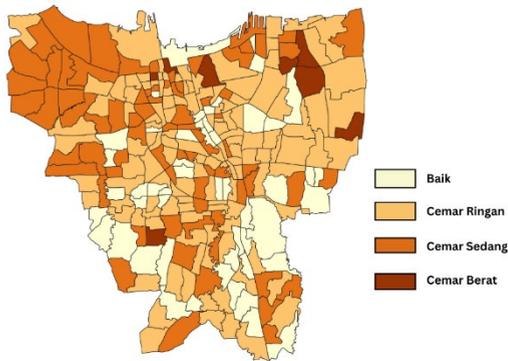
GAMBAR 10. Pemetaan Kualitas Air Tanah Tahun 2019 Periode 2

Berdasarkan data yang didapatkan dari Open Data DKI Jakarta dan visualisasi pada Gambar 10, untuk kualitas air tanah di DKI Jakarta pada tahun 2019 periode 2 terdapat kenaikan jumlah wilayah dengan status mutu air tanah Cemar Berat dengan total 10 kelurahan, diantaranya yaitu Kota Administrasi Jakarta Utara pada kelurahan Pademangan Barat, Pejagalan, Pluit, Semper Timur, dan Sukapura. Kota Administrasi Jakarta Pusat pada kelurahan Cempaka Putih Barat, Kramat dan Kwitang. Kota Administrasi Jakarta Barat pada kelurahan Kamal dan Slipi. Untuk dominasi status mutu air cemar sedang berada pada Kota Administrasi Jakarta Barat.



GAMBAR 11. Pemetaan Kualitas Air Tanah Tahun 2021 Periode 1

Berdasarkan data yang didapatkan dari laporan akhir pemantauan kualitas air tanah tahun 2021 yang diterbitkan oleh Dinas Lingkungan Hidup DKI Jakarta dan visualisasi pada Gambar 11, untuk kualitas air tanah di DKI Jakarta pada tahun 2021 periode 1 terdapat kenaikan jumlah wilayah dengan status mutu air tanah Cemar Berat dengan total 18 kelurahan, diantaranya yaitu Kota Administrasi Jakarta Utara pada kelurahan Kapuk Muara, Koja, Sunter Jaya, Tanjung Priok, Tugu Selatan, dan Warakas. Kota Administrasi Jakarta Pusat pada kelurahan Serdang. Kota Administrasi Jakarta Barat pada kelurahan, Glodok, Pegadungan, Pinangsia dan Tangki. Kota Administrasi Jakarta Selatan pada kelurahan Kalibata, Lebak Bulus, Pasar Manggis, dan Setiabudi. Kota Administrasi Jakarta Timur pada kelurahan Cijantung, Cipinang Muara dan Kebon Manggis.



GAMBAR 12. Pemetaan Kualitas Air Tanah Tahun 2021 Periode 2

Berdasarkan data yang didapatkan dari laporan akhir pemantauan kualitas air tanah tahun 2021 yang diterbitkan oleh Dinas Lingkungan Hidup DKI Jakarta dan visualisasi pada Gambar 12, untuk kualitas air tanah di DKI Jakarta pada tahun 2021 periode 2 terdapat penurunan jumlah wilayah dengan status mutu air tanah Cemar Berat dengan total 9 kelurahan, diantaranya yaitu Kota Administrasi Jakarta Utara pada kelurahan Pademangan Timur, Semper Barat, Sukapura, dan Tugu Selatan. Kota Administrasi Jakarta Barat pada kelurahan, Jembatan Lima dan Pinangsia. Kota Administrasi Jakarta Selatan pada kelurahan Cipete Selatan. Kota Administrasi Jakarta Timur pada kelurahan Ujung Menteng. Untuk dominasi status mutu air cemar sedang berada pada Kota Administrasi Jakarta Barat.

V. KESIMPULAN

Hasil evaluasi algoritma naïve bayes dalam klasifikasi kualitas kondisi air tanah di DKI Jakarta berdasarkan dua teknik, yaitu *cross-validation* dan *percentage split*, menghasilkan nilai akurasi yang berbeda. Akurasi tertinggi tercapai dengan menggunakan teknik *percentage split* 80:20, mencapai 84.36%, dengan nilai presisi 0.8566, recall 0.8436, dan F1-score 0.8436.

Selain itu, hasil pemantauan kualitas air tanah di DKI Jakarta pada tahun 2019 dan 2021, dengan dua periode pada setiap tahunnya, menggambarkan tren tertentu. Pada tahun 2019, periode 2 mengalami peningkatan jumlah pencemaran berat, dan tren serupa terlihat pada tahun 2021, periode 1. Namun, pada tahun 2021, periode 2 mengalami peningkatan pencemaran sedang, sementara terjadi penurunan pencemaran berat. Pencemaran ini terjadi akibat terkaitnya dengan lima parameter yang melebihi baku mutu, yaitu pH, deterjen, mangan, E. coli, dan total coliform.

Parameter tersebut dapat menyebabkan masalah kesehatan yang signifikan. Misalnya, ketika tingkat pH air tanah terlalu tinggi atau terlalu rendah, menurut *National Institutes of Health* ini dapat memiliki dampak serius pada tubuh manusia. Tingkat pH yang tinggi disebut "basa" atau "alkalis." Jika tingkat pH darah kita terlalu tinggi, hal ini dapat menyebabkan otot berkedut, mual, kebingungan, koma, dan dampak negatif kesehatan lainnya. Selain itu, kelebihan mangan dalam air tanah juga dapat memiliki dampak kesehatan yang serius.

Gejala keracunan mangan meliputi gemetar, kejang otot, masalah pendengaran, mania, insomnia, depresi, kehilangan nafsu makan, sakit kepala, mudah tersinggung, lemah, dan perubahan suasana hati. Oleh karena itu, penting untuk memantau dan mengendalikan parameter-parameter ini guna memastikan kualitas air tanah yang baik dan mencegah dampak negatif pada kesehatan manusia.

PERAN PENULIS

Arsya Fathiarahma: Konseptual, Persiapan Data, Penyusunan Draft Asli;

Nina Sulistiyowati: Validasi, Penulisan Penyusunan Draft Asli, Penulisan Review & Penyuntingan;

Taufik Ridwan: Penulisan Draft Asli Persiapan, Penulisan Review & Editing;

Apriade Voutama: Visualisasi, Penulisan Review & Penyuntingan;

COPYRIGHT



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

DAFTAR PUSTAKA

- [1] Detik, "Ada Berapa Banyak Jumlah Air di Bumi? Ini Jawabannya," 2022. [Online]. Available: [https://www.detik.com/edu/detikpedia/d-5880228/ada-berapa-banyak-jumlah-air-di-bumi-ini-jawabannya#:~:text=Merangkul laman Lembaga Survei Geologi,juta kilometer kubik \(km3\)](https://www.detik.com/edu/detikpedia/d-5880228/ada-berapa-banyak-jumlah-air-di-bumi-ini-jawabannya#:~:text=Merangkul%20laman%20Lembaga%20Survei%20Geologi,juta%20kilometer%20kubik%20(km3).).
- [2] H. Hamzar, S. Suprpta, and A. Amal, "Analisis Kualitas Air Tanah Dangkal Untuk Keperluan Air Minum Di Kelurahan Bontononpo Kecamatan Bontononpo Kabupaten Gowa," *J. Environ. Sci.*, vol. 3, no. 2, 2021, doi: 10.35580/jes.v3i2.20048.
- [3] A. Misa, R. S. Duka, S. Layuk, and Y. T. Kawatu, "HUBUNGAN KEDALAMAN SUMUR BOR DENGAN KADAR BESI (Fe) DAN MANGAN (Mn) DI KELURAHAN MALENDENG KECAMATAN PAAL 2 KOTA MANADO," *JKL*, vol. 9, no. 5, p. 55, 2019.
- [4] PDAM Tirta benteng, "Amankah Mengonsumsi Air Tanah, Ini Penjelasan Ahli Hidrologi," 2018. <http://www.pdamtirtabenteng.co.id/berita/amankah-mengonsumsi-air-tanah-ini-penjelasan-ahli-hidrologi>
- [5] M. J. Sodiq and E. I. Sela, "Perbandingan Metode Naive Bayes Dan K-Nearest Neighbor Pada Klasifikasi Kualitas Udara Di Dki Jakarta," 2019.
- [6] D. Purwanto, "Klasifikasi Naive Bayes Untuk Memprediksi Kualitas Udara Berdasarkan Indeks Standar Pencemar Udara (Ispu)," pp. 1–94, 2019, [Online]. Available: <https://123dok.com/document/y8rk480q-skripsi-klasifikasi-memprediksi-kualitas-berdasarkan-indeks-standar-pencemar.html>
- [7] J. H. Jaman, J. H. Jaman, . C., and A. R. Sanjaya, "Klasifikasi jenis mobil paling diminati Di Indonesia menggunakan algoritma Naive bayes," *Fakt. Exacta*, vol. 13, no. 1, p. 18, 2020, doi: 10.30998/faktorexacta.v13i1.5573.
- [8] A. Aziiz, H. Kirono, I. Asror, Y. Firdaus, and A. Wibowo, "Klasifikasi Tingkat Kualitas Udara Dki Jakarta Menggunakan Algoritma Naive Bayes," *eProceedings ...*, vol. 9, no. 3, pp. 1962–1969, 2022, [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/18002%0Ahttps://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/18002/17631>
- [9] F. Husna, R. Ramadhan, A. C. Firdaus, and I. Veritawati,

- “Prediksi Pembatalan Reservasi Hotel Menggunakan Algoritma Naive Bayes,” vol. 4, no. 1, pp. 76–80, 2023.
- [10] A. Yoga Pratama *et al.*, “Analisis Sentimen Media Sosial Twitter Dengan Algoritma K-Nearest Neighbor Dan Seleksi Fitur Chi-Square (Kasus Omnibus Law Cipta Kerja),” *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 897–910, 2021.
- [11] L. Muflikhah, D. E. Ratnawati, and R. R. MP, *Data Mining*. 2018. [Online]. Available: https://books.google.co.id/books?id=V_NqDwAAQBAJ&printec=copyright&redir_esc=y#v=onepage&q&f=false
- [12] A. Amrullah, I. Purnamasar, B. N. Sar, Garno, and A. Voutama, “Analisis Cluster Faktor Penunjang Pendidikan Menggunakan Algoritma K-Means (Studi Kasus : Kabupaten Karawang),” *JIRE (Jurnal Inform. Rekayasa Elektron.)*, vol. 5, no. 2, 2022.
- [13] M. R. Nahjan, N. Heryana, and A. Voutama, “Implementasi Rapidminer Dengan Metode Clustering K-Means untuk Analisa Penjualan pada Toko Oj Cell,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 1, 2023.
- [14] H. Nuraliza, O. N. Pratiwi, and F. Hamami, “Analisis Sentimen IMDb Film Review Dataset Menggunakan Support Vector Machine (SVM) dan Seleksi Feature Importance,” *J. Mirai Manaj.*, vol. 7, no. 1, pp. 1–17, 2022.
- [15] M. F. Andriansyah, D. Yusup, and A. Voutama, “Menggunakan Metode Naïve Bayes Berbasis Website Web-Based Expert System of Covid-19 Early Detection Using Naïve Bayes Method,” *J. Inf. Technol. Comput. Sci.*, vol. 4, no. 2, pp. 446–455, 2021.
- [16] N. M. Maghfur, F. Muhammad, and A. Voutama, “Analysis of the Relationship between Public Sentiment on Social Media and Indonesian Covid-19 Dynamics,” vol. 3, no. 3, pp. 336–345, 2021.
- [17] F. V. Sari and A. Wibowo, “Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi,” *J. SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.
- [18] F. Tempola, M. Muhammad, and A. Khairan, “Perbandingan Klasifikasi Antara KNN dan Naive Bayes pada Penentuan Status Gunung Berapi dengan K-Fold Cross Validation,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 5, p. 577, 2018, doi: 10.25126/jtiik.201855983.
- [19] S. Saadah and H. Salsabila, “Prediksi Harga Bitcoin Menggunakan Metode Random Forest,” *J. Politek. Caltex Riau*, vol. 7, no. 1, pp. 24–32, 2021.
- [20] A. Primajaya, B. N. Sari, and A. Khusaeri, “Prediksi Potensi Kebakaran Hutan dengan Algoritma Klasifikasi C4.5 Studi Kasus Provinsi Kalimantan Barat,” *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 2, p. 188, 2020, doi: 10.26418/jp.v6i2.37834.
- [21] Y. Rizki, R. Medikawati Taufiq, H. Mukhtar, and D. Putri, “Klasifikasi Pola Kain Tenun Melayu Menggunakan Faster R-CNN,” *IT J. Res. Dev.*, vol. 5, no. 2, pp. 215–225, 2021, doi: 10.25299/itjrd.2021.vol5(2).5831.

Pengendalian Suhu dan Kelembaban Budidaya Jamur Kuping dengan Kendali PID Penalaan Ziegler-Nichols

Setya Ardhi¹, Tjwanda P. Gunawan², Suhatati Tjandra³

¹Departemen Elektro, Fakultas Sains dan Teknologi, Institut Sains dan Teknologi Terpadu Surabaya, Surabaya, Indonesia

^{2,3}Departemen Informatika, Fakultas Sains dan Teknologi, Institut Sains dan Teknologi Terpadu Surabaya, Surabaya, Indonesia

Corresponding author: Penulis A. Setya Ardhi (e-mail: setyaardhi@stts.edu).

ABSTRACT Cultivating wood ear mushrooms can be an alternative source of food and income. One of the challenges in wood ear mushroom cultivation, particularly in achieving productivity, is the enlargement or incubation process. Mushroom cultivation requires a "kumbung" (mushroom house) as a maintenance facility. Optimal results in wood ear mushroom cultivation require maintaining ideal temperature and humidity within the "kumbung." The implementation of a PID controller using the Ziegler-Nichols tuning method, the first approach, in wood ear mushroom cultivation equipped with a temperature and humidity control system significantly aids the incubation process.

Applying the Ziegler-Nichols rules from the first approach, PID constants are determined after system response characteristics are identified. The prototype wood ear mushroom cultivation hardware design integrates a mist generator, heat generator, and a DHT22 sensor for temperature and humidity readings, with an Arduino Uno serving as the central controller through PID and PWM

Based on data analysis, the best setpoint for humidity control was found to be 85, resulting in an overshoot of 3.1% RH and a steady-state error of 1.7% RH. For temperature control, the optimal setpoint was 26, resulting in an overshoot of 0.7°C and a steady-state error of 0.2°C. Therefore, the total power consumption of the entire system is 31.16 watts at the best setpoint position. All of this was carried out in a prototype wood ear mushroom "kumbung" with a room volume of 7m³, accommodating 48-52 bags.

KEYWORDS Wood Ear Mushroom, PID Controller, Temperature and Humidity, Ziegler-Nichols Tuning.

ABSTRAK Budidaya tumbuhan jamur kuping bisa menjadi sumber bahan pangan alternatif serta menjadi sumber pendapatan. Salah satu yang menjadi kendala dalam budidaya tumbuhan jamur kuping dalam produktifitas budidaya jamur kuping adalah proses pembesaran atau inkubasi. Hasil optimal yang diperoleh dalam budidaya jamur kuping adalah memerlukan suhu dan kelembaban yang ideal dalam kumbung. Penerapan pengendali PID menggunakan kaidah tuning Ziegler-Nichols metode pertama pada budidaya jamur kuping yang dilengkapi dengan sistem pengendali suhu dan kelembaban akan sangat membantu proses inkubasi tersebut.

Dengan menggunakan aturan Ziegler-Nichols dari pendekatan pertama, nilai konstanta PID ditentukan setelah tanaman diidentifikasi berdasarkan karakteristik respons sistem. Desain perangkat keras kumbung jamur kuping prototipe mengintegrasikan generator kabut, generator panas, dan sensor DHT22 untuk membaca suhu dan kelembaban, dengan Arduino Uno yang berfungsi sebagai pengendali pusat melalui penggunaan sketsa program PID dan PWM.

Berdasarkan analisa data yang dilakukan dengan hasil analisa respon pada kendali kelembaban diberikan setpoint terbaik pada nilai 85 sehingga menghasilkan nilai overshoot sebesar 3,1% RH dan Error Steady State sebesar 1,7% RH, sedangkan untuk kendali suhu diberikan setpoint terbaik pada nilai 26 sehingga menghasilkan nilai overshoot sebesar 0,7°C dan error steady state sebesar 0,2°C. Oleh karena itu konsumsi daya total seluruh sistem sebesar 31,16 watt pada posisi setpoint terbaik. Semua itu dilakukan pada prototipe kumbung jamur dengan volume ruangan sebesar 7m³ yang menampung 48-52 baglog.

KATA KUNCI Jamur Kuping, Kontroler PID, Suhu dan Kelembaban, Tuning Ziegler-Nichols.

I. PENDAHULUAN

Budidaya tumbuhan jamur kuping dimana bisa menjadi alternatif bahan pangan. Pada budidaya tersebut harus diperhatikan beberapa syarat dalam tumbuh jamur kuping dimana dalam pertumbuhan jamur kuping terdiri dari penyediaan bangunan untuk mendukung proses hidup jamur, kelembaban ruangan, dan suhu ruangan. Kelembaban sangat berpengaruh terhadap pertumbuhan jamur kuping karena iklim asli di alam tempat tumbuhnya jamur kuping memiliki kelembaban tinggi. Selain kelembaban, suhu juga dibutuhkan dalam pertumbuhan jamur kuping. Suhu optimal untuk pertumbuhan jamur telinga berada dalam kisaran suhu tropis 25°C hingga 30°C. Sangat penting untuk mempertahankan suhu yang konsisten yang tidak terlalu panas atau dingin, karena suhu di atas 35°C atau di bawah 15°C dapat menghambat pertumbuhan jamur. Selain suhu, perkembangan tubuh buah jamur kuping membutuhkan tingkat kelembaban udara yang tinggi, idealnya berkisar antara 80% hingga 90%. Ventilasi yang memadai dan penyiraman yang tepat sangat penting untuk mempertahankan tingkat kelembaban yang diperlukan. Faktor-faktor ini sangat penting untuk pertumbuhan dan perkembangan jamur kuping yang sehat, yang banyak digunakan dalam masakan tradisional Cina dan Jepang karena tekstur dan nilai gizinya yang unik. Jika kelembaban udara terlalu rendah, tubuh buah jamur akan mengalami kekeringan dan mati. Sedangkan suhu yang dibutuhkan dalam pertumbuhan jamur kuping berkisar antara 25°C hingga 28°C. Tubuh reproduksi jamur sebagian besar terdiri dari air, jadi jika suhunya terlalu rendah akan menyebabkan jamur menjadi busuk karena kandungan air yang tinggi.

Pada umumnya budidaya tersebut dengan melakukan secara konvensional seperti pengaturan panas dengan sistem pemanas dan kelembaban dengan bantuan manusia, dimana terjadi kelemahan dalam monitoring secara tepat untuk mengontrol hal tersebut. Maka sistem tersebut dikembangkan dengan sistem elektronik yang memiliki sistem kontrol dengan pengendali Proportional Integral Differential (PID) sangat luas dipakai dalam dunia industri dan peralatan elektronika lainnya, di mana pengontrol ini digunakan untuk menentukan tingkat presisi dalam sistem instrumentasi dengan memberikan umpan balik karakteristik pada sistem. Dengan bantuan pengendali PID ini memiliki kelebihan pada waktu respon tanggapannya sangat cepat namun menghasilkan nilai overshoot yang tinggi.

Sistem kendali PID memakai sistem kendali loop tertutup (closed loop) yang cukup sederhana dan memiliki kemampuan yang bagus. Sistem kendali ini tidak mampu beroperasi dengan optimal apabila terjadi ketidakpastian dan ketidaklinieran pada sistem [1]. Pengendali PID terbentuk dari kombinasi dari tiga jenis pengendali, dimana ketiga jenis pendukung pengendali tersebut apabila berdiri sendiri maka hasil yang dicapai tidak bisa maksimal [2], dimana tiga parameter pengendali tersebut memiliki kelebihan dan kekurangan. Untuk memenuhi sistem kendali yang

dikehendaki harus menetapkan nilai yang optimal pada ke tiga parameter tersebut.

Pengendalian dengan PID memiliki beberapa parameter yang tetap apabila sistem tersebut lagi beroperasi, apabila pada pengendali tersebut mengalami gangguan yang tidak diketahui dari mana asalnya gangguan tersebut atau penempatan tempat yang berubah [3] dari sebelumnya, maka parameter PID yang telah ditetapkan tidak akan bisa beradaptasi [4] dengan perubahan tersebut dan juga penentuan parameter yang telah ditetapkan dalam PID tersebut akan menemui kesulitan. Ada beberapa metode penalaan untuk PID yang telah dikembangkan seperti metode trial and error (Cut and Try Method), metode tanggapan (Step Response), metode analitik, metode Ziegler-Nichols [5]. Dengan adanya beberapa analisa dan percobaan dengan menggunakan berbagai metode penalaan dapat diambil suatu kesimpulan, bahwa dalam penggunaan pengendalian PID dalam sebuah sistem mempunyai kekurangan dimana parameter-parameter tersebut harus di setting/penalaan bila terjadi perubahan didalam sistem tersebut. Pada penelitian ini untuk mengatasi hal tersebut untuk menentukan parameter PID yang diakibatkan perubahan gangguan atau perubahan lingkungan bisa dengan menggunakan metode open loop Ziegler-Nichols [6].

Penelitian ini menerapkan pada pengendali suhu dan kelembaban pada prototipe kumbung (rumah jamur) pada jamur kuping dengan kontrol PID menggunakan kaidah tuning Ziegler Nichols metode pertama. Dimana dilatar belakangi pada kendala yang dapat mengurangi produktivitas budidaya jamur kuping adalah proses pembesaran atau inkubasi. Budidaya jamur kuping membutuhkan kumbung sebagai tempat pemeliharaan [7]. Berdasarkan permasalahan tersebut pada penelitian ini dibuat sebuah sistem yang dapat mengontrol suhu dan kelembaban otomatis pada kumbung dan mendapatkan hasil yang stabil dengan PID. Parameter yang dibutuhkan dalam kontrol PID adalah nilai K_p , K_i , K_d [8] berdasarkan perhitungan plant dari kumbung. Dengan bantuan kaidah Ziegler-Nichols metode pertama, nilai tersebut dapat dengan mudah didapatkan.

Pada penelitian ini prototype yang dibuat mengikuti beberapa parameter yang menjadi syarat untuk budidaya jamur kuping terdiri dari penyediaan bangunan untuk mendukung proses hidup jamur, kelembaban ruangan, dan suhu ruangan. Kelembaban sangat berpengaruh terhadap pertumbuhan jamur kuping karena iklim asli di alam tempat tumbuhnya jamur kuping memiliki kelembaban tinggi pada tingkat kelembaban 80-90% [9] dan diperkuat menurut informasi budidaya jamur kuping pada Dinas Pemberdayaan Masyarakat dan Desa (DPMD) Provinsi Jawa Timur. Selain kelembaban, suhu juga dibutuhkan dalam pertumbuhan jamur kuping. Suhu yang dibutuhkan adalah suhu tropis yang tidak terlalu tinggi dan tidak terlalu rendah. Suhu di Indonesia sangat cocok sebagai tempat pertumbuhan jamur kuping dimana rentang suhu udara normal rata-rata dari

1981-2020 menurut data BMKG 21,7 derajat Celcius – 29,7 derajat Celcius yang diterbitkan oleh Majalah Tempo 11 Januari 2023.

Dalam merangsang pertumbuhan jamur kuping pembentukan tubuh buah jamur kuping dibutuhkan kelembaban udara berkisar antara 80% hingga 90% kelembaban [9] Jika kelembaban udara terlalu rendah, tubuh buah jamur akan mengalami kekeringan dan mati. Sedangkan suhu yang dibutuhkan dalam pertumbuhan jamur kuping berkisar antara 25°C hingga 28°C [9]. Jika suhu terlalu dingin tubuh buah jamur banyak mengandung air yang berdampak pada kebusukan.

Dengan adanya parameter-parameter yang dibutuhkan dalam budidaya ini, maka dibuatlah suatu plant untuk mengontrol parameter-parameter tersebut melalui pembuatan kontrol PID berbasis mikrokontroler untuk dapat mengontrol suhu dan kelembaban pada rancang bangun kumbung aluminium mini dengan kapasitas maksimal 48-52 baglog. Kumbung tersebut dapat menampung sebanyak 96 baglog jika dibuat dua kali lebih besar dari model aslinya. Dua sensor DHT22 memantau suhu dan kelembaban di dalam kumbung yang kecil dan menyediakan data untuk sistem kontrol. Aktuator pemanas dipasang di dalam kumbung selain sensor suhu yang disebut DHT22.

II. TINJAUAN PUSTAKA

Tinjauan literatur yang digunakan dalam penelitian ini mencakup topik-topik seperti dasar-dasar pertumbuhan kelompok jamur kuping, bagaimana kontrol proporsional-integral-derivatif (PID) bekerja dan dijelaskan, bagaimana aturan penyetelan Ziegler-Nichols untuk sistem orde pertama dapat diterapkan pada tanaman yang tidak diketahui model matematikanya, karakteristik kelompok jamur yang menunjukkan Respons Transien yang diinginkan, dan banyak lagi.

A. Budidaya Rumah Jamur Kuping

Jamur kuping, yang secara ilmiah dikenal sebagai *Auricularia auricular*, terbentuk pada kayu yang lembap dan usang. Nama jamur ini, "jamur kuping", berasal dari fakta bahwa tubuh jamur yang tumbuh melebar di bagian tengah, seperti telinga manusia. Seperti jamur tiram dan shiitake, jamur kuping berkembang biak melalui biji yang berukuran mikroskopis tetapi banyak. Hal ini terjadi ketika tubuh jamur yang sudah habis berbuah membusuk. Jika spora mendarat di lingkungan yang cocok, mereka akan berkecambah dan membentuk miselium, melalui serangkaian tahap yang meliputi fase miselium primer, fase miselium sekunder, dan akhirnya pembentukan kuncup tubuh buah yang akan tumbuh menjadi tubuh buah yang terbentuk sempurna.

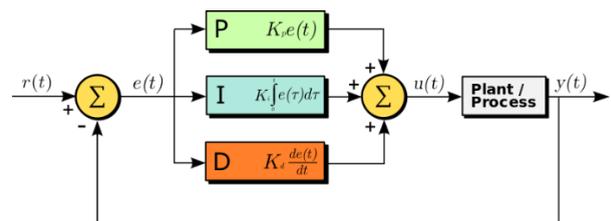
Kumbung memiliki struktur yang digunakan untuk menyimpan bahan tanam yang digunakan dalam budidaya jamur kuping. Baglog jamur akan tetap kering dan aman dari unsur-unsur dan potensi pertumbuhan jamur di dalam

kumbung. Selain itu, area kumbung dapat digunakan untuk rekayasa iklim mikro. Salah satu contoh petani jamur yang membutuhkan ruang kumbung berukuran panjang 6 meter, lebar 4 meter, dan tinggi 3 hingga 6 meter adalah petani yang berencana menanam 1.000 baglog. Ruang kumbung harus berkapasitas tertentu untuk memastikan aliran udara yang tepat serta kelembaban dan suhu yang konsisten.

Rak-rak dari bambu atau kayu disediakan di area kumbung untuk menyimpan baglog jamur. Untuk memastikan media baglog tertata rapi dan bibit yang dibudidayakan dapat tumbuh dengan maksimal, jarak antar deretan rak sekitar 70-90 cm, dan setiap rak dapat berisi 15 baglog yang disusun secara vertikal dan 20 baglog yang disusun secara horizontal.

B. Kontrol PID (Proportional Integral Derivative)

Sistem kontrol PID secara konstan menentukan nilai kesalahan dengan membandingkan variabel proses aktual dengan setpoint yang diinginkan. Berdasarkan pembacaan kesalahan, aktuator akan diperintahkan untuk bergerak seperti yang ditunjukkan pada skema blok kontrol PID Gambar 1 [10].



GAMBAR 1. Blok Diagram Kontrol PID

Kontrol P (Proporsional), kontrol I (Integral), dan kontrol D (Derivatif) adalah tiga jenis regulasi yang digunakan dalam sistem kontrol PID; masing-masing memiliki kelebihan dan kekurangan [11] [12]. Kontrol PID dapat dijabarkan seperti pada (1).

$$u(t) = K_p e(t) + K_I \int_0^t e(\tau) d\tau + K_D \frac{de(t)}{dt} \quad (1)$$

Pada persamaan 1 dengan hasil keterangan sebagai berikut Untuk $u(t)$ = Output PID, K_p = Kontanta Proportional, K_I = Kontanta Integral, K_D = Kontanta Derivative, e = Error, t = Waktu(s), τ = Variabel Integrasi.

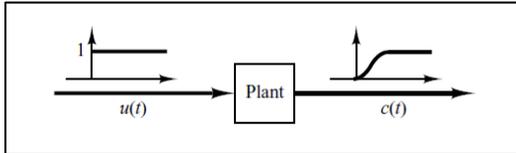
Nilai output yang dicapai oleh kontrol proporsional sebanding dengan nilai kesalahan, seperti yang didefinisikan oleh kontrol [12]. Menyesuaikan reaksi pengontrol semudah mengalikan kesalahan dengan konstanta K_p . Kontroler integral dapat memberikan angka keluaran yang sebanding dengan ukuran dan panjang kesalahan. Integral yang digunakan dalam kontrol PID membangun nilai offset yang sebanding dengan jumlah kesalahan yang telah dikompensasi. Konstanta Integral akan dikalikan dengan

kesalahan kumulatif. (Ki). Kontrol integral dengan cepat menghilangkan kesalahan kondisi-mapan yang terlihat pada kontrol proporsional dan mendorong proses lebih dekat ke setpoint. [12] Namun, kenaikan dihasilkan dari reaksi integral terhadap kesalahan sebelumnya (Overshoot).

Kontrol turunan untuk meningkatkan respons sementara sistem dan untuk memprediksi waktu tinggal dan stabilitas [12]. Dalam menghitung turunan dari kesalahan proses, perlu menemukan kemiringan kesalahan dan meningkatkan nilainya dengan konstanta turunan pada setiap iterasi. Teknologi kontrol ini mampu mengurangi overshoot dan meningkatkan stabilitas dengan meningkatkan nilai Kd. Penggunaan kontrol turunan telah terbukti efektif dalam meningkatkan respons sistem, seperti yang terlihat dalam berbagai aplikasi teknologi seperti kendaraan otonom, sistem energi terbarukan, dan teknologi pengolahan makanan. Karena itu, penggunaan teknologi kontrol turunan dapat membantu meningkatkan kinerja sistem dan memperbaiki kestabilannya dengan cara yang efektif dan efisien.

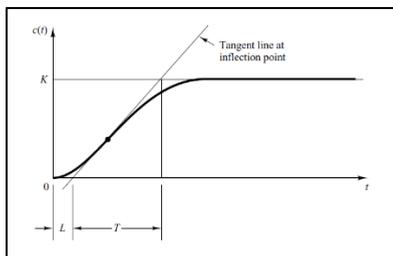
C. Tuning Ziegler-Nichols Metode Pertama

Parameter-parameter penting kontrol PID adalah nilai kontanta Kp, Ki, Kd. Untuk mendapatkan konstanta tersebut adalah salah satunya menggunakan aturan tuning Ziegler-Nichols metode pertama. Dalam metode ini, dilakukan berdasarkan respon plant terhadap respon loop terbuka seperti ditunjukkan pada gambar 2.



GAMBAR 2. Respon Unit-Step Terhadap Plant

Respons unit-step pada plant menghasilkan kurva berbentuk S dan dapat dianalisis dengan menggunakan dua konstanta, L dan T. Konstanta L merupakan jarak antara titik awal kurva dan titik maksimum, sementara konstanta T merupakan waktu yang dibutuhkan oleh sistem untuk mencapai titik maksimum. Konstanta-konstanta ini dapat diperoleh dengan menarik garis singgung yang melintasi sumbu absis dari garis maksimum, seperti yang terlihat pada gambar 3.



GAMBAR 3. Respon Kurva Berbentuk-S

Analisis respons unit-step sangat penting dalam

pengendalian sistem karena dapat memberikan informasi tentang karakteristik sistem, seperti waktu respon, overshoot, dan settling time. Oleh karena itu, dengan memahami konstanta-konstanta pada kurva respons unit-step, teknologi kontrol dapat diatur dengan lebih baik untuk mencapai respons yang diinginkan dan meningkatkan kinerja sistem secara keseluruhan.

Kontanta yang didapatkan dari respon di atas digunakan untuk menentukan parameter-parameter yang dibutuhkan oleh kontrol PID dengan tabel aturan [13]. Berdasarkan aturan tuning Ziegler-Nichols metode pertama didapatkan rumus-rumus seperti di tunjukkan pada tabel 1.

TABEL I
ATURAN TUNING ZIEGLER-NICHOLS METODE PERTAMA

Tipe Controller	Kp	Ti	Td
P	$\frac{T}{L}$	∞	0
PI	$0,9 \frac{T}{L}$	$\frac{L}{0,3}$	0
PID	$0,9 \frac{T}{L}$	2L	0,5L

Fungsi alih C(s)/U(s) dengan pendekatan sistem orde pertama didapatkan (2).

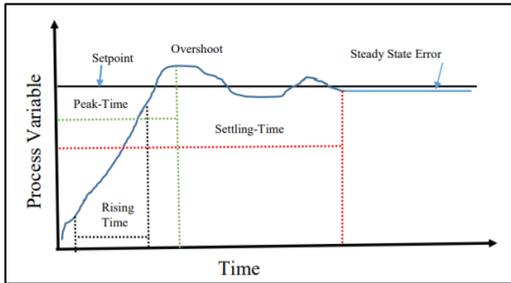
$$\frac{C(s)}{U(s)} = \frac{K e^{-Ls}}{Ts+1} \tag{2}$$

Persamaan (3) menunjukkan hasil dari penerapan rumus Ziegler-Nichols tabel 1 untuk menentukan parameter Kp, Ti, dan Td yang digunakan dalam penyetelan kontrol PID menggunakan pendekatan pertama. Parameter Kp, Ti, dan Td sangat penting dalam penyetelan kontrol PID karena mempengaruhi respons sistem secara keseluruhan. Dalam rumus Ziegler-Nichols tabel 1, parameter Kp, Ti, dan Td dapat dihitung berdasarkan respons sistem terhadap perubahan input [14]. Penggunaan rumus ini dapat membantu pengguna untuk menyetel kontrol PID dengan lebih efektif dan efisien, karena dapat menghasilkan respons sistem yang lebih cepat dan lebih stabil. Oleh karena itu, penyetelan kontrol PID menggunakan rumus Ziegler-Nichols tabel 1 dapat meningkatkan kinerja sistem secara signifikan dan memungkinkan pengguna untuk mencapai respons yang diinginkan.

$$\begin{aligned} Gc(s) &= K_p \left(1 + \frac{1}{T_i s} + T_d s \right) \\ &= 1,2 \frac{T}{L} \left(1 + \frac{1}{2Ls} + 0,5Ls \right) \\ &= 0,6T \frac{(s + \frac{1}{L})^2}{s} \end{aligned} \tag{3}$$

D. Karakteristik Respon Transient Sistem Orde Pertama

Koneksi input-ke-output (I/O), juga dikenal sebagai fungsi transfer, mendefinisikan sistem kontrol orde pertama [15]. Orde sistem dapat dihitung hanya dengan menggunakan model matematisnya, dengan bantuan kekuatan variabel s . (dalam transformasi Laplace). Jika fungsi transfer memiliki variabel berpangkat tunggal, s , maka sistem dikatakan orde pertama.

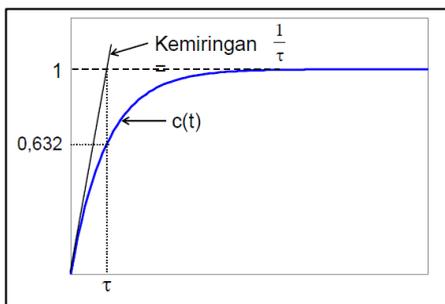


GAMBAR 4. Grafik Karakteristik Respon Transient

Rise Time didefinisikan sebagai waktu gelombang bergerak dari 0,1 ke 0,9 atau 10% sampai 90% dapat juga didapatkan nilai rise time 5% sampai 95% menuju nilai setpoint yang bisa dilihat pada gambar 4. Settling Time didefinisikan sebagai waktu untuk respon untuk mencapai kondisi diam dalam persentase nilai sebesar 2% dari setpoint. Time Constant (τ) merupakan waktu kontan dapat di artikan sebagai waktu yang dibutuhkan untuk respon mencapai 63% dari respon steady-state yang bisa dilihat pada gambar 5.

Waktu Tunda adalah waktu yang telah berlalu dari beberapa waktu awal $t = 0$ hingga respons output mencapai 50% dari nilai kondisi tunaknya sebagai respons terhadap input. Dengan menggunakan (4), dari rumus tsb dapat menghitung periode tunda.

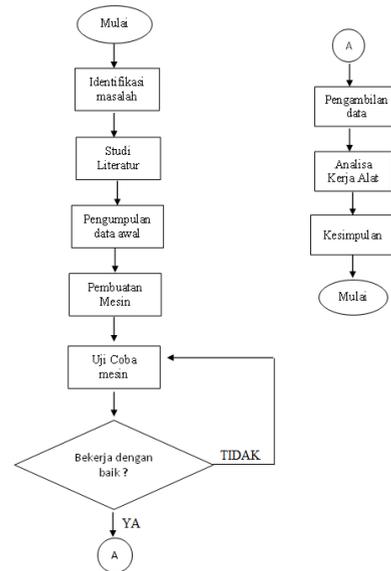
$$td = \tau \ln 2 \tag{4}$$



GAMBAR 5. Grafik Respon Time Constant

III. Metode dan Inti penelitian

Untuk menentukan desain sistem kontrol dan menemukan bagaimana sensor suhu dan kelembaban dapat disesuaikan secara efisien untuk proses pertumbuhan jamur tiram, penelitian ini menggunakan jenis metode analisis eksperimental yang dikenal sebagai studi kasus satu bidikan (penelitian eksperimental One-Shoot Case Study) [16].



GAMBAR 6. Flowchart Metode Perancangan

A. Identifikasi Masalah, Analisa dan Studi Literatur

Analisis kerja alat pada Gambar 6 mengungkapkan apakah alat ini dapat melakukan tugas yang dimaksudkan dengan sukses atau tidak. Data yang dikumpulkan untuk menentukan apakah ada perbedaan yang signifikan secara statistik antara temuan atau apakah temuan tersebut serupa. Berikut ini adalah garis besar untuk menganalisis pekerjaan yang terlibat dalam menggunakan Metode Penyetelan PID dengan aturan penyetelan Ziegler-Nichols pertama untuk secara efektif mengatur suhu dan kelembaban.

- Mengetahui proses perkembangan biakan jamur tiram pada sebuah kumbung dan berapa banyaknya baglog yang bisa ditempatkan pada sebuah kumbung.
- Mengetahui parameter apa saja yang menjadi pendukung tingkat suhu dan kelembaban perkembangan biakan jamur tiram.
- Menerapkan perangkat elektronik untuk mendukung pengendalian tingkat suhu dan kelembaban secara efisien dengan beberapa metode pengendalian.
- Pengujian respon sistem kendali terhadap pengendalian yang diatur yaitu suhu dan kelembaban.
- Pengujian prototype kumbung yang layak untuk budidaya jamur tiram dan sistem kendali yang efisien dalam kendali suhu dan temperature.

B. Pengumpulan data awal

Berikut ini adalah informasi awal yang dikumpulkan dari berbagai sumber untuk tujuan penyempurnaan dan pengembangan:

- Kumbung memiliki baglog yang berasal dari

budidaya jamur tiram dari sebuah UMKM yang berasal dari daerah Batu.

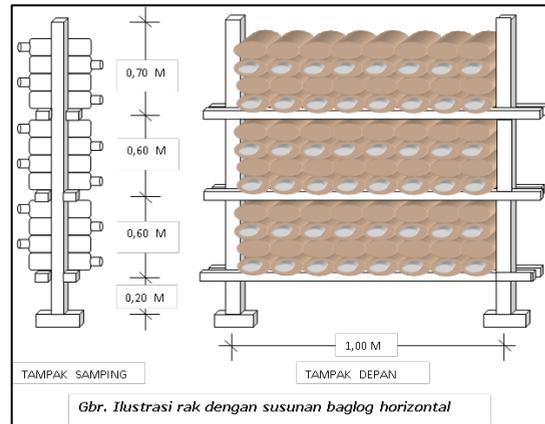
- Metode budidaya dengan cara tradisional dengan menyimpan baglog jamur tiram pada kumbung yang bersifat permanen atau tidak bisa dipindahkan dimana kumbungnya dengan kapasitas 250 baglog dan dimensi volumenya 35m³.
- Pengendalian suhu dan kelembaban dengan pengaturan manual untuk menjaga suhu dengan cara udara yang diblower dengan kipas angin serta penyiraman air secara berkala dan untuk menjaga kelembaban.

C. Perancangan Pembuatan Prototype Alat/Mesin Inti Penelitian

Perangkat keras mekanis, elektronik tambahan, dan diagram alir program, semuanya akan dijelaskan dalam bagian ini. Sejak awal, bentuk kotak atau lemari dimaksudkan untuk portabel dan sementara. Salah satu contoh petani jamur yang membutuhkan ruang kumbung berukuran panjang 6 meter, lebar 4 meter, dan tinggi 3 hingga 6 meter adalah petani yang berencana menanam 1.000 baglog. Selain memungkinkan pergerakan udara yang efektif dan kelembaban serta suhu yang stabil di dalam kumbung dengan bantuan operator, ukuran ruangan yang mirip dengan yang ditunjukkan pada Gambar 7 direkomendasikan.

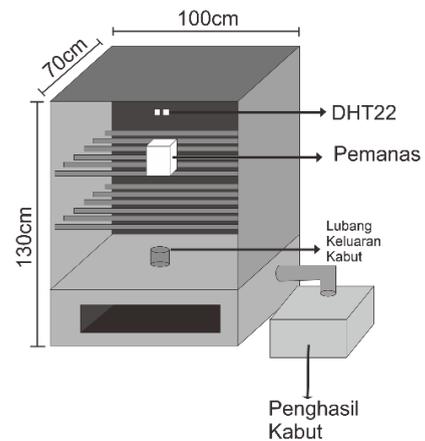
Rak-rak dari bambu atau kayu biasanya disediakan di area kumbung untuk menyimpan baglog jamur dalam budidaya jamur. Penyediaan rak-rak ini bertujuan untuk memastikan bahwa media baglog tertata rapi dan bibit yang dibudidayakan dapat tumbuh dengan maksimal. Jarak antar deretan rak juga penting untuk dipertimbangkan dalam penyusunan agar bibit jamur dapat tumbuh dengan optimal. Oleh karena itu, jarak antar deretan rak dibuat sekitar 70-90 cm, dengan setiap rak dapat menampung hingga 15 baglog yang disusun secara vertikal dan 20 baglog yang disusun secara horizontal. Dalam penyusunan rak, baik rak vertikal maupun horizontal harus dirancang dan disusun dengan baik agar bibit jamur dapat tumbuh dengan baik. Dengan penyusunan yang rapi dan teratur, bibit jamur dapat tumbuh dengan baik dan berkualitas, sehingga menghasilkan hasil panen yang optimal bagi para petani jamur. Selain itu, penyusunan rak yang baik juga membantu para petani dalam manajemen area kumbung budidaya jamur agar lebih efisien dan efektif.

Perancangan Perangkat keras yang dibuat merupakan prototipe kumbung yang didesain pada gambar 8 berdasarkan skala ukuran kumbung konvensional. Adapun bahan dari kumbung merupakan aluminium dengan ukuran dimensi luar 100 x 70 x 130 cm dan dimensi dalam 100 x 70 x 100 cm sehingga volumenya 7.1m³ sebanyak 48-52 baglog dimana lima kali lebih kecil dari sumber budidaya yang menjadi referensi.



GAMBAR 7. Rak Baglog Jamur

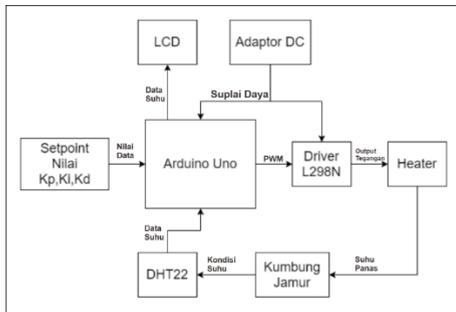
Rancangan dari prototipe sistem pengendali suhu dengan kontrol PID dan cara kerja sistem ini dapat digambarkan seperti diagram blok berikut pada gambar 9.



GAMBAR 8. Rak prototype 1: 5 Baglog Jamur

Pada gambar 9 merupakan diagram blok sistem pengendali suhu dengan kontrol PID. Hal ini dilakukan agar sistem mendapatkan nilai yang stabil dalam mengendalikan suhu yang terdapat pada kumbung. Bagian ini ada beberapa spesifikasi yang bisa dijelaskan dari sensor dan komponen elektronik, seperti Sensor DHT22 (atau AM2302) adalah sensor suhu dan kelembaban yang dikemas dalam modul. Sensor ini menggunakan teknologi NTC untuk mengukur suhu dan sensor kelembaban relatif yang dikalibrasi bersama untuk mengukur kelembaban. Berikut adalah spesifikasi teknis dari sensor DHT22 Rentang pengukuran suhu: -40 hingga 120 derajat Celsius Ketelitian pengukuran suhu: +/- 0.5 derajat Celsius Rentang pengukuran kelembaban: 0 hingga 100% RH. Komponen elektronik Driver L298N adalah sebuah IC (Integrated Circuit) yang berfungsi sebagai driver beban dimana disini sebagai pengatur pemanas (heater). Komponen tersebut memiliki spesifikasi dengan tegangan operasi: 5V hingga 35V, arus keluaran maksimum: 2A per channel atau 4A total, dapat mengendalikan 2 motor

DC atau 1 motor stepper motor, dan terdapat proteksi overheating dan overcurrent. Pemilihan Arduino Uno dimana merupakan board mikrokontroler open-source yang dirancang untuk keperluan prototyping dan DIY (do-it-yourself) projects. Dengan spesifikasi tegangan operasi: 5V, input tegangan: 7-12V, pin input/output digital: 14 (termasuk 6 output PWM), pin input analog sebanyak 6 pin, arus output per pin sebesar 20 mA, memiliki memori flash: 32 KB (termasuk 0.5 KB digunakan bootloader), memiliki SRAM: 2 KB dan EEPROM: 1 KB, dan memiliki Kecepatan clock: 16 MHz.

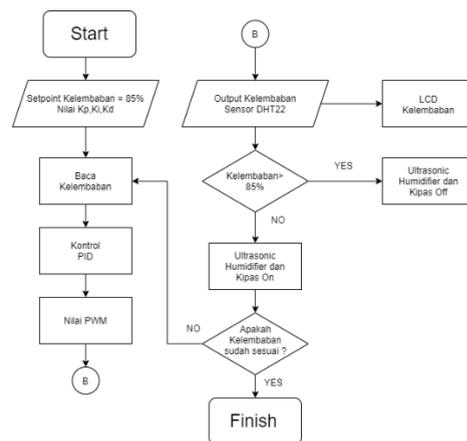


GAMBAR 9. Blok Diagram Sistem Kendali Suhu Dengan Kontrol PID

Pertama diharuskan memasang suatu adaptor untuk menyuplai tegangan pada Arduino dan Driver L298N agar seluruh sistem dapat bekerja dengan baik. Sebelum sistem dijalankan, nilai setpoint dan konstanta K_p , K_i , K_d di tentukan terlebih dahulu. Arduino uno akan diprogram sebagai kontroler PID dimana output yang dihasilkan berupa sinyal PWM yang akan diteruskan ke modul driver L298N. Untuk mengendalikan elemen pemanas/heater element dibutuhkan tegangan yang lebih tinggi. Maka sinyal PWM yang diinputkan dari Arduino ke modul driver menghasilkan persentase tegangan yang mampu menggerakkan kipas pada heater. Pemanas akan menghasilkan panas yang akan diteruskan ke dalam ruang kumbung. Perubahan suhu yang terjadi akan dibaca oleh sensor DHT22. Sensor ini menggunakan prosesor microchip dengan output sinyal digital, meniadakan kebutuhan akan rangkaian pengontrol sinyal dan ADC yang terpisah. Nilai perubahan suhu serta setpoint di ditampilkan oleh LCD dan feedback nilai suhu yang dibaca akan diproses oleh Arduino.

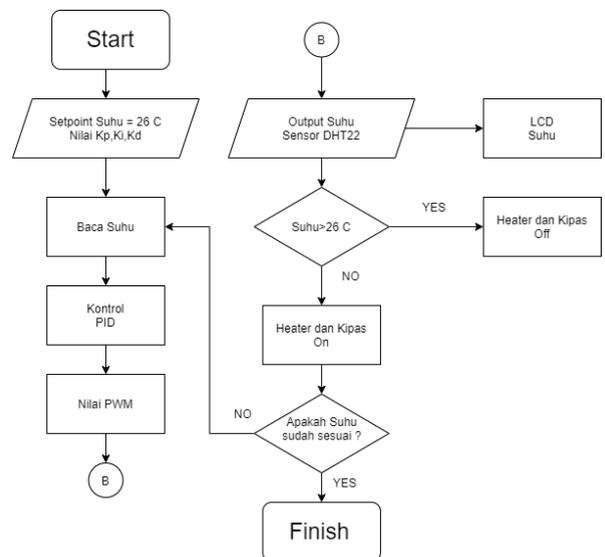
Gambar 10 menunjukkan bagaimana kontrol PID diterapkan dalam perangkat lunak sistem kontrol kelembaban. Tujuan perangkat lunak ini, singkatnya, adalah untuk mendapatkan tingkat kelembapan relatif di kumbung tersebut. Kecepatan kipas semprotan pelembab ultrasonik dihitung berdasarkan kelembapan relatif pembacaan dan angka referensi (setpoint) 85%. Dengan membandingkan persentase RH yang diinginkan dengan persentase RH aktual, sistem PID menyesuaikan tegangan ke kipas angin dan pelembap ultrasonik menggunakan modulasi lebar pulsa (PWM). Kipas angin dan pelembap ultrasonik akan hidup dan mati secara mandiri

berdasarkan tingkat kelembapan.



GAMBAR 10. Flowchart Program Kendali Kelembaban

Gambar 11 memberikan ringkasan tingkat tinggi dari proses mendapatkan suhu ruangan di kumbung tersebut. Perbandingan dibuat antara suhu yang diukur dan suhu target yang telah ditentukan sebelumnya yaitu 26 derajat Celcius. Kemudian, menghitung seberapa cepat kipas harus berputar untuk mengalirkan udara panas dari PTC Heater secara memadai.



GAMBAR 11. Blok Diagram Sistem Kendali Suhu Kontrol PID

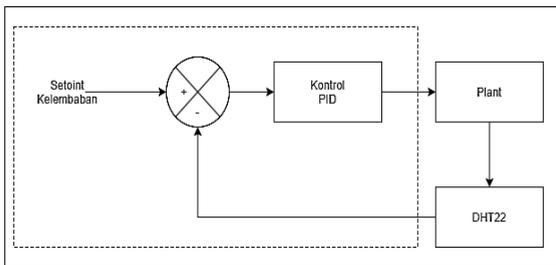
IV. Hasil eksperimen dan penelitian

Pada bab ini akan melnjutkan metode analisis eksperimen studi kasus satu tembakan (experimental research One-Shoot Case Studi) [16] pada gambar 6 dimana dilanjutkan dengan pengambilan data dari hasil permodelan, kalibrasi sensor, pengujian kontrol kelembaban dan suhu dengan kontroler PID, respon sistem terhadap jumlah baglog.

A. Permodelan, Validasi Model Dan Perancangan Kontrol Kelembaban serta Suhu Dengan Kontroler PID

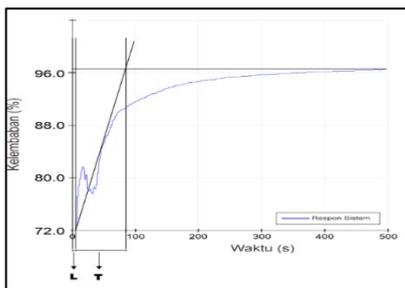
Dengan menggunakan sifat respons sistem aturan pertama Ziegler-Nichols, fungsi transfer dapat dihitung. Model pabrik diperoleh dengan menyiapkan sistem loop terbuka dan memprogram Arduino untuk merespons sinyal langkah dengan sinyal PWM 255, yang memasok daya penuh ke kipas sistem dan kelembaban ultrasonik. Pergeseran kelembaban dapat dilacak dengan DHT22.

Tata letak blok kontrol PID untuk kelembaban di kumbung ditunjukkan pada Gambar 12. Dengan kontrol PID, kelembaban di peternakan jamur dipertahankan dengan mengambil perbedaan antara setpoint dan output sebagai sinyal kesalahan dan kemudian menerapkan prinsip-prinsip proporsionalitas, integrasi, dan diferensiasi padanya.



GAMBAR 12. Diagram Blok Sistem PID Pada Kendali Kelembaban Respon Grafik Sistem Lup Terbuka Kendali Kelembaban

Ziegler-Nichols merupakan salah satu metode penyetelan parameter pada pengontrol PID yang banyak digunakan. Metode ini melibatkan pengukuran respons sistem terhadap input yang diberikan, lalu menetapkan dua konstanta - waktu tunda (L) dan waktu konstan (T) - melalui grafik respons sistem yang digambar menggunakan garis bantu. Dengan mengetahui nilai dari kedua konstanta tersebut, maka dapat ditentukan nilai-nilai parameter untuk pengontrol PID yang optimal. Metode Ziegler-Nichols terbukti efektif dalam penyetelan parameter pengontrol PID pada sistem yang stabil.



GAMBAR 13. Reaksi Plotter Serial Terhadap Pembacaan Kelembaban DHT22

Respons plant loop terbuka Ziegler-Nichols adalah dasar dari teknik pertama. Karena kurangnya Integrator dan kontrol derivatif, plant memiliki kurva reaksi berbentuk huruf S terhadap input tangga. Konstanta waktu T dan waktu

tunda L dihitung dari kurva reaksi plant. Gambar 13 kemudian menampilkan reaksi plotter serial terhadap pembacaan kelembaban DHT22 melalui variabel data DHT22 dalam program Arduino. Menurut aturan tuning metode Ziegler-Nichols pertama dengan menggunakan garis bantu untuk menentukan dua konstanta, yaitu waktu tunda (L) dan konstanta waktu (T), grafik tersebut menunjukkan bahwa model plant memiliki keluaran proses (PV) berupa kelembaban yang meningkat secara konstan hingga mencapai kondisi tunak (steady state), seperti yang ditunjukkan pada (5).

$$G(s) = \frac{96,5e^{-5,5s}}{82,5s+1} \tag{5}$$

Parameter Kp, Ti, dan Td dari kontroler PID yang dituliskan pada Tabel II diperoleh dari hasil data grafik respon sistem plant katup terbuka, mengikuti kaidah tabel parameter Ziegler-Nichols pada pendekatan pertama.

TABEL II
PARAMETER PID PADA KENDALI KELEMBABAN

Kp	Ti	Td
$1.2 \frac{T}{L}$	2L	0.5L
$1.2 \frac{82.5}{5.5}$	2 x 5.5	0.5 x 5.5

Berdasarkan hasil parameter dengan aturan tuning Ziegler-Nichols metode pertama konstanta Kp, Ti, dan Td telah ditemukan. Dari konstanta tersebut dapat juga digunakan untuk menentukan nilai konstanta Ki dan Kd dengan cara pada (6) dan (7).

$$Ki = \frac{Kp}{Ti} = \frac{18}{11} = 1,63 \tag{6}$$

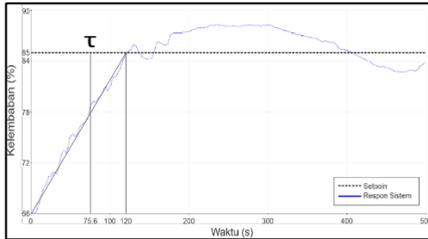
$$Kd = \frac{Kp}{Td} = \frac{18}{2,75} = 6,54 \tag{7}$$

B. Pengujian Kontrol Kelembaban Dengan Kontroler PID

Setelah Anda memiliki nilai Kp, Ki, dan Kd, Anda dapat memasukkan pengontrol PID ke dalam program dan mengirimkannya ke Arduino Uno. Beberapa teknik, termasuk pengujian dengan pengontrol PID, pengujian dengan setpoint variabel, pengujian dengan gangguan, dan pengujian output PWM yang dihasilkan untuk mengontrol kelembapan dan kipas ultrasonik, digunakan untuk mensimulasikan respons sistem dunia nyata.

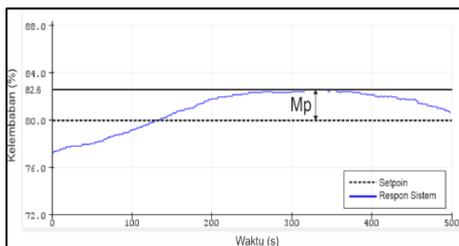
Gambar 14 menunjukkan respons sistem terhadap setpoint kelembaban 85% dengan penggunaan pengontrol PID. Hasil pengujian menunjukkan bahwa pada kondisi tersebut, sistem mampu merespons dengan Ess (Steady State Error) sebesar 1,7% RH dan Overshoot sebesar 3,1% RH. Hal ini menunjukkan bahwa pengontrol PID telah memberikan performa yang baik dalam menjaga kestabilan sistem, dengan nilai Ess yang rendah dan Overshoot yang masih

dalam batas yang dapat diterima. Dalam aplikasi industri, pengontrol PID sering digunakan untuk menjaga stabilitas sistem dengan nilai error dan overshoot yang minim, sehingga hasil pengujian ini menunjukkan bahwa pengontrol PID dapat menjadi pilihan yang tepat untuk aplikasi tersebut.



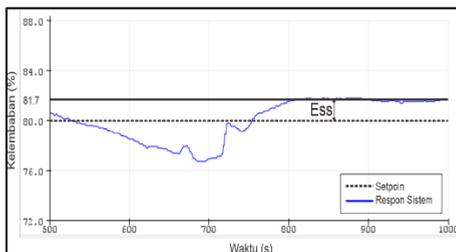
GAMBAR 14. Respon Sistem Kendali Kelembaban Dengan Kontrol PID

Sehingga mendapatkan nilai pengujian Setpoint bervariasi dilakukan dengan memberikan Setpoint 80, 85, dan 90 pada program. Pengujian pada gambar 15 ini bertujuan untuk mengetahui respon setpoint yang lebih stabil diantara setpoint yang diberikan.



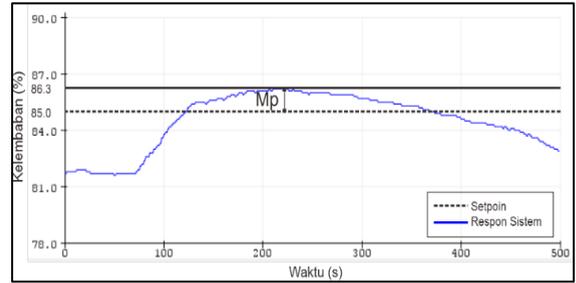
GAMBAR 15. Respon Sistem Dengan Setpoint 80%RH (Error Steady State)

Pada gambar 15 menunjukkan respon grafik dengan menggunakan setpoint 80 terjadi overshoot sebesar 2,6%RH. Keadaan Maksimum Peak atau overshoot terjadi pada 335 detik.



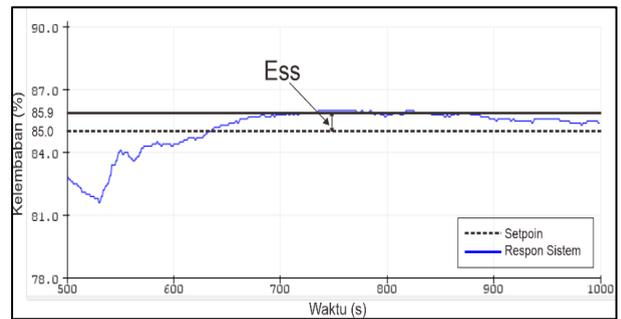
GAMBAR 16. Respon Sistem Dengan Setpoint 80%RH (overshoot)

Pada Gambar 16 merupakan respon grafik kondisi steady state yang menghasilkan error sebesar 1,7%RH. Error Steady State dengan diberikan setpoint 80 terjadi pada 820 detik. Pada Gambar 17 menunjukkan grafik respon sistem ketika diberikan setpoint 85. Overshoot yang dihasilkan sebesar 1,3% RH. Keadaan maksimum peak atau overshoot terjadi pada 220 detik.



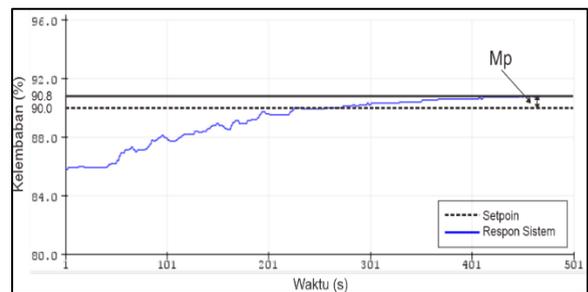
GAMBAR 17. Respon Sistem Dengan Setpoint 85%RH (Overshoot)

Pada Gambar 18 merupakan respon grafik kondisi steady state yang menghasilkan error sebesar 0,9% RH. Error Steady State dengan diberikan setpoint 85 terjadi pada 700 detik.



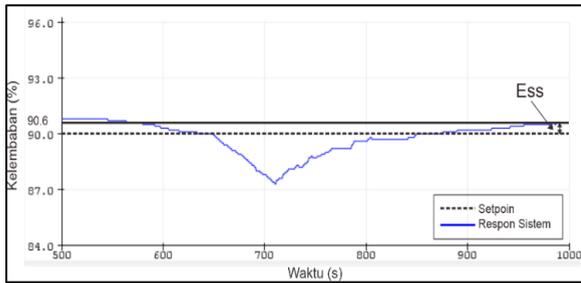
GAMBAR 18. Respon Sistem Dengan Setpoint 85%RH (Error Steady State)

Gambar 19 menunjukkan grafik respon sistem ketika diberikan setpoint 90. Overshoot yang dihasilkan sebesar 0,8% RH. Keadaan maksimum peak atau overshoot terjadi pada 411 detik. Selanjutnya Gambar 20 merupakan respon grafik kondisi steady state yang menghasilkan error sebesar 0,6% RH. Error Steady State dengan diberikan setpoint 85 terjadi pada 993 detik.

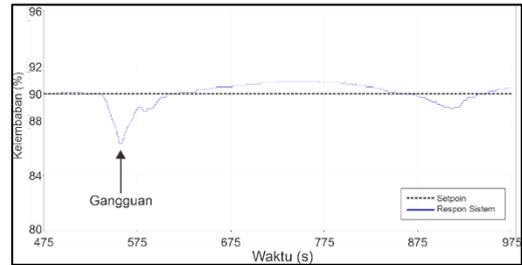


GAMBAR 19. Respon Sistem Dengan Setpoint 90%RH (Overshoot)

Dari grafik hasil respon menggunakan setpoint bervariasi dengan kontrol PID terlihat bahwa setpoint 90% RH lebih stabil dibandingkan menggunakan setpoint yang lainnya. Analisis data hasil respon sistem dengan setpoint bervariasi ditunjukkan pada tabel III berikut.



GAMBAR 20. Respon Sistem Dengan Setpoint 90%RH (Error Steady State)



GAMBAR 21. Respon Sistem Terhadap Gangguan)

Pada tabel IV merupakan hasil percobaan respon sistem dengan perubahan pada setpoint yang diset dari angka 80, 85 dan 90.

TABEL III
HASIL RESPON SISTEM KENDALI KELEMBABAN DENGAN KONTROL PID

No	Keterangan	Rumus	Penyelesaian	Hasil
1	Time delay (t_d)	$\tau \ln 2$	$75,6 \ln 2$	52,4 s
2	Rise time (t_r 5% – 95%)	$\tau \ln 19$	$75,6 \ln 19$	222,59 s
3	Rise time (t_r 10% – 90%)	$\tau \ln 9$	$75,6 \ln 9$	166,11 s
4	Settling time (5%)	3τ	$3(75,6)$	226,8 s
5	Settling time (2%)	4τ	$4(75,6)$	302,4 s
6	Settling time (0,5%)	5τ	$5(75,6)$	378 s

C. Pengujian Respon Menggunakan Kontroler PID dengan Gangguan

Pengujian tahap ini bertujuan untuk menguji ketahanan sistem dengan adanya gangguan. Gangguan di berikan dengan membuka pintu kumbang selama 10 detik. Gambar 21 menunjukkan hasil respon sistem dengan diberikan gangguan.

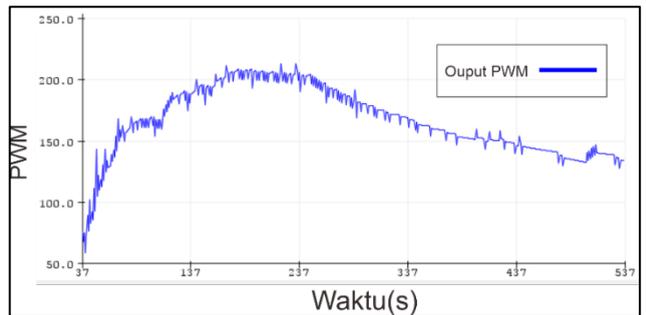
TABEL IV
ANALISIS DATA HASIL PERCOBAAN RESPON SISTEM DENGAN SETPOINT BERVARIASI

XSS (SETPPOINT)	YSS (STEADY-STATE OUTPUT)	Ess (ERROR STEADY-STATE)
80	81,7	1,7
85	85,9	0,9
90	90,6	0,6

Pada saat terjadi gangguan, sistem kontrol PID merespon perubahan kelembaban dengan cepat sehingga perubahan tersebut dapat kembali menuju setpoint dengan baik. Gambar 22 menunjukkan respon output PWM Terhadap kendali kelembaban dengan kontrol PID.

D. Hasil Prototipe Kumbang dan Kalibrasi Sensor

Prototipe penelitian ini, terbuat dari aluminium, yang dimensinya tercantum dalam tabel V. Dua sensor DHT22 mengukur kelembapan dan suhu serta memberikan masukan sistem di dalam kumbang. Sebuah aktuator pemanas dipasang di dalam kandang di samping sensor suhu yang disebut DHT22.



GAMBAR 22. Respon Output PWM Pada Kendali Kelembaban

Dalam rangka mengontrol kecepatan putaran kipas angin di bagian luar gudang, pemasangan kotak plastik dapat menjadi solusi yang efektif. Kotak plastik tersebut akan dilengkapi dengan pengendali utama berupa Arduino uno, serta adaptor dan modul driver L298N. Dengan menggunakan modul driver tersebut, tegangan 12 Volt dari adaptor dapat dikontrol untuk mengatur kecepatan putaran kipas angin. Arduino uno berfungsi sebagai pengontrol utama yang dapat diprogram untuk mengatur kecepatan kipas secara otomatis, berdasarkan input dari sensor suhu atau kelembaban, misalnya. Dengan demikian, solusi ini dapat membantu mengatur suhu atau kelembaban di dalam gudang secara lebih efektif, dengan mengoptimalkan penggunaan kipas angin sesuai dengan kebutuhan. Gambar prototipe kumbang ditunjukkan pada gambar 23. Pada

TABEL V
SPESIFIKASI PROTOTIPE ALAT KUMBUNG

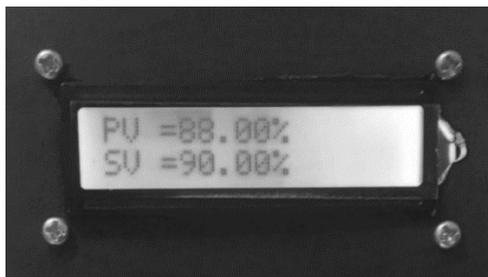
No	DATA ALAT	UKURAN
1	PANJANG	1 METER
2	LEBAR	0.7 METER
3	TINGGI	1,3 METER
4	VOLUME DALAM	7 METER KUBIK
5	JUMLAH BAGLOG	50 BAGLOG

tampilan LCD kendali kelembaban dengan kontrol PID terdapat tampilan bacaan kelembaban (PV) oleh DHT22 dan nilai setpoint (SV) yang merupakan nilai referensi kelembaban yang diinginkan.



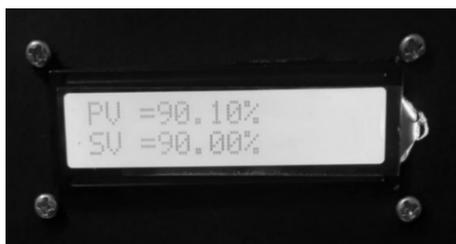
GAMBAR 23. Prototipe Kumbung Jamur

Pada gambar 24 merupakan tampilan kendali kelembaban dengan kontrol PID pada saat pertama kali dinyalakan. Terlihat nilai PV sebelum sistem bekerja terbaca 88%RH sedangkan setpoint (SV) yang diberikan adalah 90%RH.



GAMBAR 24. Tampilan Kelembaban Awal

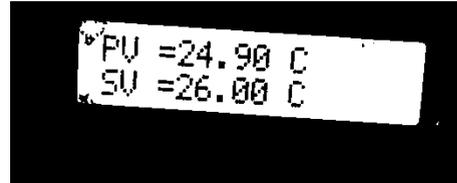
Kondisi kelembaban terhadap setpoint di dalam kumbung pada saat sistem kendali kelembaban pertama kali diaktifkan belum tercapai.



GAMBAR 25. Tampilan Kelembaban Steady State

Pada gambar 25 merupakan tampilan kendali kelembaban dengan kontrol PID pada saat kelembaban telah mencapai keadaan steady state. Terlihat nilai PV kelembaban pada kumbung telah mencapai 90.1%RH. Dengan pemberian setpoint (SV) 90%RH sistem dapat menjaga kestabilan kelembaban.

Pada gambar 26 merupakan tampilan kendali kelembaban dengan kontrol PID pada saat pertama kali dinyalakan. Terlihat nilai PV sebelum sistem bekerja terbaca 24.9°C sedangkan setpoint (SV) yang diberikan adalah 26°C. Kondisi suhu terhadap setpoint di dalam kumbung pada saat sistem kendali suhu pertama kali diaktifkan belum tercapai.



GAMBAR 26. Prototipe Kumbung Jamur

Pada gambar 27 merupakan tampilan kendali suhu dengan kontrol PID pada saat suhu telah mencapai keadaan steady state. Terlihat nilai PV suhu pada kumbung telah mencapai 25.9°C. Dengan pemberian setpoint (SV) 26°C sistem dapat menjaga kestabilan suhu pada kumbung tersebut.



GAMBAR 27. Tampilan Suhu Steady State

Penelitian menunjukkan bahwa budidaya jamur kuping dengan kontrol PID pada titik setel kelembaban dan suhu yang telah ditentukan sebelumnya menghasilkan jamur yang sehat dan berbuah. Meskipun bibit yang diperoleh memiliki usia yang bervariasi, yang menyebabkan perkembangan jamur kuping tidak konsisten. Perangkat suhu dan kelembaban yang dikendalikan PID yang digunakan untuk membudidayakan jamur kuping memberikan hasil yang ditunjukkan pada Tabel VI.

Percobaan respon sistem berdasarkan jumlah baglog di dalam kumbung, bertujuan untuk mengetahui berapa lama sistem mencapai keadaan steady state ketika baglog terisi dengan jumlah tertentu. Percobaan ini dilakukan dengan menjalankan sistem dengan setpoint kelembaban 90 dan suhu 26 ketika kondisi kumbung kosong, terisi 8 baglog, dan terisi 16 baglog. Respon sistem terhadap jumlah baglog dapat ditunjukkan pada tabel VII.

Setelah dilakukan pengujian pada beberapa percobaan sebelumnya ada beberapa yang menjadi kekurangan karena untuk memberi setpoint harus dilakukan manual dimana harus dilakukan melalui debugging lagi pada program Arduino dan dibenamkan ke dalam mikroprosesor tersebut. Adapun ada beberapa yang bisa dikembangkan dalam penelitian selanjutnya dimana bisa ditambahkan aktuator yang digunakan pada pengendalian suhu dapat ditambahkan

dengan peltier pendingin dikarenakan suhu di daerah tropis dapat mengalami perubahan iklim ekstrem pada musim kemarau.

TABEL VI
SPESIFIKASI PROTOTIPE ALAT KUMBUNG

WAKTU PERTUMBUHAN	GAMBAR
MINGGU PERTAMA, BAGLOG DILETAKKAN DIDALAM KUMBUNG TERLIHAT MISELIUM MULAI MEMBENTUK TUBUH BUAH JAMUR.	
MINGGU KEDUA, TUBUH BUAH JAMUR KUPING MEMBESAR DENGAN UKURAN DIAMETER 1CM.	
MINGGU KETIGA TERLIHAT TUBUH BUAH JAMUR KUPING MULAI MEMBENTUK SECARA SEMPURNA.	

Menambahkan keypad agar setpoint dapat diberikan di luar program Arduino. Menerapkan sistem kendali MIMO (Multiple Input-Multiple Output) dimana melibatkan dua variabel dengan satu kontroler. Untuk peneliti lanjut, sebaiknya mengembangkan sistem monitoring dengan menerapkan sistem monitoring secara online atau internet of thing (IoT). Bisa ditambahkan pengendalian suhu dan kelembaban pada kumbung jamur masih bisa dikembangkan dengan mengubah kontroler jenis adaptif, fuzzy logic maupun Jaringan Syaraf Tiruan (JST).

TABEL VII
SPESIFIKASI PROTOTIPE ALAT KUMBUNG

JUMLAH BAGLOG	OUTPUT STEADY STATE		WAKTU STEADY STATE	
	KELEMBABAN (%)	SUHU (°C)	KELEMBABAN (S)	SUHU (S)
0	89,8	26	472	402
8	90,3	26	338	393
16	91,9	26	261	367

V. KESIMPULAN

Pada bagian ini berisi tentang kesimpulan yang diperoleh dari hasil dan pembahasan dari pembuatan prototipe

pengendalian suhu dan kelembaban dengan kontrol PID pada kumbung yaitu:

- Hasil perancangan prototipe kumbung terbuat dari aluminium dengan ukuran 1m x 0,7m x1,3m dengan volume dalam 7m³ dapat menampung maksimal 48-52 baglog.
- Hasil analisa respon kelembaban menggunakan aturan tuning Ziegler-Nichols metode pertama ditemukan Konstanta Kp = 18 mampu mempercepat respon mencapai nilai setpoint, Ki = 1,63 dapat mengurangi overshoot terhadap respon yang dihasilkan, dan Kd = 6,54 mampu memperbaiki respon ketika terjadi perubahan secara mendadak. Dengan diberikan setpoint 85 menghasilkan respon menghasilkan overshoot sebesar 3,1% RH dan Error Steady State sebesar 1,7% RH.
- Hasil analisa respon suhu menggunakan aturan tuning Ziegler-Nichols metode pertama ditemukan Konstanta Kp = 5,09 mampu mempercepat respon mencapai nilai setpoint, Ki = 44,96 mampu mempercepat respon mencapai nilai setpoint, dan Kd = 11,24 mampu memperbaiki respon ketika terjadi perubahan secara mendadak. Dengan diberikan setpoint 26 menghasilkan respon overshoot sebesar 0,7°C dan error steady state sebesar 0,2°C.
- Respon berdasarkan jumlah baglog yang terisi di dalam prototipe kumbung dilakukan dengan memberi setpoint 90 pada kendali kelembaban dan 26 pada kendali suhu. Hasil respon tersebut dengan kondisi kumbung kosong dapat mencapai keadaan steady state pada kelembaban selama 472 detik dan suhu 402 detik. Sedangkan, dengan kondisi kumbung terisi 16 baglog mencapai keadaan steady state pada kelembaban selama 261 detik dan suhu 367 detik.

PERAN PENULIS

- Setya Ardhi:** Penulisan Penyusunan Draf Asli, Penulisan Review & Penyuntingan, Konseptualisasi;
- Tjwanda Putera Gunawan:** Analisis Formal, Investigasi, Administrasi Proyek, Akuisisi Pendanaan, Validasi, Visualisasi;
- Suhatai Tjandra:** Investigasi, Metodologi, Administrasi Proyek, Pengawasan, Validasi, Penulisan Draf Asli Persiapan, Penulisan Review & Editing;

COPYRIGHT



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Daftar Pustaka

[1] Bambang Sampurno, "Sistem Kendali PID pada Pengendalian Suhu untuk Kestabilan Proses Pemanasan Minuman Sari Jagung", *Prosiding Seminar Nasional Instrumentasi, Kontrol dan Otomasi (SNIKO) 2015*, Bandung, Indonesia, 10-11 Desember 2015.

- [2] Nicolaus Allu, Salma Salu, "Aplikasi Penalaan Dengan Metode Ziegler- Nichols Di Perancangan Pengendali PID Pada Putaran Motor DC" *Prosiding Seminar Nasional, Sinergitas Multidisiplin Ilmu Pengetahuan dan Teknologi, Vol 1. (April 2018)*. Hal 203-2017.
- [3] K. Sinthipsomboon, "A Hybrid of Fuzzy and Fuzzy self-tuning PID Controller for Servo Electrohydraulic System," 6th IEEE Conference on Industrial Electronics and Applications, 2011.
- [4] O. Wahyunggoro and N. B. Saad, "Development of Fuzzy-Logic-Based Self Tuning PI Controller for Servomotor," IEEE 10th Intl. Conf. on Control, Automation, Robotics and Vision Hanoi, Vietnam, 17–20 December 2, 2008.
- [5] Era Purwanto, Mukti Wibowo, Soebagio, dkk . "Pengembangan Metoda Self Tuning Parameter PID Controller Dengan Menggunakan Genetic Algorithm Pada Pengaturan Motor Induksi Sebagai Penggerak Mobil Listrik", *Seminar Nasional Aplikasi Teknologi Informasi, Yogyakarta* 20 Juni 2009.
- [6] Zaidir Jamal, "Implementasi Kendali Pid Penalaan Ziegler-Nichols Menggunakan Mikrokontroler" *Jurnal Informatika. Vol 15. No. 1 (Juni 2015)*. Hal 81-88.
- [7] Muchroji, Cahyana, Y.A. "Budidaya Jamur Kuping". PT. Penebar Swadaya. Jakarta, 2004.
- [8] Nicolaus Allu, Salma Salu, "Aplikasi Penalaan Dengan Metode Ziegler- Nichols Di Perancangan Pengendali PID Pada Putaran Motor DC" *Prosiding Seminar Nasional, Sinergitas Multidisiplin Ilmu Pengetahuan dan Teknologi, Vol 1. (April 2018)*. Hal 203-2017.
- [9] Ilona Usuman, Fitriyaningsih. "Penerapan Sistem Integrasi Elektronik dan Pengamatan Perlakuan Sifat Jamur Berdasarkan Suhu dan Kelembaban Pada Ruang Tumbuh Jamur likasi RFID untuk Sistem Kuping (*Auricularia Sp*)" *Jurnal IJEIS (Indonesian Journal of Electronics and Instrumentation Systems), Vol.1, No.2, October 2011*, Hal 11~20. ISSN: 2088-3714.
- [10] Bambang Sampurno, "Sistem Kendali PID pada Pengendalian Suhu untuk Kestabilan Proses Pemanasan Minuman Sari Jagung", *Prosiding Seminar Nasional Instrumentasi, Kontrol dan Otomasi (SNIKO) 2015*, Bandung, Indonesia, 10-11 Desember 2015.
- [11] Iwan Setiawan, "Kontrol PID Untuk Proses Industri" *PT Elex Media Komputindo*, 2008, ISBN: 978-979-27-4100-1.
- [12] Ogata, Katsuhiko, "Modern Control Engineering" fifth edition. New York PrenticeHall.Inc.
- [13] Karsid dan Rofan Aziz, "Perancangan Kendali PID dengan model Proses Reaction Curve pada Showcase Jamur Merang". *Jurnal Neutrino*, Vol 8, No. 1 pp 1-10, niversitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia, 2015.
- [14] Higuitta, Mahendra Ega dan Hendra Cordova, "Perancangan Sistem Pengendalian Suhu Kumbung Jamur dengan Logika Fuzzy", *Jurnal Teknik POMITS*, Vol 2 (2): 2337-3539, 2013.
- [15] Aosong, "Digital-output Relative Humadity & Temperatue Sensor/module DHT22 or AM2302". Aosong Electronic CO., Ltd. 2012.
- [16] Andi Arifudin, Wahyu Dwi K, "Rancang Bangun Sistem Kontrol Mesin Pengisi Dan Penakar Jamu Semi Otomatis Berbasis Timer dan Sensor Ultrasonik," *2021 JRM Journal*, Vol 06 No. 02, pp 18-25 Universitas Negeri Surabaya, 2021.

KLASIFIKASI MICRO-EXPRESSION MENGGUNAKAN K-NEAREST NEIGHBORS MENGGUNAKAN FITUR CAS DAN HOG

Nikko R. P. Wardoyo¹, Joan Santoso¹, dan Esther I. Setiawan¹

¹Magister Teknologi Informasi, Institut Sains dan Teknologi Terpadu Surabaya Jl.Ngagel jaya tengah 73-77, Surabaya 60284, Indonesia

Corresponding author: Nikko Riestian Putra Wardoyo (nikko.r21@mhs.istts.ac.id).

ABSTRACT Micro-Expressions are expressions that appear in a short time, only lasting a fraction of a second. This may be a result of communication activities between humans during social interactions. Facial micro-expression reactions occur naturally and immediately, leaving little room for manipulation. However, because Micro-Expressions are transient and have low intensity, their recognition and recognition is difficult and relies heavily on the experience of experts. Due to their specificity and intrinsic complexity, the classification of Micro-Expressions using 2 extractions namely CAS and HOG is interesting but challenging, and recently- has recently become an active area of research. context-aware saliency (CAS) which aims to detect image regions that represent the scene. The goal is to detect the dominant object. Since this descriptor is extracted on a regular grid, it is powerful for face detection with changes in occlusion, pose, and lighting. The K-Nearest Neighbors (K-NN) method is used for Micro-Expression classification based on HOG features from saliency images. The dataset used in this research was 45 students from Ma'arif NU Prambon Vocational School, majoring in Multimedia, and added a dataset from affecnet. The results obtained from a total dataset of 4116 images which were divided into 6 Micro-Expressions, namely anger, disgust, fear, happy, sad and surprise, obtained accuracy results above 80% from a comparison of a dataset of 4116 divided into 2 with a percentage of 70% training and 30 % testing data.

KEYWORDS - CAS, HOG, K-NN, Micro-Expression.

ABSTRAK *Micro-Expression* adalah ekspresi yang muncul dalam waktu singkat, hanya berlangsung sepersekian detik. Hal ini mungkin merupakan akibat dari aktivitas komunikasi antar manusia selama interaksi sosial. Reaksi ekspresi mikro wajah terjadi secara alami dan segera, sehingga hanya menyisakan sedikit ruang untuk manipulasi. Namun, karena *Micro-Expression* bersifat sementara dan memiliki intensitas rendah, pengenalan dan pengenalannya sulit dan sangat bergantung pada pengalaman para ahli. Karena kekhususan dan kompleksitas intrinsiknya, klasifikasi *Micro-Expression* menggunakan 2 ekstraksi yaitu CAS dan HOG menarik tetapi menantang, dan baru-baru ini menjadi area penelitian yang aktif. *context-aware saliency* (CAS) yang bertujuan untuk mendeteksi wilayah gambar yang mewakili pemandangan. Tujuannya adalah untuk mendeteksi objek dominan. *Histogram Oriented Gradient* (HOG) Bertujuan sebagai deskriptor yang efektif untuk pengenalan dan deteksi objek. Metode *K-Nearest Neighbors* (K-NN) digunakan untuk klasifikasi *Micro-Expression* berdasarkan fitur HOG dari citra *saliency*. Dataset yang digunakan pada penelitian ini dari data sampel siswa SMK Ma'arif NU Prambon jurusan Multimedia sebanyak 45 siswa dan ditambahkan dataset dari affecnet. Hasil yang didapatkan dari total dataset sebanyak 4116 citra yang dibagi menjadi 6 *Micro-Expression* yaitu anger, disgust, fear, happy, sad dan surprise, mendapatkan hasil akurasi diatas 80% dari perbandingan dataset sejumlah 4116 terbagi menjadi 2 dengan persentase 70% training dan 30% data testing.

KATA KUNCI CAS, HOG, K-NN, *Micro-Expression*.

I. PENDAHULUAN

Emosi adalah bagian yang tidak terpisahkan dari kehidupan manusia, dan muncul secara sukarela atau tidak sengaja melalui ekspresi wajah ketika orang berkomunikasi satu sama lain secara tatap muka. Sebagai bentuk komunikasi nonverbal yang khas, ekspresi wajah memainkan peran penting dalam analisis emosi manusia [1]–[5], dan dengan demikian telah dipelajari secara luas di berbagai domain. Emosi juga adalah perasaan atau reaksi terhadap rangsangan yang mendorong orang untuk bertindak [6]. Emosi dapat dipahami dari ekspresi mikro seseorang, dan ekspresi mikro seseorang bersifat universal [7].

Ekspresi wajah adalah cara paling langsung untuk mengekspresikan emosi manusia, dan juga merupakan cara yang sangat penting untuk memahami niat manusia dalam komputasi yang berpusat pada manusia. Ekspresi wajah dibagi menjadi ekspresi makro dan ekspresi mikro (*micro-expression*). Dari perspektif temporal, ekspresi makro berlangsung antara 0,75 detik dan 2 detik, sedangkan ekspresi mikro bertahan antara 0,04 detik dan 0,2 detik [2]. Dalam istilah spasial, gerakan otot wajah ekspresi mikro lebih sedikit dibandingkan dengan ekspresi makro. Selain itu, ekspresi mikro bersifat spontan, artinya seseorang tidak tahu kapan dia menunjukkan ekspresi mikro. Informasi mikro yang penting mudah diabaikan terutama ketika orang menggunakan mata telanjang untuk mengamati. Alat Pelatihan Ekspresi Mikro telah dikembangkan untuk membantu pengenalan. Bahkan untuk profesional dengan pelatihan khusus, hasilnya masih belum ideal. Oleh karena itu, dibutuhkan teknologi komputer untuk membantu *micro-expression recognition* (MER) [3], [4].

Penelitian ini bertujuan untuk mendeteksi emosi saat proses pembelajaran dan untuk membandingkan ekstraksi fitur menggunakan CAS dengan HOG. Hal tersebut bermanfaat untuk membantu para pengajar dalam menentukan durasi dan metode pembelajaran yang tepat bagi siswa SMK.

Sejak 2011, sembilan kumpulan data ekspresi mikro representatif yang sudah dikembangkan: yaitu, USF-HD, kumpulan data Polikovsky, York DDT, MEVIEW, SMIC, CASME, CASME II, SMM dan CAS(ME)2. Perlu dicatat bahwa ekspresi mikro yang berpose ini berbeda dari yang spontan. Selain itu, York DDT terdiri dari ekspresi mikro spontan dengan validitas ekologis yang tinggi. Namun demikian, mirip dengan deteksi kebohongan, data di York DDT dicampur dengan gerakan wajah non-emosional yang dihasilkan dari berbicara. Selain itu, semua kumpulan data (USF-HD, Polikovsky, York DDT, MEVIEW, SMIC, CASME, CASME II, SMM dan CAS(ME)2) tidak tersedia untuk umum [8].

II. PENELITIAN SEBELUMNYA

Terdapat banyak metode yang dapat digunakan dalam pengklasifikasian *micro-expression* diantaranya yaitu *Convolution Neural Network*, K-NN, *Local Binary pattern*,

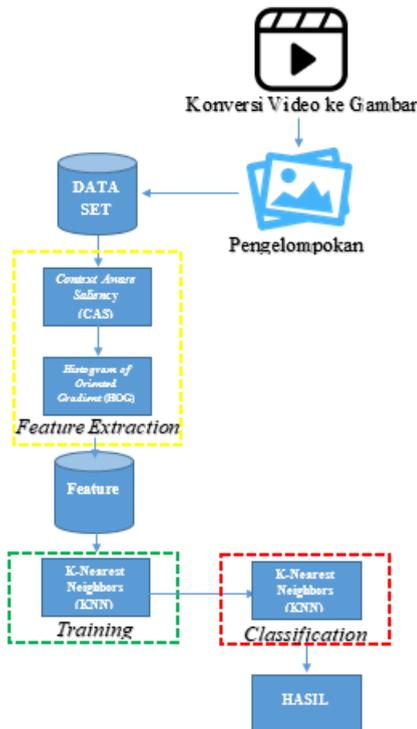
dan masih banyak lagi. Penelitian dengan memanfaatkan metode *Convolution Neural Network* pernah dilakukan oleh Sze-Teng Liong dkk pada tahun 2019 dalam jurnalnya yang berjudul “*Shallow Triple Stream Three-dimensional CNN (STSTNet) for Micro-expression Recognition*” [9]. Pada penelitian tersebut dataset yang digunakan adalah SMIC, CASME II, SMM. Makalah ini menyajikan CNN tiga dimensi dangkal (STSTNet) baru untuk mempelajari fitur panduan aliran optik untuk pengenalan ME. Representasi fitur yang ringkas dan diskriminatif dipelajari dari blok masukan yang terdiri dari tiga gambar aliran optik (yaitu aliran optik horizontal, aliran optik vertikal, dan lengkungan optik). Secara keseluruhan, pendekatan STSTNet yang diusulkan menunjukkan hasil identifikasi yang menjanjikan pada gabungan database ME sintetik yang baru, termasuk tiga database ME spontan, menghasilkan *Unweighted F1-score* sebesar 0,7353 dan *Unweighted Average Recall* sebesar 0,7605, mengungguli metode terancang saat ini [9].

Sedangkan penelitian menggunakan metode *K-Nearest Neighbor* (K-NN) Network pernah dilakukan oleh Ayu Wirdiani dkk pada tahun 2019 dalam jurnalnya yang berjudul “*Face Identification Based on K-Nearest Neighbor*” [10]. Pada penelitian tersebut dataset yang digunakan berasal dari pengambilan gambar manual. Penelitian dataset ini berisi 790 wajah dari 158 orang yang diambil dari beberapa sudut. Dataset tersebut belum melalui proses segmentasi sehingga masih terdapat background yang cukup signifikan di luar objek wajah. Dataset telah dipisahkan ke dalam setiap folder yang berisi nama pemilik individu wajah. Pada tahap pengujian, gambar yang diinput juga akan mengalami proses seperti pada tahap pelatihan, namun gambar yang digunakan merupakan gambar yang belum pernah diuji untuk pelatihan untuk mengukur keakuratan persamaan antara gambar latih dan gambar uji. Berdasarkan hasil percobaan dengan perubahan parameter diperoleh hasil yang berbeda-beda untuk setiap parameter. Pengujian k=1 memperoleh hasil akurasi sebesar 81%, pengujian k=2 memperoleh hasil akurasi sebesar 53%, pengujian k=3 memperoleh hasil akurasi sebesar 47%. Dari hasil uji coba dapat disimpulkan bahwa nilai k sangat mempengaruhi tingkat akurasi sistem. Nilai parameter k dan akurasi berbanding terbalik yaitu semakin besar nilai k maka semakin kecil keakuratan sistem identifikasi [10].

III. PERANCANGAN SISTEM

Alur sistem dijelaskan pada gambar 1, penulis menggunakan dataset siswa SMK Ma'arif Prambon sebanyak 45 siswa dan dengan dataset affecnet. Dataset dibagi menjadi data uji dan data latih. Data latih dan uji menggunakan 4116 citra pada setiap jenis *micro-expression*, Untuk data latih menggunakan 70% sedangkan data uji menggunakan 30% citra pada keseluruhan total dataset pada setiap jenis *micro expression* dengan rincian *anger* sejumlah 1065, *disgust* sejumlah 253, *fear* sejumlah 180, *happy* sejumlah 1220, *sad* sejumlah 895 dan *surprise* sejumlah 503.

Data latih kemudian dilakukan ekstraksi fitur menggunakan CAS dan setelah itu disimpan. Setelah diekstraksi menggunakan fitur CAS tersebut kemudian diekstraksi lagi menggunakan HOG. Fitur akhir hasil ekstraksi disimpan pada database. Proses selanjutnya adalah memisahkan dataset menjadi data latih dan data uji. Metode klasifikasi yang digunakan yaitu K-NN. K-NN digunakan untuk menghitung jarak antara fitur data uji dengan fitur data latih. K-NN digunakan untuk menghitung jarak antara fitur data uji dan fitur data latih. K-NN menghasilkan urutan dari jarak terkecil hingga terbesar. Jarak terkecil menunjukkan kesamaan antara data uji dan data latih.



Gambar 1. Gambaran umum sistem

A. PERANCANGAN PROSES

Proses pertama dilakukan perekaman wajah menggunakan Webcam Logitech 720P selama siswa mengikuti keseluruhan tampilan monitor dengan media visual. Siswa akan diwawancarai di studio podcast SMK Ma'arif NU Prambon dengan pendekatan *constructionist* dan *romantic* maksimal 20 menit dan ada juga yang diarahkan untuk menonton video pendek yang dapat menunjukkan ekspresi tersembunyi dari masing-masing siswa. Wawancara siswa akan direkam melalui webcam menggunakan aplikasi OBS. Hasil dari perekaman dapat dilihat pada gambar 2, dari gambar tersebut dapat dilihat bahwa proses perekaman video minimum resolusi 720 dengan 30 fps dan proses perekaman separuh badan.

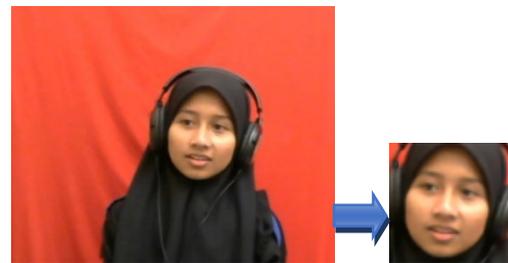
Proses pada poin (1) kemudian akan dilakukan pemotongan per frame untuk mendapatkan input gambar. Pada gambar 3 adalah proses pemotongan yang dilakukan seleksi gambar secara otomatis dan manual. Input gambar

terdiri dari area wajah untuk kebutuhan pengenalan ekspresi wajah dan setengah badan yang terekam kamera untuk keperluan pengenalan ekspresi mikro



Gambar 2. Proses pengambilan Video

Setelah melalui konversi ke gambar, maka dilakukan proses pelabelan oleh guru untuk menentukan klasifikasi *micro expression*. Klasifikasi *micro-expression* terdiri dari 6 kelas yaitu bahagia, sedih, marah, takut, jijik, dan terkejut.



Gambar 3. Proses Cropping wajah

B. FEATURE EXTRACTION

B.1. CONTEXT AWARE SALIENCY

Pada proses feature extraction akan diterapkan metode *Context Aware Saliency* (CAS). CAS mengikuti empat prinsip dasar penglihatan visual manusia, yang didukung oleh bukti psikologis [11]:

1. Pertimbangan lokal tingkat rendah, termasuk faktor seperti kontras dan warna.
2. Pertimbangan global, yang menekan ciri-ciri yang sering muncul, dengan tetap mempertahankan ciri-ciri yang menyimpang dari norma.
3. Rule organisasi visual, yang menyatakan bahwa bentuk visual dapat memiliki satu atau beberapa pusat gravitasi di mana bentuk tersebut diatur.
4. Faktor tingkat tinggi, seperti wajah manusia [11].

B.1.1 LOCAL-GLOBAL SINGLE-SCALE SALIENCY

Ada dua tantangan untuk mendefinisikan *Saliency*. Yang pertama adalah mendefinisikan keunikan baik secara lokal maupun global. Opsi kedua adalah memasukkan informasi lokasi titik koordinat.

Menurut prinsip 1 dan 2, suatu piksel dianggap menonjol jika tampilannya unik. Namun, daripada berfokus pada piksel yang terisolasi, kita perlu fokus pada patch di sekitarnya yang memberikan konteks langsung.

Secara khusus, misalkan $d_{color}(p_i, p_j)$ adalah *Euclidean distance* antara patch tervektorisasi p_i dan p_j dalam ruang warna CIE L^*a^*b yang dinormalisasi ke

rentang [0;1]. Piksel i dianggap menonjol jika tinggi $d_{color}(p_i, p_j)$ tingginya $\forall j$.

Misalkan $d_{position}(p_i, p_j)$ adalah *Euclidean distance* antara posisi patch p_i dan p_j , yang dinormalisasi dengan dimensi gambar yang lebih besar. Berdasarkan observasi tersebut dapat didefinisikan ukuran ketidaksamaan antara sepasang patch sebagai:

$$d(p_i, p_j) = \frac{d_{color}(p_i, p_j)}{1+c.d_{position}(p_i, p_j)} \quad (1)$$

Dalam praktiknya, untuk mengevaluasi keunikan suatu *patch*, tidak perlu menggabungkan perbedaannya dengan semua *patch* gambar lainnya. Cukup untuk mempertimbangkan K *patch* yang paling mirip (jika *patch* yang paling mirip sangat berbeda dari p_i , maka jelas semua *patch* gambar sangat berbeda dari p_i). Oleh karena itu, untuk setiap *patch* p_i , dapat mencari K *patch* yang paling mirip $\{q_k\}_{k=1}^K$ dalam gambar, sesuai dengan (1). Sebuah piksel i menonjol ketika $d(p_i, q_k)$ tingginya $\forall k \in [1, k]$. Nilai *single-scale saliency* pada piksel i dalam skala r didefinisikan sebagai ($K = 64$):

$$S_i^r = 1 - \exp\left\{-\frac{1}{K} \sum_{k=1}^K d(p_i^r, q_k^r)\right\}, \quad (2)$$

B.1.2 MULTI-SCALE SALIENCY ENHANCEMENT

Piksel latar belakang (*patches*) cenderung memiliki *patches* serupa pada berbagai skala, misalnya, di wilayah sebagian besar yang homogen atau buram. Ini berbeda dengan piksel yang lebih menonjol yang dapat memiliki *patches* serupa pada beberapa skala tetapi tidak semuanya. Oleh karena itu, untuk menggabungkan beberapa skala agar semakin mengurangi arti-penting dari piksel latar belakang/*background*, meningkatkan kontras untuk wilayah yang lebih menonjol dan tidak menonjol.

Untuk *patch* p_i dengan skala r , dianggap sebagai kandidat tetangga semua *patches* pada gambar yang skalanya adalah $R_q = \{r, 1/2 r, 1/4 r\}$. Di antara semua *patch* ini, K *patch* yang paling mirip menurut (1) digunakan untuk menghitung *saliency*. Oleh karena itu, (2) dapat ditulis ulang menjadi [11]–[13]:

$$S_i^r = 1 - \exp\left\{-\frac{1}{K} \sum_{k=1}^K d(p_i^r, q_k^r)\right\}, \quad (3)$$

Selain itu, mewakili setiap piksel dengan kumpulan tambalan gambar multiskala yang berpusat padanya. Misalkan $R = \{r_1, \dots, r_M\}$ menunjukkan kumpulan ukuran *patches* yang harus dipertimbangkan untuk piksel i . Arti-penting pada piksel i diambil sebagai rata-rata arti-pentingnya pada skala yang berbeda [11]–[13]:

$$\bar{S}_i = \frac{1}{M} \sum_{r \in R} S_i^r, \quad (4)$$

B.1.3 INCLUDING THE IMMEDIATE CONTEXT

Menurut hukum Gestalt, bentuk visual dapat memiliki satu atau beberapa pusat gravitasi di mana bentuk diatur (prinsip (3)). Hal ini menunjukkan bahwa area yang dekat dengan fokus perhatian harus lebih banyak dieksplorasi secara signifikan daripada area yang jauh. Ketika wilayah di sekitar fokus menyampaikan konteksnya. Untuk mensimulasikan efek kontekstual visual ini terdapat dalam dua langkah. Pertama, area terlokalisasi yang paling banyak dikunjungi diekstraksi dari peta arti-penting yang dihasilkan oleh (4) [11]. Sebuah piksel dianggap ada jika nilai arti-pentingnya melebihi ambang tertentu ($S_i > 0.8$ dalam contoh yang ditunjukkan dalam paper ini). Kemudian, setiap piksel di luar area yang ada diberi bobot sesuai dengan jarak Euclidean-nya ke piksel terdekat. misalkan $d_{foci}(i)$ adalah posisi *Euclidean distance* antara piksel i dan piksel fokus perhatian terdekat, dinormalisasi ke kisaran [0;1]. *Saliency piksel* didefinisikan ulang sebagai [11]–[13]:

$$\hat{S}_i = \bar{S}_i \left(1 - d_{foci}(i)\right). \quad (5)$$

Perhatikan, bahwa *Saliency* dari daerah yang tidak menarik, seperti daerah yang kabur atau homogen, tetap rendah, karena (4) akan mendominasi. Namun, *Saliency* dari latar belakang yang menarik di sekitar objek yang menonjol akan ditingkatkan dengan (5) [12].

Hasil dari ekstraksi menggunakan CAS dapat dilihat pada gambar 4, pada gambar 4 ekstraksi menggunakan software python versi 3.9.13.

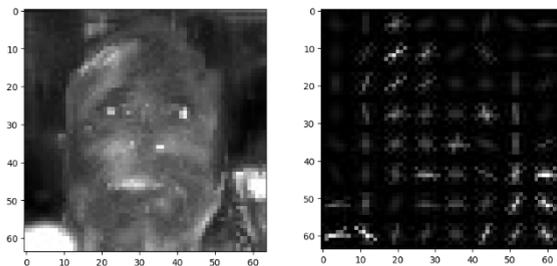


Gambar 4. Membandingkan hasil Saliency pada gambar satu objek dengan latar belakang yang tidak menarik

B.2. HISTOGRAM ORIENTED GRADIENT (HOG)

Sebagai deskriptor bentuk lokal berbasis tepi yang sukses, HOG pada awalnya digunakan untuk memecahkan masalah deteksi objek dengan mendeskripsikan informasi tepi lokal objek. Ide yang sama dari HOG dimanfaatkan di sini untuk menemukan arah gradien yang lebih dominan untuk setiap blok di citra [14]. Nilai besaran gradien untuk setiap piksel diakumulasikan tergantung pada sudut yang sesuai. Kemudian, untuk menemukan sudut dominan θ_d untuk

setiap blok pada gambar, kemudian memilih sudut gradien yang sesuai dengan nilai magnitudo gradien akumulasi maksimum G_d [15]. Berdasarkan kriteria tersebut, langkah-langkah algoritma yang digunakan untuk menemukan arah tepi yang dominan dirangkum dalam Algoritma 1. Contoh ilustrasi gambar grayscale yang berisi arah tepi yang berbeda untuk setiap BOI (*Block Of Interest*) [14]. Terdapat Tiga arah tepi, horizontal, vertikal dan diagonal, hampir ada dan tersebar sepanjang semua piksel gambar [14]–[17].



Gambar 5. contoh BOI untuk gambar *Micro-Expression Fear* yang berisi arah tepi yang berbeda

Dari gambar 5 merupakan hasil dari ekstraksi HOG dengan input gambar hasil dari ekstraksi CAS yang dilakukan menggunakan platform google colab. Proses ini memerlukan waktu sekitar 45-60 menit untuk semua gambar.

C. TAHAPAN K-NN

KNN adalah suatu algoritma sederhana yang menyimpan semua kondisi dan mengklasifikasikan ke dalam kondisi baru berdasarkan kemiripan nilai. Algoritma KNN telah digunakan sejak 1970 pada banyak aplikasi estimasi statistik dan pengenalan pola [18]. Metode ini dipilih dikarenakan dari penelitian sebelumnya mendapatkan hasil akurasi rata-rata maksimum 97% dalam jurnal “*Facial Expression Classification using KNN and Decision Tree Classifiers*” [19] dan 81% pada jurnal “*Face Identification Based on K-Nearest Neighbor*” [10]. Pada metode ini terdapat (6) adalah perhitungan jarak euclidian dimana K adalah jumlah fitur data, x dan y adalah data uji dan data latih. Algoritma K-NN pada dasarnya dilakukan dengan langkah berikut. Dimisalkan klasifikasi label ada dua yaitu label *Anger* dan label *Happy*.

Jika nilai $K=3$ dan data testing lebih dekat ke label *Anger* dengan dua tetangga dan satunya lebih dekat ke label *Happy* maka hasilnya akan diputuskan menjadi label *Anger*. Penentuan label diambil dari mayoritas terhadap K [10], [18]–[25].

$$\sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (6)$$

IV. HASIL DAN PEMBAHASAN

A. IMPLEMENTASI PADA PSEUDOCODE

Proses implementasi Script program pada python memiliki beberapa urutan yaitu proses ekstraksi fitur

menggunakan CAS. Hasil gambar dari CAS harus disimpan terlebih dahulu ke dalam folder dikarenakan proses ini memakan waktu 15 menit tiap gambar. Proses CAS selesai dalam 7 hari menggunakan 20 Notebook. Setelah itu hasil gambar yang sudah di ekstraksi menggunakan CAS dilakukan ekstraksi lagi menggunakan HOG. Hasil ekstraksi disimpan dan dijadikan .zip, kemudian dilakukan upload ke google drive. Setelah itu proses klasifikasi menggunakan K-NN, proses ini memisahkan 70% data training dan 30% data testing secara random. Berikut pseudocode Algoritma K-NN [22]:

Input:

X: data training,

Y: label kelas dari X,

K: jumlah tetangga terdekat.

Output:

Kelas sampel uji x.

Start

Tentukan Nilai K

Klasifikasikan (X, Y, x)

1. for setiap sampel x do

 Hitung jarak: $d(x, X) = \sqrt{\sum_{i=1}^n (x_i - X_i)^2}$

end for

2. Temukan tetangga terdekat

3. Tentukan kelas mayoritas

4. Klasifikasikan titik data baru x pada kelas mayoritas:

$C(x) = \text{argmax}_j \sum_{x_j \in \text{KNN}} C(X_j, Y_k)$

End.

Berikut langkah-langkah utama cara kerja algoritma KNN:

- Langkah pertama adalah memilih nilai K yang menentukan jumlah tetangga yang akan diperiksa untuk menentukan klasifikasi suatu bagian data. Menunjukkan Misalnya, jika K sama dengan 3, algoritma akan mempertimbangkan tiga tetangga terdekat.
- Hitung jarak Selanjutnya algoritma menghitung jarak antara titik data baru yang akan diklasifikasikan dengan seluruh titik data yang ada di set pelatihan. Untuk menghitung jarak ini menggunakan *Euclidean distance*.
- Mencari tetangga terdekat Setelah menghitung jarak, KNN akan menentukan K tetangga terdekat yang memiliki jarak terpendek ke titik data baru.
- Identifikasi kelas mayoritas Dari tetangga yang teridentifikasi, KNN menghitung jumlah titik data pada setiap kelas atau kategori. Kelas yang memiliki jumlah tetangga terbanyak akan dianggap sebagai klasifikasi untuk titik data baru tersebut.
- Klasifikasikan titik data baru Akhirnya, KNN akan menetapkan titik data baru ke dalam kelas yang telah ditentukan sebagai mayoritas oleh tetangga-tetangga terdekatnya. Dengan demikian, titik data baru akan mendapatkan label atau kategori yang sesuai.

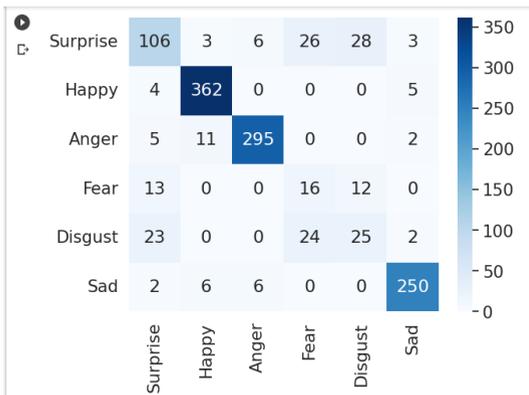
B. EVALUASI HASIL

Hasil dari evaluasi menggunakan 4 indikator utama yaitu akurasi, *precision*, *recall*, dan *f1-Score*. Hasil dari ketiga

perbandingan K=1, K=3 dan K=5 disajikan pada Tabel 1, 2 dan 3. Pada Gambar 6, Gambar 8 dan Gambar 10 merupakan *Confusion Matrix* dari masing-masing pengujian K-NN dan Ekstraksi RGB. Hasil *scatter plot* menunjukkan bahwa perubahan nilai K terhadap nilai *precision*, *recall* dan *precision* semakin meningkat. Adapun observasi yang dihasilkan adalah percobaan dua nilainya selalu lebih tinggi dibandingkan percobaan pertama.

TABEL I
NILAI EVALUASI DENGAN K=1
(PERSENTASE DARI GAMBAR 7)

	Precision	Recall	F1-Score
Surprise	69%	62%	65%
Happy	95%	98%	96%
Anger	96%	94%	95%
Fear	24%	39%	30%
Disgust	38%	34%	36%
Sad	95%	95%	95%
Akurasi Rata-rata	85%		



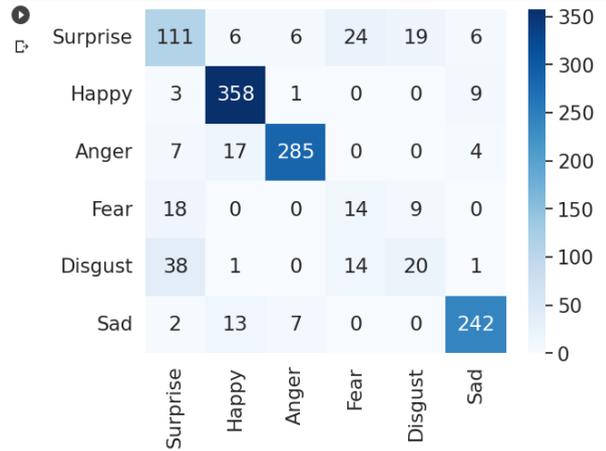
Gambar 6. *Confusion Matrix* K=1 dari hasil pengujian menggunakan google colab

	precision	recall	f1-score	support
0	0.69	0.62	0.65	172
1	0.95	0.98	0.96	371
2	0.96	0.94	0.95	313
3	0.24	0.39	0.30	41
4	0.38	0.34	0.36	74
5	0.95	0.95	0.95	264
accuracy			0.85	1235
macro avg	0.70	0.70	0.70	1235
weighted avg	0.86	0.85	0.86	1235

Gambar 7. Implementasi K-NN dengan K=1 dari hasil pengujian menggunakan google colab

TABEL II
NILAI EVALUASI DENGAN K=3
(PERSENTASE DARI GAMBAR 9)

	Precision	Recall	F1-Score
Surprise	62%	65%	63%
Happy	91%	96%	93%
Anger	95%	91%	93%
Fear	27%	34%	30%
Disgust	42%	27%	33%
Sad	92%	92%	92%
Akurasi Rata-rata	83%		



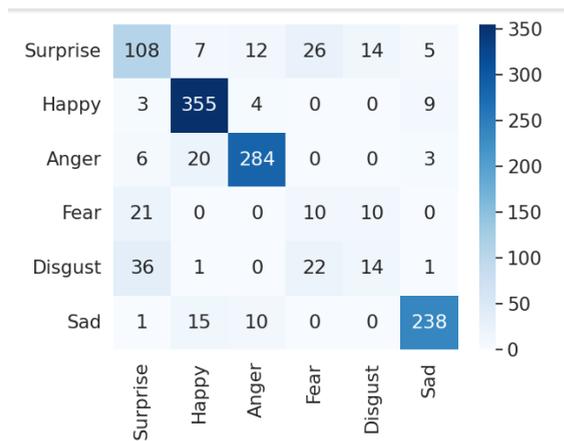
Gambar 8. *Confusion Matrix* K=3 dari hasil pengujian menggunakan google colab

	precision	recall	f1-score	support
0	0.62	0.65	0.63	172
1	0.91	0.96	0.93	371
2	0.95	0.91	0.93	313
3	0.27	0.34	0.30	41
4	0.42	0.27	0.33	74
5	0.92	0.92	0.92	264
accuracy			0.83	1235
macro avg	0.68	0.67	0.67	1235
weighted avg	0.83	0.83	0.83	1235

Gambar 9. Implementasi K-NN dengan K=3 dari hasil pengujian menggunakan google colab

TABEL III
NILAI EVALUASI DENGAN K=5
(PERSENTASE DARI GAMBAR 11)

	Precision	Recall	F1-Score
Surprise	62%	63%	62%
Happy	89%	96%	92%
Anger	92%	91%	91%
Fear	17%	24%	20%
Disgust	37%	19%	25%
Sad	93%	90%	92%
Akurasi Rata-rata	82%		



Gambar 10. *Confusion Matrix* K=5 dari hasil pengujian menggunakan google colab

	precision	recall	f1-score	support
0	0.62	0.63	0.62	172
1	0.89	0.96	0.92	371
2	0.92	0.91	0.91	313
3	0.17	0.24	0.20	41
4	0.37	0.19	0.25	74
5	0.93	0.90	0.92	264
accuracy			0.82	1235
macro avg	0.65	0.64	0.64	1235
weighted avg	0.81	0.82	0.81	1235

Gambar 11. Implementasi K-NN dengan K=5 dari hasil pengujian menggunakan google colab

Dari evaluasi hasil klasifikasi K-NN yang dilakukan dengan menggunakan ekstraksi CAS dan HOG terhadap 6 jenis *Micro-Expression* yaitu *surprise*, *happy*, *anger*, *fear*, *disgust*, dan *sad*. Akurasi yang didapatkan dibagi menjadi tiga yaitu dengan secara rata-rata nilai K=1 sebesar 85%, nilai K=3 sebesar 83% dan nilai K=5 sebesar 82%. Hasil Akurasi yang didapatkan menurun dengan penentuan K yang semakin tinggi sehingga ditetapkan K=1 untuk hasil yang maksimal dalam klasifikasi *micro-expression* enam label.

Adapun akurasi setiap jenis label pada K=1 adalah 69% untuk ekspresi *surprise*, 95% untuk ekspresi *happy*, 96% untuk ekspresi *anger*, 24% untuk ekspresi *fear*, 38% untuk ekspresi *disgust* dan 95% untuk ekspresi *sad*. Adapun hal yang paling mencolok adalah akurasi ekspresi *fear* dan *disgust* yaitu dibawah 40%. Ketika diobservasi dalam percobaan pertama, kedua dataset ini memiliki jumlah data yang paling sedikit dikarenakan siswa tidak dapat mengeluarkan ekspresi tersebut meskipun sudah dilakukan wawancara. Hal ini merupakan salah satu faktor kelemahan dari algoritma K-NN dengan ekstraksi fitur CAS dan HOG.

Pada observasi percobaan kedua dengan nilai K=3 terjadi penurunan akurasi pada 4 ekspresi tetapi pada ekspresi *disgust* terjadi peningkatan yang sebelumnya adalah 38% menjadi 42%. Dan pada observasi percobaan ke tiga dengan nilai K=5 terjadi penurunan akurasi pada 5 ekspresi dan yang paling mencolok adalah pada ekspresi *fear* yaitu dari 27% ke 17%. Hal ini disebabkan rata-rata ekspresi *fear* mengambil dataset dari affecnet sebanyak 180, dan dataset dari siswa tidak ada. Siswa sendiri sangat sulit memunculkan *micro-Expression* secara spontan pada ekspresi *fear* dan *disgust*. Hal ini menunjukkan bahwa banyaknya dataset mempengaruhi akurasi tiap ekspresi dari pengujian klasifikasi *micro-expression* menggunakan K-NN menggunakan fitur CAS dan HOG.

V. KESIMPULAN

Kesimpulan pada penelitian ini adalah ekstraksi fitur menggunakan CAS sangat tidak cocok untuk pengujian realtime dikarenakan untuk memproses 3 skala membutuhkan waktu sekitar 15-20 menit untuk 1 citra. CAS juga harus membutuhkan dataset yang RGB dikarenakan prinsip CAS yang pertama pertimbangan lokal tingkat rendah, termasuk faktor seperti kontras dan warna sangat mempengaruhi dari saliency dari citra tersebut. Dataset *fear* dan *disgust* sangat sulit didapat dikarenakan sangat sulit sekali para siswa menunjukkan kedua ekspresi tersebut

secara spontan. Hasil dari klasifikasi menggunakan K-NN didapatkan akurasi 85% dengan menggunakan K=1, sedangkan dengan K=3 mendapatkan akurasi 83%, sedangkan dengan K=5 mendapatkan akurasi 82%. Proses klasifikasi K-NN sangat cepat dan akurasinya juga sangat bagus untuk mengklasifikasikan *micro expression* dan sangat disarankan untuk dijadikan metode computer vision yang sangat mudah difahami.

PERAN PENULIS

Nikko Riestian Putra Wardoyo: Analisa formal, Pengumpulan dataset, Metodologi, Perangkat lunak, Visualisasi & Administrasi proyek.

Joan Santoso: Penulisan Review & Penyuntingan;

Esther Irawati Setiawan: Penulisan Review & Penyuntingan;

COPYRIGHT



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

DAFTAR PUSTAKA

- [1] X. Ben *et al.*, "Video-based facial micro-expression analysis: A survey of datasets, features and algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, pp. 5826–5846, 2021.
- [2] L. Lei, T. Chen, S. Li, and J. Li, "Micro-expression recognition based on facial graph representation learning and facial action unit fusion," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 1571–1580.
- [3] Y. Liu, H. Du, L. Zheng, and T. Gedeon, "A neural micro-expression recognizer," in *2019 14th IEEE international conference on automatic face & gesture recognition (FG 2019)*, 2019, pp. 1–4.
- [4] N. Van Quang, J. Chun, and T. Tokuyama, "CapsuleNet for micro-expression recognition," in *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, 2019, pp. 1–7.
- [5] N. Amynarto, Y. A. Sari, and R. C. Wihandika, "Pengenalan emosi berdasarkan ekspresi mikro menggunakan metode local binary pattern," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 10, pp. 3230–3238, 2018.
- [6] D. Goleman, R. Boyatzis, and A. McKee, "The emotional reality of teams," *J. Organ. Excell.*, vol. 21, no. 2, pp. 55–65, 2002.
- [7] P. Ekman, "What scientists who study emotion agree about," *Perspect. Psychol. Sci.*, vol. 11, no. 1, pp. 31–34, 2016.
- [8] W.-J. Yan, Q. Wu, Y.-J. Liu, S.-J. Wang, and X. Fu, "CASME database: A dataset of spontaneous micro-expressions collected from neutralized faces," in *2013 10th IEEE international conference and workshops on automatic face and gesture recognition (FG)*, 2013, pp. 1–7.
- [9] S.-T. Liong, Y. S. Gan, J. See, H.-Q. Khor, and Y.-C. Huang, "Shallow triple stream three-dimensional cnn (ststnet) for micro-expression recognition," in *2019 14th IEEE international conference on automatic face & gesture recognition (FG 2019)*, 2019, pp. 1–5.
- [10] N. K. A. Wirdiani, P. Hridayami, N. P. A. Widiari, K. D. Rismawan, P. B. Candradinata, and I. P. D. Jayantha, "Face identification based on K-nearest neighbor," *Sci. J. Informatics*, vol. 6, no. 2, pp. 150–159, 2019.
- [11] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, pp. 1915–1926, 2011.

- [12] W. Kim and C. Kim, "A novel image importance model for content-aware image resizing," in *2011 18th IEEE International Conference on Image Processing*, 2011, pp. 2469–2472.
- [13] M. Ahmadi, N. Karimi, and S. Samavi, "Context-aware saliency detection for image retargeting using convolutional neural networks," *Multimed. Tools Appl.*, vol. 80, pp. 11917–11941, 2021.
- [14] L. R. Cerna, G. Camara-Chavez, and D. Menotti, "Face detection: Histogram of oriented gradients and bag of feature method," in *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (ICCV)*, 2013, p. 1.
- [15] O. Déniz, G. Bueno, J. Salido, and F. la Torre, "Face recognition using histograms of oriented gradients," *Pattern Recognit. Lett.*, vol. 32, no. 12, pp. 1598–1603, 2011.
- [16] M. A. Hameed, M. Hassaballah, S. Aly, and A. I. Awad, "An adaptive image steganography method based on histogram of oriented gradient and PVD-LSB techniques," *IEEE Access*, vol. 7, pp. 185189–185204, 2019.
- [17] R. G. Guendel, F. Fioranelli, and A. Yarovoy, "Phase-based classification for arm gesture and gross-motor activities using histogram of oriented gradients," *IEEE Sens. J.*, vol. 21, no. 6, pp. 7918–7927, 2020.
- [18] M. M. Fouad, H. M. Zawbaa, T. Gaber, V. Snasel, and A. E. Hassanien, "A fish detection approach based on BAT algorithm," in *The 1st International Conference on Advanced Intelligent System and Informatics (AISII2015), November 28-30, 2015, Beni Suef, Egypt*, 2016, pp. 273–283.
- [19] M. Murugappan *et al.*, "Facial expression classification using KNN and decision tree classifiers," in *2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP)*, 2020, pp. 1–6.
- [20] R. Mohamed, M. M. Yusof, N. Wahid, N. Murli, and M. Othman, "Bat algorithm and k-means techniques for classification performance improvement," *Indones J Electr Eng Comput Sci.*, vol. 15, no. 3, pp. 1411–1418, 2019.
- [21] L. Farokhah, "Implementasi K-Nearest Neighbor untuk Klasifikasi Bunga Dengan Ekstraksi Fitur Warna RGB," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 6, pp. 1129–1135, 2020.
- [22] K. A. Sugiarta, I. Cholissodin, and E. Santoso, "Optimasi K-Nearest Neighbor Menggunakan Bat Algorithm Untuk Klasifikasi Penyakit Ginjal Kronis," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 10, pp. 10301–10308, 2019.
- [23] K. W. Mahardika, Y. A. Sari, and A. Arwan, "Optimasi K-Nearest Neighbour Menggunakan Particle Swarm Optimization pada Sistem Pakar untuk Monitoring Pengendalian Hama pada Tanaman Jeruk," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 9, pp. 3333–3344, 2018.
- [24] W. Nugroho, "Optimasi Metode K-Nearest Neighbours dengan Backward Elimination Menggunakan Dataset Software Effort Estimation," *Bianglala Inform.*, vol. 8, no. 2, pp. 129–133, 2020.
- [25] F. Fandiansyah, J. Y. Sari, and I. P. Ningrum, "Pengenalan Wajah Menggunakan Metode Linear Discriminant Analysis dan k Nearest Neighbor," *Ultim. J. Tek. Inform.*, vol. 9, no. 1, pp. 1–9, 2017.

Implementasi Metode Fuzzy Analytica Hierarchy Process (FAHP) Dalam Penentuan Bobot Seleksi Mahasiswa Program Pendidikan Dokter Spesialis

Ni N. Murni¹, Gede R. Dantes¹, dan I Made Candiasa¹

¹ Program Studi Ilmu Komputer, Universitas Pendidikan Ganesha Singaraja, Bali Indonesia

Corresponding author: Ni Nyoman Murni (e-mail:nyomanmurni.ilkom@gmail.com).

ABSTRACT The purpose of this study was to implement the Fuzzy Analytical Hierarchy Process (F-AHP) method in determining the weight of admission of specialist medical education students in Skin and Gender Health Sciences at the Faculty of Medicine, Udayana University and measure the level of accuracy using confusion matrix. The F-AHP method is able to reduce the subjective assessment of criteria from the AHP method and produce better decisions. The calculation process starts from compiling a pairwise comparison matrix from the interview results, calculating the eigenvalue, calculating the consistency ratio value, converting the criteria weighting into the form of a paired matrix of F-AHP criteria, calculating the fuzzy syntethic extent value for each criterion until getting the weight of each variable. The resulting weights are: C1 = basic competency test by 30%, C2 = interview test by 30%, C3 = academic potential test by 17%, and C4 = TOEFL by 23%. Accuracy testing was carried out using confusion matrix and obtained a total accuracy of 89%, while per-predicate graduation accuracy was obtained for cumlaude = 93%, very satisfying = 84% and satisfying = 90%. The findings indicate that the weights obtained are significant to be applied in the selection process of specialist medical education program students.

KEYWORDS: Confusion Matrix, Fuzzy Analytical Hierarchy Process, Weighting

ABSTRAK Tujuan dari penelitian ini adalah mengimplementasi metode Fuzzy Analytical Hierarchy Process (F-AHP) dalam menentukan bobot penerimaan mahasiswa pendidikan dokter spesialis Ilmu Kesehatan Kulit dan Kelamin di Fakultas Kedokteran Universitas Udayana dan mengukur tingkat akurasi menggunakan confusion matrix. Metode F-AHP mampu mengurangi penilaian kriteria yang bersifat subjektif dari metode AHP dan menghasilkan keputusan yang lebih baik. Proses perhitungan dimulai dari menyusun matriks perbandingan berpasangan dari hasil wawancara, menghitung nilai eigen, menghitung nilai rasio konsistensi, mengubah pembobotan kriteria ke dalam bentuk matrik berpasangan kriteria F-AHP, menghitung nilai fuzzy syntethic extent untuk tiap kriteria sampai mendapatkan bobot dari masing-masing variabel. Bobot yang dihasilkan adalah: C1 = tes kompetensi dasar sebesar 30%, C2 = tes wawancara sebesar 30%, C3= tes potensi akademik sebesar 17%, dan C4 = TOEFL sebesar 23%. Pengujian akurasi dilakukan dengan menggunakan confusion matrix dan diperoleh akurasi total sebesar 89%, sedangkan akurasi per-predikat kelulusan diperoleh untuk cumlaude = 93%, sangat memuaskan = 84% dan memuaskan = 90%. Temuan tersebut mengindikasikan bahwa bobot yang diperoleh signifikan untuk diterapkan dalam proses seleksi mahasiswa program pendidikan dokter spesialis.

KATA KUNCI: *Confusion Matrix, Fuzzy Analytical Hierarchy Process, Pembobotan*

I. PENDAHULUAN

Pendidikan dokter spesialis adalah lanjutan pendidikan profesi dokter, merupakan perpaduan pendidikan keprofesian yang dilandasi kemampuan bidang keilmuan (akademik). Dalam pendidikan dokter spesialis proses pembelajaran dilaksanakan melalui supervisi untuk meningkatkan kompetensi mahasiswa agar setelah lulus bisa melakukan praktek secara mandiri [1], [2]. Kegiatan seleksi penerimaan mahasiswa baru pada Program studi Ilmu Kesehatan Kulit dan Kelamin Fakultas Kedokteran Universitas Udayana yang telah berjalan masih memiliki kendala yaitu, persentase dari kriteria kemampuan yang diuji pada saat seleksi ditentukan berdasarkan keputusan rapat sehingga belum teruji apakah persentase pada seleksi dapat mencerminkan presentasi mahasiswa, banyaknya jumlah peminat dan sedikitnya jumlah daya tampung menyebabkan proses seleksi berlangsung cukup lama [3]. Pembobotan berdasarkan hasil keputusan rapat memiliki kelemahan yaitu persentase pembobotan dari masing-masing variabel tidak bisa dijelaskan atau diuji secara ilmiah apakah bobot yang dihasilkan berdasarkan keputusan tersebut sudah sesuai untuk masing-masing variabel. Metode F-AHP mampu mengurangi penilaian kriteria yang bersifat subjektif dari metode AHP dan menghasilkan persentase pembobotan yang lebih akurat. variable dari masing-masing variabel yang lebih.

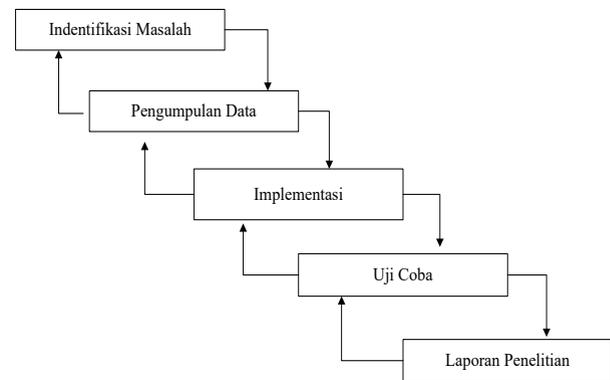
Metode *F-AHP* dapat diimplementasikan dalam penentuan peminatan siswa dengan tahapan yaitu, membuat matriks perbandingan berpasangan dengan skala *AHP*, transformasi matriks perbandingan berpasangan ke dalam skala TFN, menghitung nilai sintesis *fuzzy* (S_i), nilai vektor (V) dan ordinat defuzzyfikasi (d'), normalisasi, menghitung rasio konsistensi, dan menghitung nilai bobot alternatif [4], [5]. Penulis melakukan penelitian pada bobot seleksi mahasiswa Pendidikan dokter spesialis Ilmu Kesehatan Kulit dan Kelamin karena Pembobotan hasil seleksi masih menggunakan model klasik yaitu melalui keputusan rapat pimpinan, perankingan skor akhir masih menggunakan model klasik dengan menggunakan bobot yang sudah disetujui oleh rapat pimpinan. Setelah mendapatkan bobot dari hasil perhitungan *F-AHP* akan dilakukan pengujian akurasi yang bertujuan untuk mengetahui apakah bobot hasil perhitungan dengan metode *F-AHP* menghasilkan bobot yang lebih baik jika dibandingkan dengan bobot yang sudah digunakan. Akurasi diuji dengan metode *confusion matrix* [6].

Confusion matrix adalah metode yang digunakan pengukuran akurasi nilai prediksi dan nilai aktual. Pada *confusion matrix* memiliki tabel dengan 4 kombinasi. Ada empat istilah pada *confusion matrix* yaitu *true positif*, *true negatif*, *false positif*, dan *false negatif* [7], [8].

II. METODE PENELITIAN

Pada Penelitian ini menggunakan metode Research and Development, seperti yang ditunjukkan pada gambar 1, yang digunakan untuk menghasilkan produk tertentu, dan menguji

keefektifan produk tersebut [9]. Penelitian ini mengimplementasikan metode F-AHP dalam menentukan bobot penerimaan mahasiswa pendidikan dokter spesialis.



Gambar 1. Rancangan Penelitian

1) Identifikasi Masalah

Pada tahap ini, mengidentifikasi masalah penentuan bobot pada proses seleksi mahasiswa, menentukan rumusan masalah dan ruang lingkup penelitian.

2) Pengumpulan data

Pada tahap ini dilakukan pengumpulan data sesuai dengan identifikasi masalah untuk memperoleh informasi yang dibutuhkan

3) Implementasi

Tahap implementasi adalah tahap untuk menerapkan dan melakukan perhitungan dengan menggunakan metode *F-AHP* untuk memperoleh persentase bobot di setiap variabel.

4) Pengujian (Testing)

Proses pengujian untuk mengukur tingkat akurasi metode *F-AHP* untuk menentukan penerimaan mahasiswa dengan menggunakan metode *confusion matrix*.

5) Penulisan Laporan

Merupakan penyelesaian akhir dengan membuat laporan berdasarkan penelitian yang dilakukan.

III. HASIL DAN PEMBAHASAN

Pada penelitian ini terdapat empat kriteria utama yang harus dibandingkan yaitu: tes kompetensi dasar (C1), tes wawancara (C2), tes potensi akademik (C3) dan TOEFL (C4). Pengumpulan data yang digunakan dalam menentukan tingkat prioritas dari variabel untuk perhitungan matrik berpasangan pada proses AHP menggunakan metode wawancara dengan tiga responden yang merupakan pimpinan di Program Studi Ilmu Kesehatan Kulit dan Kelamin Fakultas Kedokteran Universitas Udayana.

1. Menyusun matriks perbandingan berpasangan seperti pada tabel 1, menghitung vektor prioritas (*Eigenvector*) seperti pada tabel 2 untuk kriteria utama.

TABEL I
MATRIK PERBANDINGAN BERPASANGAN

Kriteria	C1	C2	C3	C4
C1	1,00	1,00	3,00	3,00
C2	1,00	1,00	3,00	3,00
C3	0,33	0,33	1,00	0,33
C4	0,33	0,33	3,00	1,00

TABEL II
VEKTOR PRIORITAS

Kriteria	C1	C2	C3	C4	Total	Vektor Prioritas
C1	0,376	0,376	0,300	0,409	1,461	0,365
C2	0,376	0,376	0,300	0,409	1,461	0,365
C3	0,124	0,124	0,100	0,045	0,393	0,098
C4	0,124	0,124	0,300	0,136	0,685	0,171

$$S_2 = (3,67, 5,00, 7,00) \times (0,0444, 0,0594, 0,0789) = (0,163, 0,297, 0,553)$$

$$S_3 = (2,50, 3,00, 4,00) \times (0,0444, 0,0594, 0,0789) = (0,111, 0,178, 0,316)$$

$$S_4 = (3,00, 3,83, 5,00) \times (0,0444, 0,0594, 0,0789) = (0,133, 0,228, 0,395)$$

Perbandingan ini akan digunakan untuk nilai bobot pada setiap kriteria dengan nilai minimumnya, seperti pada tabel 3.

TABEL III
NILAI PERBANDINGAN BILANGAN FUZZY KONVEKS

	S1	S2	S3	S4
S1		1,000	0,574	0,775
S2	1,000		0,56261	0,770
S3	1,419	1,368		1,000
S4	1,000	1,000	0,787	
Minimum	1,000	1,000	0,563	0,770

Vektor bobot dilakukan agar mempermudah interpretasi yang diperoleh dari nilai minimum perbandingan bilangan fuzzy konveks. Jika vektor bobot dinormalisasi, seperti pada tabel 4, maka nilainya menjadi analog bobot dan terdiri dari bilangan non-fuzzy dari proses ini akan diperoleh nilai bobot kriteria utama [11].

TABEL IV
NORMALISASI VEKTOR BOBOT

	C1	C2	C3	C4
W	0,30	0,30	0,17	0,23

B. HASIL PERHITUNGAN BOBOT SELEKSI DENGAN METODE F-AHP

Dari proses perhitungan diatas didapatkan hasil akhir dari perhitungan bobot seleksi mahasiswa dengan menggunakan metode F-AHP dapat dilihat pada tabel 5.

TABEL V
HASIL AKHIR NILAI BOBOT PER VARIABEL

Variabel	Nilai Bobot
C1 = tes kompetensi dasar	30
C2 = tes wawancara	30
C3 = tes potensi akademik	17
C4 = tes kemampuan bahasa inggris / TOEFL	23

Pada tabel 5 dapat kita lihat bobot dari masing-masing variabel yaitu C1 = kompetensi dasar (%) mendapatkan bobot 0,30 atau 30%, C2 = tes wawancara (%) mendapatkan bobot 0,30 atau 30%, C3= tes potensi akademik (%) mendapatkan bobot 0,17 atau 17% dan C4 = tes kemampuan bahasa inggris / TOEFL (%) mendapatkan bobot r 0,23 atau 23%. Jumlah total bobot dari C1 + C2 + C3 + C4 = 30 + 30 + 17 + 23 = 100%.

2. Menghitung Nilai Eigen (*Eigenvalue*)

a. Matrik perbandingan berpasangan dikalikan dengan vektor prioritas kemudian hasilnya disebut vektor jumlah bobot.

$$\begin{bmatrix} 1,00 & 1,00 & 3,00 & 3,00 \\ 1,00 & 1,00 & 3,00 & 3,00 \\ 0,33 & 0,33 & 1,00 & 0,33 \\ 0,33 & 0,33 & 3,00 & 1,00 \end{bmatrix} \times \begin{bmatrix} 0,37 \\ 0,37 \\ 0,10 \\ 0,17 \end{bmatrix} = \begin{bmatrix} 1,539 \\ 1,539 \\ 0,396 \\ 0,707 \end{bmatrix}$$

b. Membagi vektor jumlah bobot dengan entri yang berpasangan dari vektor prioritas kemudian hasil disebut bobot prioritas (nilai eigen). Bobot prioritas = $\frac{1,539}{0,365} \quad \frac{1,539}{0,365} \quad \frac{0,396}{0,098} \quad \frac{0,396}{0,171} = [4,213 \quad 4,213 \quad 4,028 \quad 4,132]$

c. Menghitung rata-rata nilai bobot prioritas kemudian hasilnya dinotasikan dengan λ_{maks} .
 $\lambda_{maks} = \frac{4,213 + 4,213 + 4,02 + 4,132}{4} = 4,146$

3. Menghitung rasio konsistensi (CR)

a. Menghitung *Consistency Index* (CI).

$$CI = \frac{\lambda_{maks} - n}{n - 1} = CI = \frac{4,146 - 4}{4 - 1} = 0,048$$

b. Menghitung *Consistency Ratio* (CR).

$$CR = \frac{CI}{RI} \text{ dengan } n = 4 \text{ maka } RI = 0,9$$

$$CR = \frac{0,048}{0,9} = 0,054$$

Menurut Saaty [10], jika $CR \leq 10\%$ maka matriks perbandingan berpasangan disebut konsisten. Konsisten memiliki arti setiap element sudah dikelompokkan secara homogen dan relasi antara kriteria membenarkan secara logis.

A. FUZZY AHP DAN FUZZY SYNTHETIC EXTENT

1. Merubah pembobotan kriteria ke dalam bentuk matrik berpasangan kriteria fuzzy AHP dan perhitungan jumlah bilangan *triangular fuzzy* setiap baris.
2. Menghitung nilai *fuzzy syntethic extent* untuk tiap kriteria.

$$S_1 = (3,50, 5,00, 6,50) \times (0,0444, 0,0594, 0,0789) = (0,156, 0,297, 0,513)$$

C. Pengujian Tingkat Akurasi

Pada penelitian ini tahap pengujian akurasi dilakukan dengan menggunakan validasi metode confusion matrix seperti pada tabel 6. Tahap pengujian akurasi dilaksanakan agar mengetahui tingkat akurasi dari metode yang digunakan.

Pengujian akurasi dilakukan dengan membandingkan hasil perhitungan sistem dengan data latih dengan menggunakan (1). Pengujian akurasi dibuat dalam bentuk tabel perbandingan hasil. Hasil sistem yang digunakan adalah implementasi dari algoritma F-AHP sedangkan data latih yang digunakan adalah IPK mahasiswa angkatan 2013 sampai dengan 2017 yang berjumlah 66 orang. IPK mahasiswa kemudian dikelompokkan sesuai dengan syarat predikat kelulusan Magister / Spesialis yang terdapat pada pedoman akademik universitas Udayana.

$$\frac{TP+TN}{(TP+FP+FN+TN)} \quad (1)$$

Kelompok pertama adalah mahasiswa dengan nilai IPK >3,75 dimasukkan kategori *cumlaude*. Kelompok kedua adalah mahasiswa dengan nilai IPK 3,50–3,75 dimasukkan kategori sangat memuaskan. Kelompok ketiga adalah mahasiswa dengan nilai IPK 3,00 <3,50 dimasukkan kategori memuaskan.

Tahap berikutnya adalah membandingkan nilai IPK mahasiswa yang dibandingkan dengan hasil perhitungan yang diperoleh dengan metode F-AHP seperti pada tabel 7, tabel 8, dan tabel 9. Kolom sesuai diisi jika hasil sistem dan data uji bernilai sama dan pada kolom tidak sesuai jika hasil sistem dan data uji tidak sama. Jumlah mahasiswa untuk kategori *cumlaude* sebanyak 15 orang, untuk kategori sangat memuaskan sebanyak 25 orang, untuk kategori memuaskan sebanyak 26 orang, dan untuk kategori cukup memuaskan sebanyak 48 orang. Hasil perhitungan metode F-AHP akan dimasukkan ke dalam masing-masing kategori berdasarkan jumlah mahasiswa dalam masing-masing kelompok. Untuk kategori *cumlaude* adalah mahasiswa peringkat 1-15, untuk kategori sangat memuaskan adalah mahasiswa peringkat 15-40, untuk kategori memuaskan adalah mahasiswa peringkat 41-66.

TABEL VI
CONFUSION MATRIX

Kelas aktual	Kelas prediksi		
	<i>cumlaude</i>	sangat memuaskan	memuaskan
<i>cumlaude</i>	12	3	0
sangat memuaskan	1	20	4
memuaskan	0	2	24

TABEL VII
CONFUSION MATRIX CUMLAUDE

Kelas aktual	Kelas prediksi	
	<i>cumlaude</i>	<i>non cumlaude</i>
<i>cumlaude</i>	12	3

<i>non cumlaude</i>	1	50
Akurasi <i>Cumlaude</i>	$= (12 + 50) / (12+3+1+50)$ $= 0,93 * 100 \%$ $= 93 \%$	

TABEL VIII
CONFUSION MATRIX SANGAT MEMUASKAN

Kelas aktual	Kelas prediksi	
	sangat memuaskan	<i>non</i> sangat memuaskan
sangat memuaskan	20	5
<i>non</i> sangat memuaskan	5	36

Akurasi sangat memuaskan = $(20 + 36) / (20+5+5+36)$
 = $0,84 * 100 \%$
 = 84%

TABEL IX
CONFUSION MATRIX MEMUASKAN

Kelas aktual	Kelas prediksi	
	memuaskan	<i>non</i> memuaskan
memuaskan	24	2
<i>non</i> memuaskan	4	36

Akurasi memuaskan = $(24 + 36) / (24+2+4+36)$
 = $0,90 * 100 \%$
 = 90%

Akurasi Total = $\frac{93\%+84\%+90\%}{3}$
 = 89%

Dari perhitungan akurasi diperoleh akurasi total sebesar 89 %. Apabila akurasi dihitung berdasarkan kelas prediksi, maka mendapatkan hasil sebagai berikut.

- Cumlaude* = 93%
- Sangat Memuaskan = 84%
- Memuaskan = 90%

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat ditarik kesimpulan sebagai berikut.

- Penggunaan metode *F-AHP* pada penelitian ini mampu menentukan bobot penerimaan mahasiswa pendidikan dokter spesialis ilmu Kesehatan kulit dan kelamin di Fakultas Kedokteran Universitas Udayana sehingga dapat membantu Program Studi Spesialis Ilmu Kesehatan Kulit dan Kelamin Universitas Udayana dalam proses selesi mahasiswa untuk menentukan mahasiswa yang akan diterima.
- Pada penelitian selanjutnya dapat ditambahkan variabel penelitian yang berhubungan dengan bidang ilmu spesialis yang dipilih misalkan untuk jurusan spesialis Kulit dan Kelamin bisa ditambahkan variabel pengetahuan tentang ilmu kesehatan kulit dan kelamin (menjelaskan salah satu kasus penyakit kulit dan kelamin; tentang penyebab, patogenesis, epidemiologi, gambaran klinis, diagnosis penatalaksanaan)

PERAN PENULIS

Ni Nyoman Murni: Konseptualisasi, Metodologi dan Implementasi Sistem, Uji Coba, Penyusunan Draft Asli;

Gede Rasben Dantes: Penyusunan Draf Asli, Penulisan Review dan Penyuntingan;

I Made Candiasta: Penyusunan Draf Asli, Penulisan Review dan Penyuntingan;

COPYRIGHT



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

DAFTAR PUSTAKA

- [1] Konsil Kedokteran Indonesia, *Standar Pendidikan Profesi Dokter Spesialis*. Jakarta, 2006.
- [2] Kolegium Dermatologi dan Venereologi, *Panduan Kolegium Dermatologi dan Venereologi*. Jakarta, 2015.
- [3] Universitas Udayana, *POB Penerimaan Mahasiswa Baru Udayana*. Denpasar, 2015.
- [4] M. Fajri, R. R. M. Putri, and L. Muflikhah, "Implementasi Metode Fuzzy Analytic Hierarchy Process (F-AHP) Dalam Penentuan Peminatan di MAN 2 Kota Serang," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 5, pp. 2109–2117, 2018.
- [5] A. Santoso, R. Rahmawati, and S. Sudarno, "Aplikasi Fuzzy Analytical Hierarchy Process Untuk Menentukan Prioritas Pelanggan Berkunjung Ke Galeri (Studi Kasus Di Secondhand Semarang)," *J. Gaussian*, vol. 5, no. 2, pp. 239–248, 2016.
- [6] Rusyati, *Buku Panduan Peserta PPDS Ilmu Kesehatan Kulit dan Kelamin*. Denpasar, 2017.
- [7] F. Gorunescu, *Data Mining: Concepts, models and techniques*, vol. 12. Springer Science & Business Media, 2011.
- [8] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier Science, 2011.
- [9] P. Sugiyono, "Metodologi penelitian kuantitatif kualitatif dan R&D," *Alfabeta, Bandung*, pp. 62–70, 2011.
- [10] T. L. Saaty, "Pengambilan Keputusan Bagi Para Pemimpin. L. Setiono [Penerjemah]; K. Peniwati [Editor]. Terjemahan dari," *Decis. Mak. Leaders Anal. Hierarchy Process Decis. Complex World*, 1993.
- [11] I. D. M. A. B. Joni and A. A. G. B. Ariana, "Sistem Pendukung Keputusan Seleksi Penerimaan Dosen Tetap Yayasan Dengan Metode Fuzzy-AHP," *Netw. Eng. Res. Oper.*, vol. 1, no. 2, 2014.

Algoritma *Random Forest* pada Prediksi Status Kredit Usaha Rakyat untuk Mengurangi *Nonperforming Loan Rate*

Wulansari¹ dan Diana Purwitasari²

¹Magister Manajemen Teknologi, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

²Departemen Teknik Informatika, Fakultas Teknologi Elektro dan Informatika Cerdas, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Corresponding author: Wulansari (e-mail: wulansarirachmat27@gmail.com).

ABSTRACT Small and Micro credit program (KUR) is one of Indonesian government's programs in improving access to financing for small and micro segments, channeled through financial institutions with guarantees. KUR distribution funds originate 100% from KUR-distributing financial institutions. The government provides subsidies in the form of credit guarantees, whereby the government secures 70% of the total loan value and financial institutions, such as banks, assume the remaining 30%. Ining a low non-performing loan (NPL) rate for KUR is crucial to averting the formation of loss-reserve expenses that will reduce bank profits. Predicting the variables that influence the borrowing impasse is essential for preventing an increase in NPLs, particularly in the early phases of lending. No research has been conducted to predict KUR default using machine learning. The purpose of this research Bank could predict through the use of machine learning and determine which variables contribute to KUR default. The SMOTE technique is also used to solve data imbalance. The results show that random forest classifiers outperform logistic regression with accuracy 88% respectively on data testing. Contributions of each variable that affects KUR default rank by feature importance are credit restructuring, availability of debt service payment account, unit operations, and business sector. The Bank may use this as a foundation to assess the risk profile of the new KUR procedure, hence mitigating the likelihood of problematic loans in the future and decreasing the NPL rate.

KEYWORDS KUR, Logistic Regression, NPL, *Random Forest*

ABSTRAK Program Kredit Usaha Rakyat (KUR) adalah salah satu program pemerintah Indonesia dalam meningkatkan akses pembiayaan kepada Usaha Mikro, Kecil, dan Menengah (UMKM) yang disalurkan melalui lembaga keuangan dengan pola penjaminan. Dana penyaluran KUR 100% berasal dari dana lembaga keuangan penyalur KUR. Pemerintah memberikan subsidi berupa pembayaran penjaminan kredit yang menjamin kredit sebesar 70% dari nilai total kredit yang ditetapkan dan 30% sisanya ditanggung lembaga keuangan salah satunya adalah bank. Mempertahankan *non performing loan (NPL) rate* yang rendah sangat penting untuk mencegah terjadinya pembentukan biaya pencadangan kerugian yang akan mengurangi keuntungan bank. Memprediksi variabel yang berpengaruh terhadap macetnya KUR diperlukan untuk mencegah kredit macet terutama pada tahap awal pemberian kredit. Belum terdapat penelitian untuk memprediksi kredit macet pada KUR menggunakan *machine learning*. Penelitian ini bertujuan agar Bank dapat melakukan prediksi dengan pendekatan *machine learning* dan mengetahui kontribusi variabel yang mempengaruhi KUR macet. Teknik SMOTE juga digunakan dalam penelitian untuk mengatasi ketidakseimbangan data. Hasil yang diperoleh menunjukkan bahwa klasifikasi *random forest* memberikan akurasi lebih tinggi dari regresi logistik yaitu sebesar 88% pada data uji. Urutan tingkat kepentingan dari kontribusi variabel yang mempengaruhi macet pada KUR adalah status restrukturisasi kredit, ketersediaan *debt service payment account*, unit operasional pemroses kredit, dan sektor usaha dapat menjadi dasar Bank untuk menganalisis profil resiko proses KUR baru sehingga dapat menurunkan potensi kredit bermasalah kedepannya dan menekan *NPL rate*.

KATA KUNCI KUR, NPL, *Random Forest*, Regresi Logistik

I. PENDAHULUAN

Bank adalah badan usaha yang menghimpun dana dari masyarakat dalam bentuk simpanan dan menyalurkannya kepada masyarakat dalam bentuk kredit dan atau bentuk-bentuk lainnya dalam rangka meningkatkan taraf hidup rakyat banyak [1]. Kredit usaha rakyat (KUR) merupakan pembiayaan kepada UMKM yang disalurkan melalui lembaga keuangan dengan pola penjaminan. Dana penyaluran KUR 100% berasal dari dana lembaga keuangan penyalur KUR. Pemerintah memberikan subsidi berupa pembayaran penjaminan kredit yang menjamin kredit sebesar 70% dari nilai total kredit yang ditetapkan sementara 30% sisanya ditanggung oleh penyalur KUR dalam hal ini lembaga keuangan [2]. Berdasarkan hal tersebut dari sudut pandang lembaga keuangan, mempertahankan rasio kredit macet KUR yang rendah sangat penting untuk mencegah terjadinya kerugian akibat kredit macet yang akan mengurangi keuntungan perusahaan.

Nonperforming loan (NPL) *rate* adalah rasio yang dapat menunjukkan seberapa baik kualitas kredit yang disalurkan oleh perbankan, dan jika rasio NPL semakin rendah maka penyaluran kredit suatu bank dinilai semakin berkualitas. Laba suatu bank akan tergerus apabila terjadi peningkatan NPL *rate*, sejalan dengan teori manajemen likuiditas *Commercial Loan Theory* yang menyatakan penyaluran kredit harus memperhatikan tingkat pengembalian kredit agar kegiatan operasional bank terus berjalan dan tetap menjaga kepercayaan nasabah [3].

Terdapat beberapa penelitian terkait prediksi kredit macet dan faktor yang menyebabkan kredit menjadi macet dengan menggunakan metode *machine learning* meskipun belum ada untuk KUR. Beberapa penelitian tersebut yaitu dengan metode Regresi Logistik dan *Random Forest*. Tujuan dari penelitian tersebut melakukan *assesment* terhadap beberapa faktor untuk membantu bank dalam melakukan seleksi terhadap permohonan kredit [4]. Penelitian lainnya untuk *loan risk prediction* model dengan menggunakan metode *random forest* adalah penelitian [5] yang menilai signifikansi masing-masing variabel (faktor), dimungkinkan untuk menentukan fitur mana yang memiliki pengaruh lebih besar terhadap hasil akhir gagal bayar, sehingga dapat dilakukan keputusan resiko kegagalan bayar yang lebih bijaksana pada bank.

Algoritma klasifikasi berguna untuk memisahkan dan mengelompokkan data untuk melakukan prediksi sebagaimana akan digunakan pada *loan prediction* [6]. Pada kasus kredit macet akan terjadi data *imbalanced* karena perbandingan yang nilainya signifikan antara klasifikasi satu kelas dengan kelas lainnya (macet/tidak macet). Salah satu metode yang digunakan untuk permasalahan *imbalanced* tersebut adalah SMOTE, teknik *oversampling* yang menghasilkan data baru untuk menyerupai data asli dari kelas yang lebih kecil sehingga dapat menciptakan keseimbangan data dan terhindar dari hasil akurasi yang bias dan menyebabkan ketidaktepatan prediksi [7].

Belum terdapat penelitian untuk memprediksi kredit macet pada KUR menggunakan *machine learning*. Penelitian ini bertujuan agar Bank dapat melakukan prediksi dengan pendekatan *machine learning* dan mengetahui kontribusi variabel yang mempengaruhi KUR macet dengan *feature importance* sehingga dapat menjadi mitigasi saat analisis permohonan kredit yang dapat mengurangi NPL *rate* KUR.

Dalam penelitian ini diusulkan menggunakan dua algoritma *machine learning* sebagai komparasi untuk akurasi terbaik dari model yang dibangun, metode tersebut dikombinasikan dengan SMOTE untuk mengatasi masalah data kelas yang tidak seimbang.

II. TINJAUAN PUSTAKA

Mempertahankan NPL *rate* KUR yang rendah sangat penting untuk mencegah terjadinya kerugian akibat kredit macet yang akan mengurangi keuntungan perusahaan atas biaya pencadangan kerugian.

Penelitian yang telah dilakukan untuk prediksi kredit macet [8] dari data German *credit* dataset dengan jumlah data sebanyak 1.000 dengan metode *Naïve Bayes*, MLP, SMO, *K-nearest neighbour's*, *Rotation Forest*, *Bagging*, *RealADABOOST*, dan *Random Forest* dengan dan tanpa SMOTE. Hasil yang diperoleh adalah *random forest* berkinerja terbaik dengan nilai akurasi sebelum SMOTE sebesar 75,3% dan nilai akurasi setelah SMOTE meningkat menjadi sebesar 83,4%. Hasil dari urutan *feature* 5 variabel yang paling berkontribusi terhadap kredit macet adalah *Over draft*, *Credit History*, *Purpose*, *Average Credit Balance* dan *Employment*.

Penelitian lainnya dari [9] dengan menggunakan metode SVM, KNN, C4.5, dan *Random Forest* dengan SMOTE untuk dataset kredit dari Kaggle dengan jumlah 150.000 diperoleh hasil *random forest* dengan SMOTE berkinerja terbaik dibandingkan metode lainnya dengan nilai akurasi sebesar 93,7%.

Metode Bayesian juga digunakan dalam penelitian [10] dengan kesimpulan mode *inpection games* dikembangkan dimana kreditur sebagai pemeriksa dan debitur sebagai yang diperiksa dan variabel penyebab NPL dari sisi debitur adalah utilitas kredit atau penggunaan kredit. Sepuluh literatur yang dikumpulkan dalam penelitian [11] dengan menggunakan metode SVM, regresi logistik, *decision tree*, dan *neural network* disimpulkan regresi logistik dan *decision tree* dinilai efektif dalam memprediksi gagal bayar pinjaman dan mengelola resiko kredit. Variabel yang berkontribusi terhadap kredit macet berdasarkan penelitian ini adalah usia, pendapatan, angsuran kredit, dan pendidikan.

Perbandingan metode *random forest* dan metode *XGBoosting* dilakukan dalam penelitian [12] dengan sumber data dari *lending club website* sebanyak 272.401. Penelitian ini menyimpulkan metode *random forest* berkinerja terbaik berdasarkan evaluasi AUC, dengan nilai AUC sebesar 0,99 lebih tinggi jika dibandingkan dengan *XGBoosting classifier* sebesar 0,97. Penelitian dari [13] dari German *credit* dataset

dengan jumlah data sebanyak 1.000 dan jumlah variabel yang digunakan 21 variabel, diperoleh hasil nilai akurasi model dengan menggunakan metode *random forest* sebesar 83% termasuk kedalam klasifikasi “*very good model*”.

Untuk penelitian dengan penekanan pada skenario data latih dan data uji dilakukan pada [14] dengan menggunakan tiga skenario perbandingan yaitu:

1. Data latih 60%: data uji 40%
2. Data latih 70%: data uji 30%, dan
3. Data latih 80%: data uji 20%.

Metode *machine learning* yang digunakan adalah metode *random forest* dengan sumber data dari data kredit BRI cabang Pelaihari. Kesimpulan dari penelitian ini skenario terbaik yaitu data latih berbanding data uji sebesar 70%:30% dan 80%:20% dengan nilai akurasi sama yaitu sebesar 97,1%, dan nilai *precision* sebesar 97,02% dan *recall* sebesar 100%.

Pada saat menafsirkan model, dalam penelitian [15] digunakan metode *feature importance* untuk menghitung peringkat kepentingan dan menyimpulkan kontribusi dari masing-masing variabel input. Selain itu dapat digunakan untuk mengukur kekuatan pengaruh dari suatu variabel prediktor terhadap variabel respon. Hal tersebut diperlukan untuk mengurutkan kontribusi terhadap kredit macet yang akan digunakan pada penelitian ini.

III. METODOLOGI

Tujuan dari penelitian ini adalah membangun model klasifikasi untuk memprediksi status KUR (lancar/macet) dengan menggunakan metode *Random Forest* dengan SMOTE dan mengukur serta menganalisis kontribusi setiap variabel yang mempengaruhi kredit macet pada KUR untuk mengurangi NPL *rate*. Metode klasifikasi lainnya yaitu regresi logistik digunakan untuk membangun model sebagai komparasi.

A. DATASET

Data pada penelitian ini merupakan data sekunder yang diperoleh dari Bank XYZ regional Surabaya. Data yang digunakan adalah data pinjaman aktif KUR dari Januari 2016 sampai dengan posisi 31 Desember 2022.

Definisi untuk kredit macet dalam penelitian ini adalah kredit yang memiliki tunggakan dengan hari tunggakan telah melampaui 90 hari. Dataset KUR yang diperoleh dengan total data pinjaman sebanyak 45.938 dengan variabel sebanyak 17. NPL *rate* KUR Bank XYZ pada posisi 31 Desember 2022 sebesar 1,0% dengan total debitur macet sebanyak 427. Terdapat 17 variabel pada data KUR yang dikumpulkan yaitu 16 variabel prediktor dan 1 variabel respon.

Dari 16 variabel prediktor awal pada tabel 1, akan dilakukan *exclude* untuk variabel TglBukaRek yang merupakan tanggal menjadi nasabah dan Tgl_PK yang merupakan tanggal menjadi debitur, disebabkan variabel tersebut tidak dapat dilakukan “*adjustment*” di masa mendatang untuk perbaikan proses kredit sebagai masukan kepada perusahaan.

TABEL I
VARIABEL PENELITIAN

Variabel	Nama	Keterangan	Skala
Y	Kol	Status kelancaran Kredit yaitu Lancar/Macet	Nominal
X1	Unit_Processing	Cabang Pemroses Kredit, terdapat 13 cabang pemroses	Nominal
X2	Produk	Pembayaran kewajiban secara Musiman/ Bulanan	Nominal
X3	Peruntukan	KUR Mikro: maksimum kredit sd Rp100Juta KUR Ritel: maksimum kredit sd Rp500 Juta	Nominal
X4	Maks Krd	Maksimum kredit	Rasio
X5	Bk Dbt	Nilai sisa pokok kredit	Rasio
X6	Suku Bunga	Nilai suku bunga	Interval
X7	JW	Jangka waktu kredit 1- 60 bulan	Interval
X8	Jaminan	Ketersediaan Jaminan: Ada/Tidak ada jaminan	Nominal
X9	Umur_Tgk	Jumlah hari menunggak kewajiban	Interval
X10	Sektor_Ek	Sektor Ekonomi usaha debitur, Terdapat 9 sektor ekonomi usaha	Nominal
X11	Nilai_Angsuran	Kewajiban per bulan debitur	Rasio
X12	DSPA	<i>Debt service reserve account</i> yaitu ketersediaan saldo dana 1x kewajiban Tersedia/Tidak tersedia	Nominal
X13	TglBukaRek.	Tanggal pembukaan rekening dana	Nominal
X14	Tgl_PK	Tanggal menjadi debitur sd 2019, 2020, 2021, 2022	Nominal
X15	R3	Restrukturisasi kredit Restruk/Non Restruk	Nominal
X16	Tipe_Debitur	Perorangan atau Non Perorangan	Nominal

B. ALUR PENELITIAN

Berdasarkan gambar 1 alur penelitian dimulai dengan identifikasi masalah, tinjauan literatur komprehensif dilakukan untuk mencari metodologi dan pendekatan yang ada yang dapat menyelesaikan masalah secara efektif.

Tahap selanjutnya sebelum analisis dimulai, data diproses terlebih dahulu untuk mengidentifikasi dan memperbaiki nilai yang hilang dan data outlier dihilangkan untuk memenuhi persyaratan penelitian. Tahap ini sangat penting untuk memastikan bahwa dataset dibersihkan secara menyeluruh dan dipersiapkan untuk analisis lebih lanjut dan pembuatan model prediksi. Untuk memahami karakteristik data, dilakukan *exploratory data analysis* (EDA) sehingga dapat diketahui karakteristik dari data set, fitur apa yang membuat data tersebut berbeda atau sama, bagaimana cara terbaik dalam menjelaskan data, karakteristik terpenting yang diperlukan. Kemudian dilakukan penghilangan data *outlier* yang tidak sesuai dengan karakteristik data serta pengkodean variabel (*encoding*) untuk dilakukan proses pengolahan data lebih

lanjut. Dari hasil pemeriksaan tidak ditemukan nilai *null* sehingga dapat dilanjutkan pada tahapan selanjutnya yaitu pemeriksaan *outlier* yaitu apakah terdapat maksimum KUR diatas Rp.500 Juta. Hasil yang diperoleh pada data tidak ditemukan adanya nilai *outlier*.

Hal yang selanjutnya dilakukan adalah penerapan *one hot encoding* terhadap data yang bersifat kategorikal yaitu unit *processing*, produk, peruntukan, ketersediaan jaminan, sektor ekonomi, *debt service payment account*, tahun menjadi debitur, status restrukturisasi kredit, dan tipe debitur. Dilakukan pembagian data latih dan data uji dengan perbandingan yaitu 80%:20%. Data latih untuk menghasilkan model klasifikasi untuk kebutuhan prediksi, dan data uji untuk melakukan pengujian terhadap model klasifikasi.

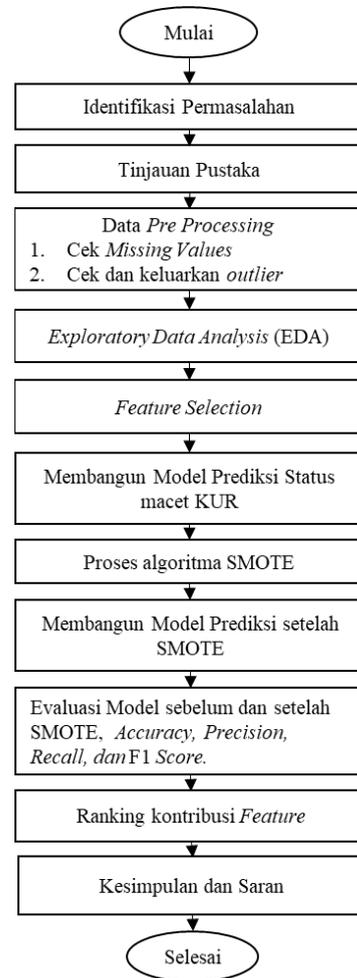
Feature selection, dilakukan untuk memilih fitur yang relevan dan menghilangkan atribut yang tidak perlu. *Feature selection* dilakukan dengan metode filter yaitu penerapan Lasso (*Least Absolute Shrinkage and Selection Operator*) dan menghasilkan 7 *feature* yaitu Unit_Processing_JEMBATAN MERAH, Unit_Processing_MADURA, Unit_Processing_GRAHA_PANGERAN, Sektor_Ek_Industri_Pengolahan, Sektor_Ek_Pertambangan, DSPA_TERSEDIA, dan R3_RESTRUK.

Selanjutnya dilakukan pemodelan dengan menggunakan metode regresi logistik dan *random forest*. Metode tersebut dipilih karena merupakan model klasifikasi yang dinilai memiliki akurasi yang tinggi sesuai dengan data yang diperoleh dari penelitian terdahulu yang dijadikan pedoman dalam penelitian ini.

Model selanjutnya dievaluasi berdasarkan *confusion matrix* dengan beberapa parameter yaitu *accuracy*, *precision*, *recall*, dan *F1 score*. *Accuracy* mewakili proporsi data yang diklasifikasikan dengan benar dari jumlah total data, *Precision* menggambarkan tingkat akurasi prediksi kelas positif (*default*) sebagai kelas positif (*default*). *Recall*, menunjukkan seberapa akurat model dapat memprediksi kelas positif dan disebut juga sensitivitas. *F1 score* adalah metrik yang mempertimbangkan *precision* dan *recall*. *F1 score* yang lebih tinggi menunjukkan *precision* dan *recall* yang lebih kuat, sehingga memperkuat keunggulannya dibandingkan pengukuran *accuracy*.

Untuk mencegah ketidakseimbangan dataset dan meningkatkan akurasi model klasifikasi, teknik SMOTE dilakukan untuk membuat dataset seimbang. SMOTE dilakukan pada data *training*, selanjutnya model yang dibangun sebelum dan sesudah SMOTE dievaluasi, dan model dengan nilai *recall* tertinggi dipilih. Perlu dilakukan SMOTE pada penelitian ini karena dataset tidak seimbang yaitu kelas macet sebanyak 427 dan kelas tidak macet sebanyak 45.511.

Model ini dibedah lebih lanjut untuk menentukan *feature importance* setiap fitur. Karakteristik signifikan tersebut kemudian digunakan untuk memberikan rekomendasi bagi perusahaan dalam memitigasi risiko saat menganalisis usulan KUR guna mencegah kredit macet pada KUR.



GAMBAR 1. Diagram Alir Penelitian.

IV. HASIL DAN PEMBAHASAN

Bagian ini menjelaskan secara rinci hasil penelitian, termasuk data dan temuan penting, dan menjawab pertanyaan dalam perumusan masalah terkait penelitian.

A. MODEL PREDIKSI SEBELUM DILAKUKAN SMOTE

Untuk mengetahui efektivitas dan kinerja model prediktif perlu dilakukan evaluasi atas model tersebut dengan menggunakan berbagai metrik atau estimator. *Precision* menunjukkan akurasi prediksi kasus positif dibandingkan dengan total prediksi kelas positif, *recall* menunjukkan akurasi prediksi kasus positif dibandingkan dengan jumlah total kelas positif terjadi, *accuracy* menunjukkan tingkat akurasi kedua kelas diprediksi dibandingkan dengan semua catatan, akan menjadi ukuran yang baik ketika kumpulan data seimbang, *F1-score* yaitu mean harmonik dari *precision* dan *recall*. Performa model prediksi, seperti terlihat pada tabel 2 model klasifikasi *random forest* menunjukkan tingkat kinerja yang lebih tinggi dibandingkan model regresi logistik.

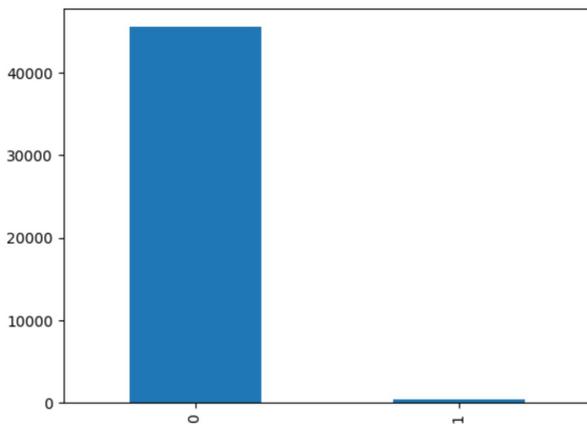
TABEL II
EVALUASI MODEL PREDIKSI SEBELUM SMOTE

Metrik	Regresi Logistik	Random Forest
Accuracy	0.99	0.99
Precision	0.00	0.60
Recall	0.00	0.07
F1-score	0.00	0.13

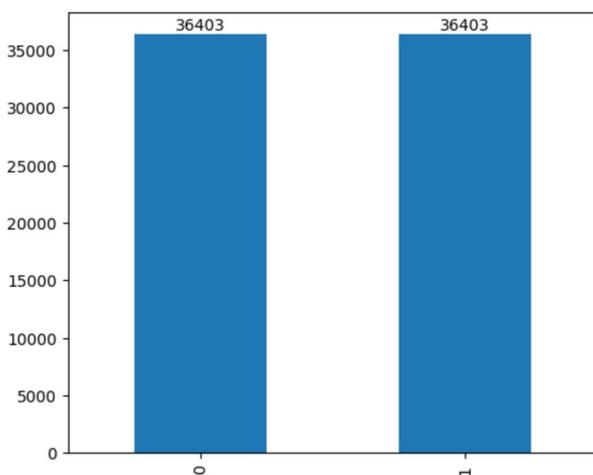
B. MODEL PREDIKSI SETELAH DILAKUKAN SMOTE

Untuk mencegah ketidakseimbangan dataset dan meningkatkan akurasi model klasifikasi, teknik SMOTE digunakan dalam model prediksi untuk membuat dataset seimbang dengan menambahkan data ke kelas minor.

Gambar 2 dan gambar 3 menggambarkan variabel respon sebelum dan sesudah dilakukan SMOTE. Pada saat dilakukan kembali model prediksi hasil yang diharapkan tidak bias karena perbedaan antar kelas klasifikasi yang terlalu jauh.



GAMBAR 2. KUR Macet Sebelum SMOTE.



GAMBAR 3. KUR Macet Setelah SMOTE.

Setelah melakukan algoritma SMOTE, model klasifikasi dibuat ulang menggunakan regresi logistik dan *random forest*, sehingga menghasilkan *confusion matrix* berikutnya

dengan *accuracy*, *precision*, *recall*, dan *F1 score* baru yang ditunjukkan pada tabel 3.

TABEL III
EVALUASI MODEL PREDIKSI SETELAH SMOTE

Metrik	Regresi Logistik	Random Forest
Accuracy	0.80	0.88
Precision	0.03	0.06
Recall	0.75	0.86
F1-score	0.06	0.11

Meskipun terjadi penurunan *accuracy*, *precision*, *recall*, dan *F1 score* setelah dilakukan SMOTE, perlu dilihat bahwa algoritma *random forest* terus menunjukkan kinerja yang lebih unggul dibandingkan dengan regresi logistik. Sehingga untuk model prediksi yang dipilih adalah model dari *random forest*. Selanjutnya dilakukan evaluasi untuk model dengan metode *random forest* yaitu analisis perbandingan model *random forest* sebelum dan sesudah penerapan teknik SMOTE pada kelas 1 (macet).

TABEL IV
EVALUASI *RANDOM FOREST*: MODEL PREDIKSI SEBELUM DAN SETELAH SMOTE UNTUK KELAS 1 (MACET)

Metrik	Random Forest sebelum SMOTE	Random Forest setelah SMOTE
Accuracy	0.99	0.88
Precision	0.60	0.06
Recall	0.07	0.86
F1-score	0.13	0.11

Berdasarkan tabel 4, terlihat bahwa penggunaan SMOTE menyebabkan penurunan nilai *accuracy*, *precision*, dan *F1 score*, sedangkan nilai *recall* mengalami peningkatan. *Recall* atau sensitivitas mengukur kapasitas model untuk secara efektif mengurangi kejadian *false negative*, yang mengacu pada prediksi macet pada KUR yang benar-benar macet namun secara keliru diklasifikasikan sebagai tidak macet oleh model atau diprediksi sebagai kredit yang baik. Hal ini penting bagi Bank XYZ karena adanya peningkatan biaya pencadangan kerugian yang akan mengurangi keuntungan bank.

Penurunan akurasi pada model setelah dilakukan SMOTE, karena pada data yang tidak seimbang memungkinkan model akan memilih prediksi pada kelas mayoritas dimana dalam kasus kredit macet, kelas mayoritas adalah kredit lancar. Nilai *recall* memiliki peningkatan yang signifikan setelah dilakukan SMOTE sehingga model prediksi yang dipilih adalah *random forest* setelah SMOTE.

C. FEATURE IMPORTANCE

Melakukan interpretasi klasifikasi dalam penelitian sebelumnya dengan menggunakan metode *feature importance* untuk menghitung peringkat kontribusi dari masing-masing variabel prediktor terhadap variabel respon.

TABEL V
 FEATURE IMPORTANCE DARI VARIABEL PREDIKSI

	Feature	Importance
0	R3 RESTRUK	0,36
1	DSPA TERSEDIA	0,24
2	Unit Processing MADURA	0,21
3	Sektor_Ek_Industri Pengolahan	0,15

Fitur yang berulang kali menunjukkan kecenderungan untuk mengurangi ketidakmurnian dan meningkatkan performa model akan diberi skor relevansi yang lebih tinggi, sehingga menandakan signifikansinya dalam menghasilkan prediksi yang akurat. Tabel 5 menampilkan kontribusi dari 4 karakteristik teratas yang memiliki pengaruh paling besar, dan mewakili 96% dari keseluruhan fitur. Fitur yang paling berpengaruh terhadap gagal bayar/macet KUR adalah restrukturisasi kredit, ketersediaan *debt service payment account*, unit pemroses kredit dari Madura, dan sektor usaha dari sektor industri pengolahan.

V. KESIMPULAN

Setelah menggunakan teknik SMOTE, algoritma *random forest* konsisten menunjukkan kinerja yang lebih unggul dibandingkan dengan regresi logistik. Pengujian data performa untuk memprediksi kelas 1 (macet) nilai *recall* meningkat setelah SMOTE yang menunjukkan kapasitas model secara efektif mengurangi kejadian *false negative*. Hasil dari model terpilih dapat menjadi input untuk Bank XYZ sebagai pertimbangan dalam memprediksi kredit macet pada KUR.

Fitur dengan kontribusi tertinggi berdasarkan hasil *feature important* yang paling besar pengaruhnya terhadap macet pada KUR adalah status restrukturisasi kredit, ketersediaan *debt service payment account*, unit pemroses kredit dari Madura, dan sektor usaha dari sektor industri pengolahan dapat dijadikan pertimbangan oleh Bank XYZ untuk mengkaji bobot risiko saat ini.

Untuk mengurangi NPL *rate* dalam ekspansi KUR baru yang berkualitas tindakan mitigasi yang dapat dilakukan oleh Bank XYZ adalah pemberian bobot risiko yang lebih tinggi dalam hal aplikasi permohonan KUR yang memiliki historikal restrukturisasi kredit pada bank lain, sektor ekonomi industri pengolahan, dan unit proses Madura.

Untuk penelitian kedepannya, data secara nasional dapat ditambahkan sehingga dapat meningkatkan akurasi prediksi kredit macet pada Bank XYZ.

PERAN PENULIS

Wulansari: Analisis Formal, Pengumpulan dataset, Metodologi, Administrasi Proyek, Perangkat Lunak, Validasi, Visualisasi, Penulisan Penyusunan Draf Asli.

Diana Purwitasari: Penulisan Review & Penyuntingan;

COPYRIGHT



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

DAFTAR PUSTAKA

- [1] Bank Indonesia, "Undang-Undang Republik Indonesia Nomor 10 Tahun 1998 Tentang Perbankan." Jakarta, 1998.
- [2] Menteri Koordinator Bidang Perekonomian, "Peraturan Menteri Koordinator Bidang Perekonomian Nomor 1 Tahun 2022 tentang Pedoman Pelaksanaan Kredit Usaha Rakyat." Jakarta, 2022.
- [3] S. B. Haryanto and E. T. Widyarti, "Analisis Pengaruh NIM, NPL, BOPO, BI Rate dan CAR terhadap Penyaluran Kredit Bank Umum Go Public Periode Tahun 2012-2016," vol. 6, no. Nomor 4, pp. 1–11, 2017.
- [4] M. Madaan, A. Kumar, C. Keshri, R. Jain, and P. Nagrath, "Loan default prediction using decision trees and random forest: A comparative study," *IOP Conf Ser Mater Sci Eng*, vol. 1022, no. 1, p. 012042, Jan. 2021, doi: 10.1088/1757-899X/1022/1/012042.
- [5] Q. Zhang, "Loan Risk Prediction Model based on Random Forest," *Advances in Economics, Management and Political Sciences*, vol. 5, no. 1, pp. 216–222, Apr. 2023, doi: 10.54254/2754-1169/5/20220082.
- [6] J. Sanjaya, E. Renata, V. E. Budiman, F. Anderson, and M. Ayub, "Prediksi Kelalaian Pinjaman Bank Menggunakan Random Forest dan Adaptive Boosting," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 6, no. 1, Apr. 2020, doi: 10.28932/Jutisi.v6i1.2313.
- [7] L. Zhu, D. Qiu, D. Ergu, C. Ying, and K. Liu, "A study on predicting loan default based on the random forest algorithm," *Procedia Comput Sci*, vol. 162, pp. 503–513, 2019, doi: 10.1016/j.procs.2019.12.017.
- [8] M. K. Pandey, M. Mittal, and K. Subbiah, "Optimal balancing & efficient feature ranking approach to minimize credit risk," *International Journal of Information Management Data Insights*, vol. 1, no. 2, Nov. 2021, doi: 10.1016/j.ijime.2021.100037.
- [9] L. Zhou and H. Wang, "Loan Default Prediction on Large Imbalanced Data Using Random Forests," *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 10, no. 6, Sep. 2012, doi: 10.11591/Telkonnika.v10i6.1323.
- [10] E. Widodo, O. A. Rochmadhan, Lukmandono, and Januardi, "Modeling Bayesian inspection game for non-performing loan problems," *Operations Research Perspectives*, vol. 9, p. 100218, 2022, doi: 10.1016/j.orp.2021.100218.
- [11] U. Aslam, H. I. T. Aziz, A. Sohail, and N. K. Batcha, "An empirical study on loan default prediction models," *J Comput Theor Nanosci*, vol. 16, no. 8, pp. 3483–3488, 2019, doi: 10.1166/jctn.2019.8312.
- [12] A. K. Sharma, L.-H. Li, and R. Ahmad, "Default Risk Prediction Using Random Forest and XGBoosting Classifier," 2023, pp. 91–101. doi: 10.1007/978-3-031-05491-4_10.
- [13] B. Prasajo and E. Haryatmi, "Analisa Prediksi Kelayakan Pemberian Kredit Pinjaman dengan Metode Random Forest," *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 7, no. 2, pp. 79–89, Sep. 2021, doi: 10.25077/TEKNOSI.v7i2.2021.79-89.
- [14] M. I. Putra, A. Yusuf, and N. Yalina, "Klasifikasi Kelancaran Kredit Dengan Metode Random Forest," *Systemic: Information System and Informatics Journal*, vol. 5, no. 2, pp. 7–12, Mar. 2020, doi: 10.29080/Systemic.v5i2.713.
- [15] G. K. Rajbahadur, S. Wang, Y. Kamei, and A. E. Hassan, "The impact of feature importance methods on the interpretation of defect classifiers," Feb. 2022, doi: 10.1109/TSE.2021.3056941.

Sistem Deteksi dan Klasifikasi Truk Air Menggunakan YOLO v5 dan EfficientNet-B4

Ardian K. L. Santoso¹ dan Daniel M. Wonohadidjojo¹

¹Program Studi Informatika, Universitas Ciputra Surabaya, Surabaya, Indonesia

Corresponding author: Daniel M. Wonohadidjojo (e-mail: daniel.m.w@ciputra.ac.id).

ABSTRACT Recording activities in the water filling business carried out using water trucks experienced problems due to human error, for example missed recording and the efficiency of the time required. This requires system automation using technology. To overcome this problem, in this research a method in Computer Vision is used, by using Object Detection and Classification algorithms. In this research, a system was built that takes video frames using CCTV which is then inserted into the object detection algorithm with the YOLOv5 (You Only Look Once version 5) architecture. The detection results are then classified using an algorithm with the EfficientNet-B4 architecture. The results of this classification determine specifically which water trucks are filling and are recorded. Then a recapitulation of the recording results is sent using the Telegram messaging application using Tkinter library to the business owner who collects the water. This recapitulation is then used by the business owner to monitor the business and make payments according to the amount of water taken. Test results for object detection and classification using evaluation metrics show that the method is successful in detecting and classifying. Testing of the entire system shows that all stages of testing were successfully carried out well. This shows that the system can be used to overcome the problems faced.

KEYWORDS CCTV, EfficientNetB4, Telegram, YOLOv5

ABSTRAK Kegiatan pencatatan dalam usaha pengisian air yang dilakukan dengan menggunakan truk air mengalami masalah karena kesalahan manusia (human error) misalnya pencatatan yang terlewat dan efisiensi waktu yang diperlukan. Untuk itu diperlukan otomatisasi sistem dengan menggunakan teknologi. Untuk mengatasi masalah tersebut, pada penelitian ini digunakan metode yang termasuk dalam Computer Vision dengan penggunaan algoritma Object Detection dan Classification. Pada penelitian ini dibangun suatu sistem yang mengambil frame video menggunakan CCTV yang kemudian dimasukkan pada algoritma object detection dengan arsitektur YOLOv5 (You Only Look Once versi 5). Hasil deteksi kemudian diklasifikasikan dengan menggunakan algoritma dengan arsitektur EfficientNet-B4. Hasil klasifikasi tersebut akan menentukan secara spesifik truk air yang mana yang sedang melakukan pengisian dan dicatat. Kemudian rekapitulasi hasil pencatatan tersebut dikirimkan dengan menggunakan aplikasi messaging Telegram menggunakan library Tkinter kepada pemilik usaha yang mengambil air tersebut. Rekapitulasi tersebut kemudian digunakan oleh sang pemilik usaha dalam memantau usaha tersebut dan melakukan pembayaran sesuai dengan jumlah pengambilan air. Hasil pengujian untuk object detection dan classification dengan menggunakan evaluation metrics menunjukkan bahwa metode tersebut berhasil melakukan deteksi dan klasifikasi dengan baik. Pengujian keseluruhan sistem menunjukkan bahwa semua tahap pengujian berhasil dilakukan dengan baik. Hal ini menunjukkan bahwa sistem tersebut dapat digunakan untuk mengatasi masalah yang dihadapi.

KATA KUNCI CCTV, EfficientNetB4, Telegram, YOLOv5

I. PENDAHULUAN

Usaha pengisian air merupakan salah satu usaha yang biasanya dilakukan oleh warga di Indonesia. Pemilik usaha menjelaskan bahwa kegiatannya berupa truk air

yang datang ke suatu tempat yang menjadi sumber pengambilan air dan biasanya dimiliki oleh suatu perusahaan besar maupun usaha perorangan. Di dalam proses kegiatan usaha tersebut terdapat pencatatan berapa kali truk-truk

mengambil air dalam sehari. Pencatatan tersebut dilakukan karena penyedia sumber air akan menerima pembayaran dari pihak pengambil air. Pencatatan tersebut biasanya dilakukan secara konvensional terutama bagi usaha perseorangan yang skala usahanya lebih kecil dari perusahaan. Pencatatan atau penghitungan biasanya dilakukan secara manual dengan sistem laporan dan seringkali mengalami masalah seperti ada pencatatan yang terlewat.

Pemilik bisnis menyebutkan bahwa alternatif solusi yang lain adalah dengan mengambil gambar atau foto bukti bahwa ada pengambilan air pada waktu tersebut agar di kemudian hari dapat dilakukan penagihan terhadap perusahaan truk tangki air beserta bukti berupa foto. Namun cara ini memerlukan waktu yang panjang karena menunggu keluar masuknya truk air hanya untuk menunggu foto. Selain itu juga rentan terhadap terjadinya human error pada saat mencatat truk air yang keluar dan masuk.

Saat ini sudah banyak digunakan sistem kamera CCTV di tempat umum, dalam bisnis, perusahaan, jalan, sekolah, dan bahkan di sepanjang jalur jalan raya umum [1]. Dalam hal itu CCTV dapat dimanfaatkan bagi pemilik bisnis untuk membantu dalam pencatatan atau pengarsipan aktivitas bisnis. Akan tetapi pencatatan tersebut menjadi tidak efisien karena pemilik bisnis harus melihat rekaman CCTV berdurasi 24 jam yang juga mengharuskan pemilik bisnis memeriksa rekaman tersebut secara mendetail.

Berdasarkan latar belakang masalah tersebut, maka pada penelitian ini diusulkan suatu metode untuk mendeteksi truk air yang masuk dan keluar secara otomatis dengan menggunakan bantuan teknologi. Teknologi yang dimaksud adalah computer vision dengan memanfaatkan CCTV yang dilengkapi dengan kemampuan object detection dan classification. Dalam klasifikasi citra, tugas yang dilakukan adalah menetapkan suatu citra tertentu ke dalam suatu kelas, yang mana kelas citra tersebut bergantung pada jenis objek yang direpresentasikan pada citra tersebut. Untuk beberapa tahun, metode paling sukses dalam aplikasi dunia nyata didasarkan pada Convolutional Neural Network (CNN) [2].

Pada metode ini, truk air tersebut dideteksi dengan menggunakan algoritma YOLOv5 dan kemudian diklasifikasikan dengan algoritma EfficientNet-B4. Algoritma YOLOv5 dipilih karena lebih cepat dan lebih akurat dalam mengidentifikasi item dalam foto, membuatnya ideal untuk digunakan dengan real-time object dalam bentuk video [3]. Selain itu YOLOv5 adalah model yang lebih cepat, lebih terukur, dan lebih ringan dibandingkan pesaing lainnya [4]. Sedangkan EfficientNet dipilih karena efektivitas metode ini dalam meningkatkan MobileNets dan ResNet [5]. Pada penelitian tersebut digunakan penelusuran arsitektur saraf untuk merancang jaringan dasar baru dan meningkatkannya untuk mendapatkan serangkaian model, yang disebut EfficientNets, yang mencapai akurasi dan efisiensi jauh lebih baik daripada ConvNets sebelumnya. EfficientNet juga digunakan pada penelitian lainnya disebabkan oleh akurasi kerjanya yang tinggi meskipun ukurannya kecil [6].

Setelah melakukan deteksi dan klasifikasi, selanjutnya sistem pada penelitian ini mengirimkan bukti informasi berupa gambar dan teks yang mengandung nama sopir truk air yang terdeteksi. Kemudian dilanjutkan dengan pengiriman informasi rekapitulasi harian di malam hari yang memuat informasi jumlah masing-masing dan keseluruhan truk air yang masuk beserta jumlah total tagihan ke perangkat smartphone pemilik bisnis menggunakan aplikasi Telegram agar pemilik bisnis dapat lebih mudah memeriksa aktivitas bisnis. Aplikasi Telegram sendiri adalah aplikasi pesan berbasis cloud namun menyediakan bot-bot tertentu yang dapat digunakan untuk perangkat IoT [7].

II. TINJAUAN PUSTAKA

Pada penelitian [8] dibuat sistem pendeteksi objek memakai metode YOLOv5 untuk mendeteksi jenis kendaraan di jalan raya. Pada penelitian tersebut digunakan dataset sebanyak 1332 gambar dengan kelas bajaj, becak, bus, mobil, mobil molen, mobil pikup, sepeda, sepeda motor, dan truk. Penelitian ini dapat mengenali objek secara konsisten dengan tingkat akurasi 90%.

Selain itu terdapat penelitian yang bertujuan untuk mengembangkan suatu mekanikal sistem seperti robot dengan tujuan untuk menjadi lebih baik dengan meningkatkan kapabilitas untuk pengenalan citra, yang nantinya juga akan meningkatkan akurasi objek deteksi dari robot yang dibuat untuk membantu di berbagai macam bidang [9]. Penelitian ini menggunakan 4 versi dari YOLOv5 yaitu YOLOv5s, YOLOv5m, YOLOv5l, dan YOLOv5x.

Pada bidang image classification terdapat penelitian yang mengeksplorasi penggunaan arsitektur jaringan saraf konvolusi mendalam (CNN) untuk klasifikasi kendaraan bervolume berlebih. Studi ini membandingkan kinerja arsitektur CNN yang berbeda, termasuk ResNet50, InceptionV3, dan EfficientNetB4, dengan menggunakan strategi pelatihan dan teknik augmentasi data yang berbeda. Studi ini menemukan bahwa EfficientNet-B4 mengungguli model-model lainnya, dan mencapai akurasi 99,91% pada tugas klasifikasi. Kinerja EfficientNet-B4 yang unggul disebabkan oleh penggunaan network parameter dan metode scaling yang efisien [10]. Secara keseluruhan, penelitian ini menyoroti potensi arsitektur CNN yang mendalam untuk mengklasifikasikan kendaraan dengan volume berlebih secara akurat, dengan EfficientNet-B4 sebagai model yang paling efektif di antara model-model yang dievaluasi.

Dalam penelitian selanjutnya, dilakukan identifikasi pasien COVID-19 menggunakan gambar sinar-X dada dan transfer-based learning. Pada penelitian tersebut dievaluasi berbagai model pre-trained deep neural network sebelumnya, termasuk Sequential, DenseNet121, ResNet152, dan EfficientNet-B4, untuk menentukan hasil yang paling efisien. Hasil penelitian menunjukkan EfficientNet-B4 memiliki area under curve (AUC) yang tertinggi yaitu 0,997, diikuti oleh ResNet-50 sebesar 0,967, DenseNet121 sebesar 0,874, dan Sequential sebesar 0,762. EfficientNet-B4 juga mencapai akurasi tertinggi pada 98,86%. Hasil ini menunjukkan bahwa EfficientNet-B4

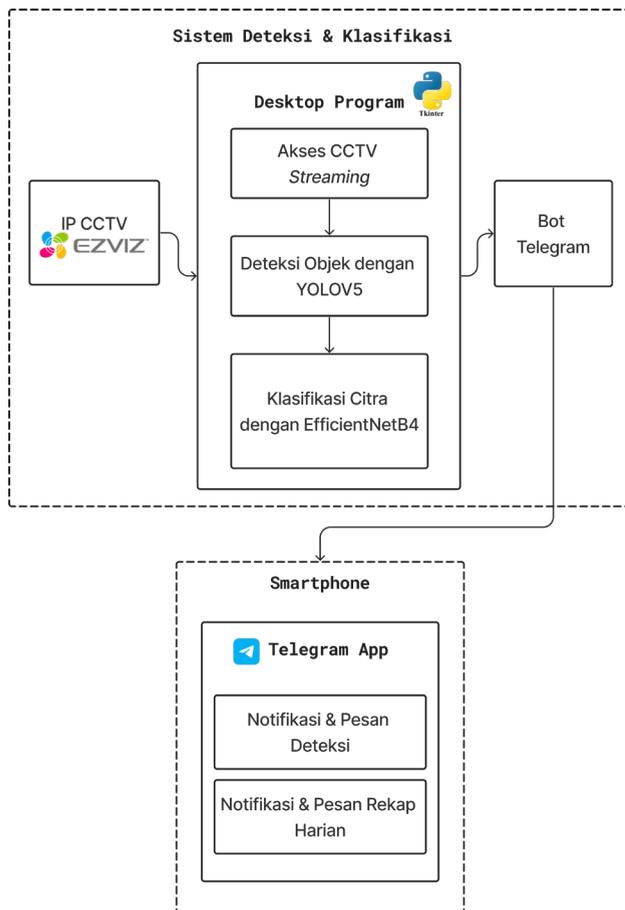
adalah model yang menjanjikan untuk mengidentifikasi pasien COVID-19 dari gambar rontgen dada [11].

Untuk pengukuran kinerja, pada penelitian sebelumnya digunakan mAP, Precision dan Recall untuk mengevaluasi kinerja YOLOv5 dibandingkan dengan model yang telah ditingkatkan [12].

III. METODE

A. TAHAPAN PENELITIAN

Penelitian sistem deteksi dan klasifikasi truk air dengan CCTV ini dilakukan dengan menggunakan beberapa tahap yang terdiri dari: pengumpulan dataset, anotasi data, pembuatan model YOLOv5, pembuatan model EfficientNetB4, iterasi model klasifikasi dan deteksi, dan implementasi sistem. Sebelum dilakukan tahap-tahap penelitian tersebut, dirancang terlebih dahulu arsitektur sistem yang akan dibangun. Arsitektur sistem tersebut ditunjukkan pada Gambar 1.



GAMBAR 1. Arsitektur sistem.

1) PENGUMPULAN DATASET

Dataset diambil dari Kamera IP CCTV EZVIZ yang menghadap ke jalan raya dengan pencahayaan yang cenderung mengalami backlight dengan jumlah sebanyak

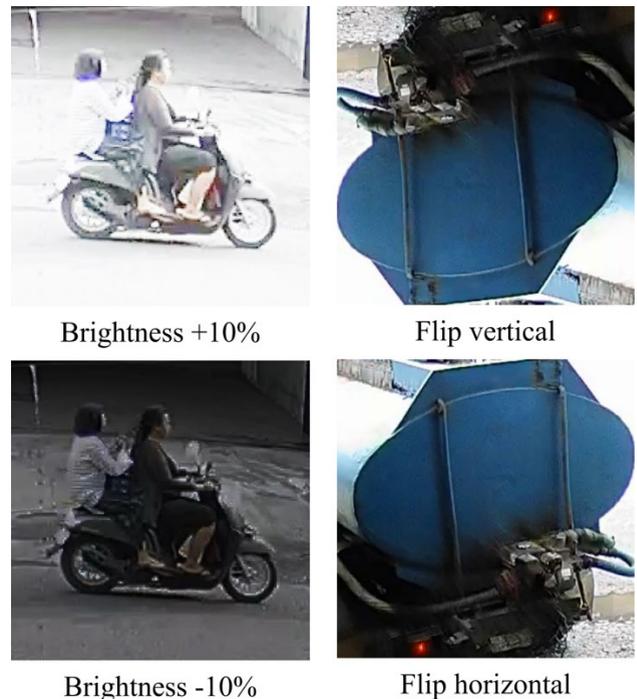
1596 yang memuat gambar 4 jenis truk air, orang, sepeda motor, mobil.

2) ANOTASI DATA

- DATASET MODEL DETEKSI

Pada proses deteksi, untuk membantu memvalidasi keefektifan sistem dalam kondisi yang berbeda maka dilakukan peningkatan jumlah dan ragam dataset. Hal ini dilakukan dengan proses anotasi dan pelabelan dataset dengan menggunakan platform Roboflow. Pada tahap ini dilakukan anotasi dan pelabelan gambar, implementasi pra-pemrosesan, augmentasi, untuk menghasilkan dataset dalam format yang diperlukan pada penelitian ini. Dataset yang didapatkan dari Roboflow ini dibagi menjadi dua jenis dataset yaitu dataset untuk deteksi objek dan klasifikasi.

Pada tahap ini, untuk dataset YOLOv5 dilakukan anotasi, augmentasi, dan eliminasi. Augmentasi dataset yang dilakukan adalah membuat dengan lebih terang dan lebih gelap sebanyak 10% lalu mengubah orientasi gambar secara *vertical* dan *horizontal* dengan contoh yang ditunjukkan pada Gambar 2.



GAMBAR 2. Augmentasi dataset untuk model deteksi

Untuk eliminasi yang dilakukan adalah menghilangkan gambar-gambar dengan objek yang terlalu gelap, objek yang terlalu blur dan objek yang terhalangi 50% dengan contoh yang ditunjukkan pada Gambar 3.



GAMBAR 3. Eliminasi dataset untuk model deteksi

Dari hasil anotasi, augmentasi dan eliminasi didapatkan dataset yang terdiri dari 1835 gambar yang terdiri dari 152 gambar 'car', 967 gambar 'motorcycle', 582 gambar 'person', dan 1207 gambar 'water-truck'.

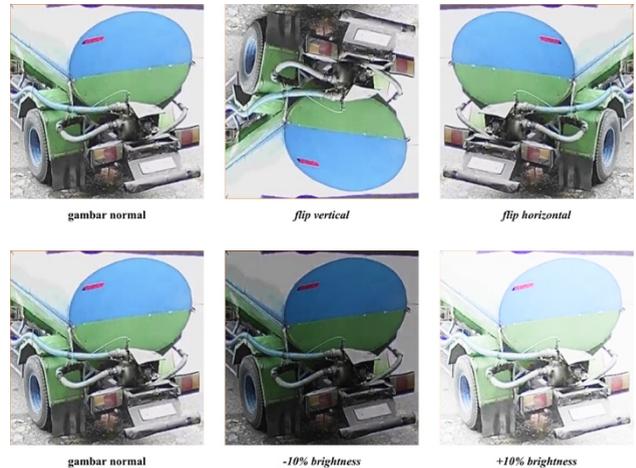
• DATASET MODEL KLASIFIKASI

Pada proses klasifikasi, dataset model EfficientNet-B4 sendiri berjumlah 512 gambar yang dibagi menjadi 4 kelas yaitu 'water_truck_bagus' sebanyak 144 gambar, 'water_truck_dasor' sebanyak 156 gambar, 'water_truck_parman' sebanyak 103 gambar, dan 'water_truck_sugeng' sebanyak 109 gambar dengan contoh yang ditunjukkan pada Gambar 4.



GAMBAR 4. Kelas dataset untuk model klasifikasi

Untuk menghindari *overfitting*, pada dataset ini juga dilakukan augmentasi dengan cara yang sama seperti pada dataset model deteksi yaitu *flip horizontal & vertical* dan membuat gambar lebih terang dan lebih gelap dengan *brightness level -10% sampai +10%* dengan contoh yang ditunjukkan pada Gambar 5.



GAMBAR 5. Augmentasi dataset untuk model klasifikasi

Setelah dilakukannya proses augmentasi jumlah total dataset menjadi 4359 lalu dibagi menjadi 3803 gambar untuk *training*, 370 gambar untuk *validation*, dan 186 gambar untuk *testing*.

3) PEMBUATAN MODEL YOLOV5

Model YOLOv5 dilatih melalui transfer learning dengan initial weights menggunakan YOLOv5s.pt, 100 epochs, 16 batch_size, dengan input image sebesar 800x800. Training model YOLOv5 ini juga mengaktifkan Cache untuk mempercepat pelatihan secara signifikan. Cache dalam konteks pelatihan ini berfungsi menyimpan gambar pelatihan yang sudah diproses sebelumnya dalam memori, yang mengurangi waktu yang diperlukan untuk memuat dan memproses gambar selama setiap epoch.

4) PEMBUATAN MODEL EFFICIENTNET-B4

EfficientNet-B4 adalah pre-trained model, yang telah diinisialisasi dengan bobot ImageNet dan tidak termasuk layer klasifikasi teratas. Model EfficientNet-B4 pada penelitian ini terdiri dari 'Input Layer' dengan resolusi 380x380, 'trainable_layer' yang sudah di-frozen, layer 'global_average_pooling2D', Fully connected Dense layer dengan 64 unit dan aktivasi Rectified Linear Unit (ReLU), Regularization L2 dengan kekuatan 0.01, layer dropout dengan tingkat 0,5, di layer terakhir terdapat layer Dense dengan 4 unit output dan aktivasi softmax.

5) ITERASI MODEL

Pada tahap iterasi ini, dilakukan evaluasi kinerja model dengan menggunakan performance metrics yang sesuai. Untuk model YOLOv5 digunakan beberapa metric yaitu precision, recall, dan mAP, sedangkan untuk model EfficientNetB4 digunakan metric accuracy, loss, precision, recall, dan F1 Score.

6) IMPLEMENTASI

Setelah pada kedua model CNN dilakukan pengujian, maka model tersebut diimplementasikan dengan menggunakan kode python yang dapat mengakses kamera IP CCTV dan telah disambungkan dengan bot Telegram.

B. RANCANGAN PENGUJIAN

Tujuan pengujian yang dilakukan pada penelitian ini adalah untuk memastikan bahwa sistem deteksi dapat menjalankan semua fitur dan dapat memprediksi secara benar sesuai dengan alur dari arsitektur sistem pada Gambar 1. Tahap pengujian pada penelitian ini terdiri dari:

- 1) Ketika program komputer dijalankan akan memunculkan jendela yang berisikan Graphical User Interface (GUI) yang dibuat menggunakan library Tkinter. Pada GUI tersebut terdapat tombol bertuliskan ‘Start’.
- 2) Ketika pengguna menekan tombol ‘Start’, program akan langsung memulai sistem deteksi. Program tersebut akan mengakses kamera IP CCTV lalu memunculkan jendela baru yang berisi real time video yang bersumber dari kamera IP CCTV. Real time video tersebut juga menampilkan real time object detection yang ditandai dengan munculnya bounding box dan confidence level dari objek yang berhasil dideteksi oleh sistem dengan menggunakan model YOLOv5.
- 3) Ketika sistem deteksi berhasil mendeteksi truk air dengan confidence level sama dengan atau lebih besar dari 0.80 maka sistem langsung melakukan frame capture dan memotong sesuai dengan region of interest yang dihasilkan bounding box di sekitar truk air. Lalu hasil foto yang sudah dipotong tersebut akan diklasifikasikan secara benar oleh model klasifikasi EfficientNet-B4. Selanjutnya sistem akan mengirim pesan berupa informasi hasil deteksi beserta bukti foto truk air hasil dari frame capture ke aplikasi Telegram pemilik bisnis.
- 4) Pada akhir hari sistem akan mengirimkan pesan berupa rekap harian hasil deteksi yang mengandung informasi akumulasi truk air yang mengambil air dan akumulasi tagihan sesuai dengan supir truk air tersebut.

IV. HASIL DAN PEMBAHASAN

A. PENGUJIAN MODEL YOLOV5

Hasil training model YOLOv5 yang menggunakan hyperparameter yang telah ditetapkan dengan menggunakan total 362 gambar adalah Precision sebesar 86%, Recall sebesar 87%, dan mAP 90%. Untuk data yang lebih lengkap dapat dilihat data pada tabel 1.

TABEL I
HASIL PENGUJIAN MODEL YOLOV5

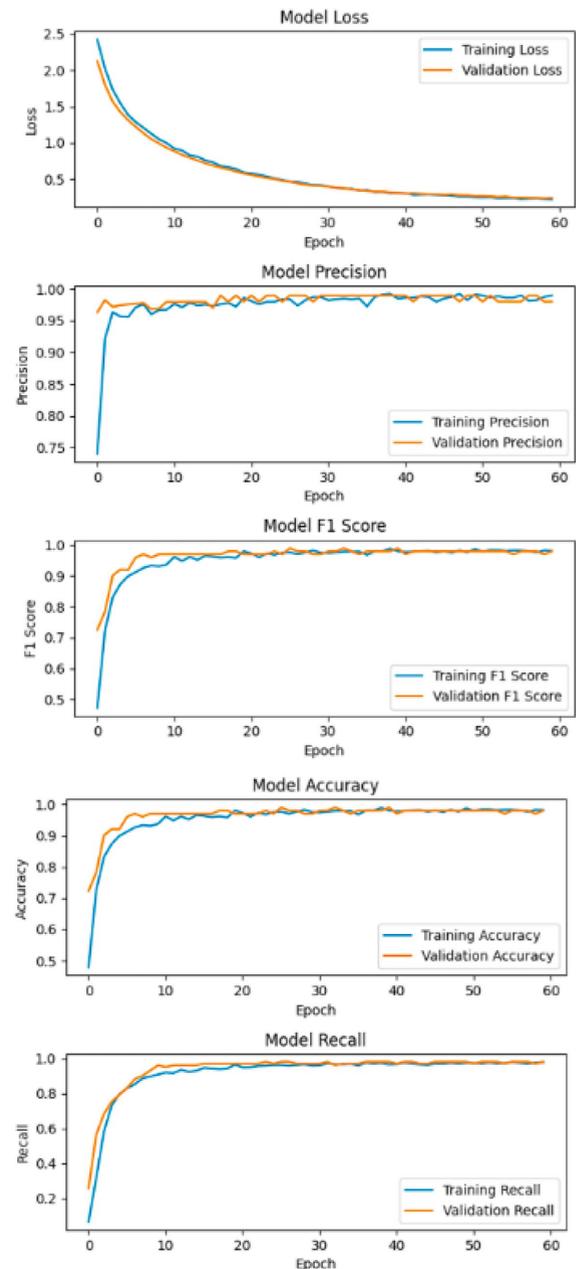
Kelas	Instances	Precision	Recall	mAP
Car	31	94%	97%	96%
Motorcycle	232	82%	79%	85%

Person	145	78%	78%	80%
Water-truck	239	92%	97%	98%
All	647	86%	87%	90%

Tabel 1 menunjukkan bahwa hasil deteksi untuk 4 kelas yaitu kelas ‘Car’, ‘Motorcycle’, ‘Person’ dan ‘Water-truck’. Pada tiap kelas tersebut didapatkan hasil pengujian untuk metrics Precision, Recall dan mAP. Semua metrics menunjukkan angka di atas 80% kecuali untuk ‘Person’.

B. PENGUJIAN MODEL EFFICIENTNETB4

Hasil training model klasifikasi EfficientNetB4 yang menggunakan Adam optimizer, 60 epochs, dan 0.0001 learning rate ditunjukkan pada grafik yang terdapat pada Gambar 6.



GAMBAR 6. Grafik hasil pelatihan model EfficientNet-B4

Pada Gambar 6 tersebut ditunjukkan bahwa grafik training untuk Accuracy mengalami kenaikan sampai mencapai nilai stabil mendekati 1 (100%). Sedangkan grafik Loss menunjukkan penurunan sampai mencapai nilai stabil mendekati 0. Pada kedua grafik tersebut, baik training maupun validation menunjukkan konsistensi nilai yang berdekatan selama 60 epochs. Demikian juga untuk grafik Precision, Recall dan F1 Score, grafik ketiganya mengalami kenaikan sampai mendekati nilai 1 (100%). Pada ketiga grafik tersebut juga menunjukkan konsistensi kedekatan nilai antara training dan validation selama 60 epochs. Nilai yang didapatkan untuk semua metrics tersebut ditunjukkan pada tabel 2.

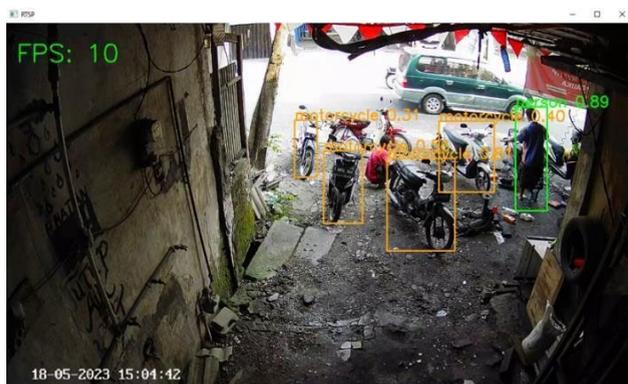
TABEL II
METRIK HASIL PELATIHAN MODEL

Metrik	Training	Validation	Mean
Loss	22%	24%	23%
Accuracy	98%	98%	98%
Precision	99%	98%	99%
Recall	97%	98%	97%
F1 Score	98%	98%	98%

Tabel 2 menyajikan hasil pengujian untuk tiap metric Accuracy, Loss, Precision, Recall dan F1 Score. Hasil tersebut menunjukkan bahwa rata-rata metric Loss, Accuracy, Precision, Recall dan F1Score adalah berturut-turut 23%, 98%, 99%, 97% dan 98%.

C. PENGUJIAN IMPLEMENTASI SISTEM DETEKSI DAN KLASIFIKASI

Pada tahap pengujian pertama, pengguna menekan tombol ‘Start’, dan program berhasil memulai sistem deteksi. Program kemudian mengakses kamera IP CCTV, dan berhasil memunculkan jendela baru yang berisi real time video yang bersumber dari kamera IP CCTV. Video tersebut juga berhasil menampilkan object detection yang ditandai dengan munculnya bounding box dan confidence level dari object yang berhasil dideteksi oleh sistem dengan model YOLOv5. Hal ini ditunjukkan pada Gambar 7. Hasil pengujian ini menunjukkan bahwa pengujian tahap pertama berhasil.



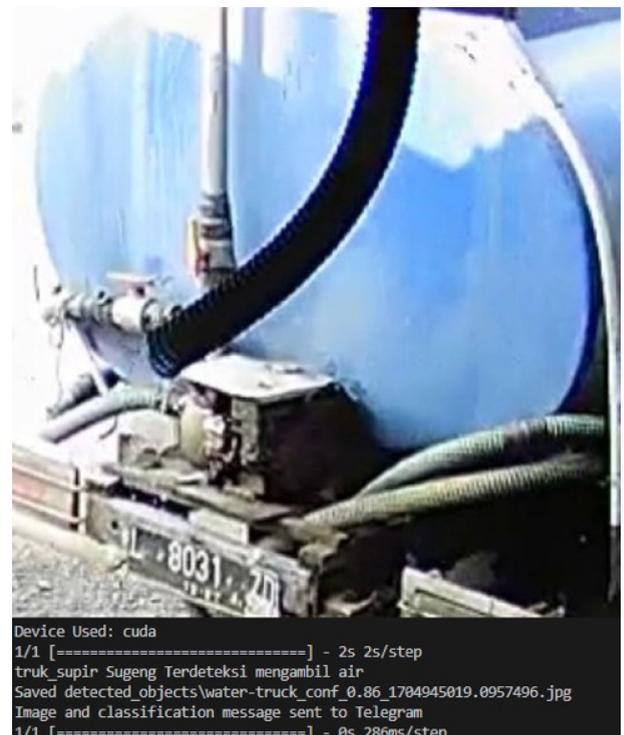
GAMBAR 7. Jendela hasil deteksi dengan bounding box

Pengujian tahap ke kedua dimulai ketika sistem deteksi berhasil mendeteksi truk air dengan confidence level sebesar 0.80 yang ditunjukkan pada Gambar 8.



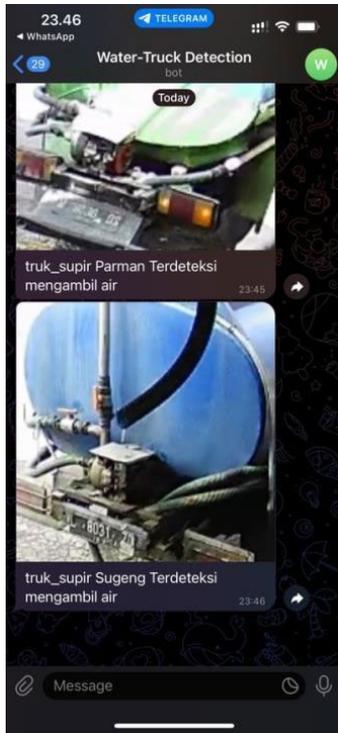
GAMBAR 8. Hasil deteksi truk air dengan confidence level lebih tinggi atau sama dengan 0.80.

Selanjutnya sistem langsung melakukan frame capture dan memotong sesuai dengan region of interest yang dihasilkan bounding box di sekitar truk air. Kemudian hasil foto yang sudah dipotong tersebut berhasil diklasifikasikan secara benar oleh model klasifikasi EfficientNet-B4. Hasil pengujian tahap ini ditunjukkan pada Gambar 9 yang menunjukkan keberhasilan pengujian pada tahap ke kedua.



GAMBAR 8. Model EfficientNetB4 berhasil mengklasifikasi truk air

Pada pengujian tahap ketiga hasil foto yang sudah diklasifikasikan pada tahap kedua, selanjutnya dikirimkan ke aplikasi Telegram pemilik bisnis. Pesan yang dikirimkan berupa informasi hasil deteksi beserta bukti foto truk air hasil dari frame capture. Pesan yang dikirimkan tersebut ditunjukkan pada Gambar 9. Pesan tersebut menunjukkan bahwa pengujian tahap ketiga berhasil.



GAMBAR 10. Pesan berupa informasi hasil deteksi beserta foto

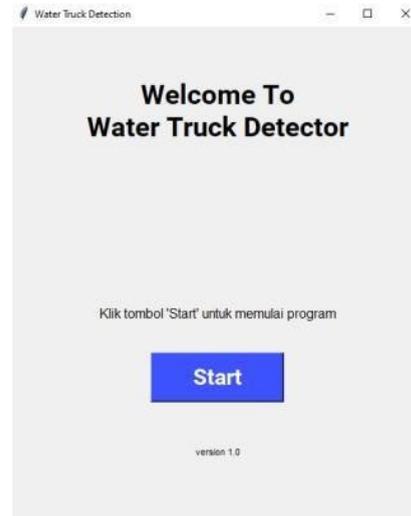
Pada pengujian tahap terakhir yaitu tahap keempat, sistem berhasil mengirimkan pesan berupa rekapitulasi harian hasil deteksi yang berisi informasi akumulasi truk air yang mengambil air dan akumulasi tagihan untuk pengusaha truk air tersebut. Pesan tersebut ditunjukkan pada Gambar 11. Hal ini menunjukkan bahwa pengujian tahap keempat berhasil.



GAMBAR 11. Pesan rekapitulasi di akhir hari yang dikirimkan.

D. IMPLEMENTASI PADA OPERASIONAL BISNIS

Pengimplementasian sistem dalam bisnis pengisian air dirancang untuk mengintegrasikan teknologi ke dalam operasi sehari-hari. Awalnya, pemilik bisnis dapat menjalankan program yang telah dibuat dalam format .exe, yang kemudian akan menampilkan antarmuka pengguna grafis (GUI) yang dibangun menggunakan Tkinter yang ditunjukkan pada Gambar 12.



GAMBAR 12. GUI sistem yang menggunakan Tkinter

Dengan antarmuka yang intuitif, pengguna cukup menekan tombol 'START' untuk memulai proses deteksi, memanfaatkan RTSP untuk preview deteksi. Setelah sistem beroperasi, pemilik bisnis dapat memantau proses pengambilan air melalui *chatbot* di aplikasi mobile Telegram, memungkinkan mereka untuk mengidentifikasi truk mana yang sedang mengambil air. Sistem ini juga menyediakan rekap harian yang mencakup informasi tentang jumlah truk per hari dan total uang yang seharusnya diterima, memudahkan pemilik bisnis dalam mengelola dan melacak transaksi. Selain itu pemilik bisnis dapat menginformasikan sopir truk tentang total yang harus dibayarkan, baik untuk pembayaran tunggal atau beberapa transaksi sekaligus, berdasarkan rekap harian di Telegram. Terakhir, sistem ini mendukung proses penyelesaian keluhan, dimana pemilik bisnis dapat memvalidasi klaim dari sopir truk menggunakan foto yang dikirimkan oleh sistem ke Telegram, menjamin transparansi dan keadilan dalam setiap transaksi. Keseluruhan implementasi ini bertujuan untuk meningkatkan efisiensi, akurasi, dan keandalan dalam operasi bisnis pengisian air.

E. KETERBATASAN SISTEM

Keterbatasan utama dalam sistem deteksi truk air ini terletak pada kemampuannya untuk mengklasifikasikan truk. Sistem ini bergantung pada dataset model yang telah ada. Sehingga ketika truk air

yang belum terdaftar dalam dataset diperkenalkan, sistem cenderung mengalokasikan truk tersebut ke kelas yang paling mirip yang ada dalam dataset. Hal ini mengakibatkan potensi kesalahan dalam klasifikasi. Untuk mengatasi keterbatasan ini, kami mengusulkan penambahan fitur self-learning dalam sistem. Mekanisme ini memungkinkan sistem untuk mengidentifikasi truk baru sebagai 'unknown_water_truck'. Selanjutnya, sistem akan mengumpulkan data dari truk-truk yang dikategorikan sebagai 'unknown' ini. Pengguna kemudian dapat memberikan label pada truk-truk ini melalui GUI yang tersedia. Proses pelabelan ini tidak hanya mengurangi kesalahan klasifikasi tetapi juga memperkaya dataset model dengan kelas-kelas baru, sehingga meningkatkan kemampuan adaptasi dan akurasi sistem dalam pengenalan truk air. Pendekatan ini menunjukkan potensi perbaikan yang signifikan dalam hal pengelolaan dan peningkatan data, yang merupakan aspek penting dalam pengembangan sistem pengenalan berbasis pembelajaran mesin.

V. KESIMPULAN

Sistem yang direncanakan untuk memberi solusi kepada pemilik usaha berhasil diimplementasikan. Sistem ini terbukti dapat mendeteksi truk air dan juga mengklasifikasikan setiap truk air dengan benar. Model YOLOv5 sebagai objek detector dan EfficientNetB4 sebagai model untuk melakukan klasifikasi dapat digunakan pada sistem tersebut.

Semua performance metrics Precision, Recall dan mAP untuk model YOLOv5 menunjukkan hasil di atas 90%. Sedangkan performance metrics untuk model EfficientNetB4 yaitu Accuracy, Precision, Recall dan F1 Score semuanya juga menunjukkan hasil di atas 90%.

Pada penelitian ini terdapat beberapa hal yang dapat ditingkatkan yaitu tingkat accuracy untuk object detection yang dilakukan YOLOv5 dalam mendeteksi object lain, misalnya 'Person' dengan tingkat Precision dan Recall 78%.

Hasil pengujian untuk object detection dan classification dengan menggunakan evaluation metrics menunjukkan bahwa metode tersebut berhasil melakukan deteksi dan klasifikasi dengan baik. Pengujian keseluruhan sistem menunjukkan bahwa semua tahap pengujian berhasil dilakukan dengan baik. Hal ini menunjukkan bahwa sistem tersebut dapat digunakan untuk mengatasi masalah yang dihadapi.

PERAN PENULIS

Ardian Kurniawan Langgeng Santoso: Implementasi arsitektur sistem, implementasi object detector dan algoritma classifier, implementasi sistem dan perangkat lunak.

Daniel Martomangolo Wonohadidjojo: Metodologi, Perancangan arsitektur sistem, perancangan training, validation dan testing robust model, perancangan dataset, dan penyusunan artikel.

COPYRIGHT



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

DAFTAR PUSTAKA

- [1] Aballe *et al.*, "Security Measures: Effectiveness of the Installation of CCTV Cameras in Relation to Crime Prevention as Perceived by the Community," *Middle East Journal of Applied Science & Technology*, vol. 05, no. 02, pp. 149–160, 2022, doi: 10.46431/mejast.2022.5216.
- [2] M. Kohler and B. Walter, "Analysis of convolutional neural network image classifiers in a rotationally symmetric model," May 2022, [Online]. Available: <http://arxiv.org/abs/2205.05500>
- [3] J. S. Murthy, G. M. Siddesh, W. C. Lai, B. D. Parameshachari, S. N. Patil, and K. L. Hemalatha, "ObjectDetect: A Real-Time Object Detection Framework for Advanced Driver Assistant Systems Using YOLOv5," *Wirel Commun Mob Comput*, vol. 2022, 2022, doi: 10.1155/2022/9444360.
- [4] K. Sharma, S. Singh Rawat, D. Parashar, S. Sharma, S. Roy, and S. Sahoo, "State-of-the-Art Analysis of Multiple Object Detection Techniques using Deep Learning." [Online]. Available: www.ijacsa.thesai.org
- [5] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," May 2019, [Online]. Available: <http://arxiv.org/abs/1905.11946>
- [6] R. N. Lazuardi, N. Abiwinanda, T. H. Suryawan, M. Hanif, and A. Handayani, "Automatic diabetic retinopathy classification with efficientnet," in *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020, pp. 756–760. doi: 10.1109/TENCON50793.2020.9293941.
- [7] IEEE Staff, *2019 International Conference on Sustainable Information Engineering and Technology (SIET)*. IEEE, 2019.
- [8] D. I. Mulyana and M. A. Rofik, "Implementasi Deteksi Real Time Klasifikasi Jenis Kendaraan Di Indonesia Menggunakan Metode YOLOV5."
- [9] Z. Chen, X. Li, L. Wang, Y. Shi, Z. Sun, and W. Sun, "An Object Detection and Localization Method Based on Improved YOLOv5 for the Teleoperated Robot," *Applied Sciences (Switzerland)*, vol. 12, no. 22, Nov. 2022, doi: 10.3390/app122211441.
- [10] S. S. Reka, V. D. Murthy Voona, P. V. Sai Nithish, D. S. Paavan Kumar, P. Venugopal, and V. Ravi, "Performance Analysis of Deep Convolutional Network Architectures for Classification of Over-Volume Vehicles," *Applied Sciences (Switzerland)*, vol. 13, no. 4, Feb. 2023, doi: 10.3390/app13042549.
- [11] T. Hasan Rafi, "A Holistic Approach to Identification of Covid-19 Patients from Chest X-Ray Images utilizing Transfer Based Learning", doi: 10.1101/2020.07.08.20148924.
- [12] H. Liang, J. Chen, W. Xie, X. Yu, and W. Wu, "Defect detection of injection-molded parts based on improved-YOLOv5," in *Journal of Physics: Conference Series*, Institute of Physics, 2022. doi: 10.1088/1742-6596/2390/1/012049.

INSYST

Journal of Intelligent System and Computation

Volume 05 Nomor 02 Oktober 2023

Author Guidelines

- Manuscript should be written in Indonesia and be submitted online via journal website. Online Submission will be charged at no Cost
- Manuscript should not exceed 15 pages including embedded figures and tables, without any appendix, and the file should be in Microsoft Office (.doc/.docx). [download template](#)
- Title should be less than 15 words
- Abstracts consists of no more than 200 words, contains the essence of the article and includes a brief background, objectives, methods and results or findings of the study. Abstract is written in one paragraph.
- Keywords are written in Indonesia three to five words/phrases, separated with coma and consist of important words/phrases from the article.
- Author's name, affiliation, affiliation address and email. State clearly and include country's name on your affiliation address.
- The main text of the writing should be consists of: Introduction, Method, Result and Discussion, and Conclusion; followed by Acknowledgment and Reference
- Introduction State adequate background, issues and objectives, avoiding a detailed literature survey or a summary of the results. Explain how you addressed the problem and clearly state the aims of your study.
- Used method is the scientific in the form of study of literature, observation, surveys, interviews, Focus Group Discussion, system testing or simulation and other techniques commonly used in the world of research. It is also recommended to describe analysis techniques used briefly and clearly, so that the reader can easily understand.
- Results should be clear, concise and not in the form of raw data. Discussion should explore the significance of the results of the work, not repeat them. Avoid extensive citations and discussion of published literature. INSYST will do the final formatting of your paper.
- Conclusion should lead the reader to important matter of the paper. Authors are allowed to include suggestion or recommendation in this section. Write conclusion, suggestion and/or recommendation in narrative form (avoid of using bulleting and numbering)
- Acknowledgments. It is highly recommended to acknowledge a person and/or organizations helping author(s) in many ways. Sponsor and financial support acknowledgments should be included in this section. Should you have lots of parties

to be acknowledged, state your acknowledgments only in one paragraph. Avoid of using bulleting and numbering in this section

- The number of references are not less than 10 with at least 8 primary references. Primary references are include journal, thesis, disertasion and all kinds of research reports. All refferences must come from source published in last 7 years.
- Figure and table should be in black and white, and if it is made in color, it should be readable when it is later printed in black and white.
- Figure and table should be clearly readable and in a proportional measure to the overall page.

Tim Redaksi

Journal of Intelligent System and Computation

Departement of Informatics

Institut Sains dan Teknologi Terpadu Surabaya

Jl. Ngagel Jaya Tengah 73-77 Surabaya

Email: insyst@istts.ac.id

Website: <https://jurnal.stts.edu/index.php/INSYST/index>