

Air Pollution Level Estimation In Public Areas Using Image Segmentation And Validation Against Air Quality Sensor Data

Muhamad Reyipi ¹⁾; Rozali Toyib ²⁾; Yulia Darmi ³⁾; Yopi Apridiansyah ⁴⁾; Yusa Virginiawan ⁵⁾

^{1,2,3,4,5)}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Bengkulu

Email: ¹⁾muhamadreypi07@gmail.com

How to Cite :

Reyipi, M., Toyib, R., Darmi, Y., Apridiansyah, Y., Virginiawan, Y. (2026). Air Pollution Level Estimation In Public Areas Using Image Segmentation And Validation Against Air Quality Sensor Data, Support Vector Machine and Logistic Regression Algorithms. Jurnal Media Computer Science, 5(4).

ARTICLE HISTORY

Received [23 Juli 2026]

Revised [23 Oktober 2026]

Accepted [26 Oktober 2026]

KEYWORDS

Air Pollution, Image Segmentation, HSV, CO₂ Sensor, Classification, Confusion Matrix, Cross-Validation, Python.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



ABSTRAK

Penelitian ini mengembangkan sebuah sistem estimasi tingkat polusi udara berbasis pengolahan citra digital yang diintegrasikan dengan validasi menggunakan sensor CO₂. Dibangun dalam platform aplikasi desktop berbasis Python. Metode yang digunakan meliputi segmentasi citra pada ruang warna HSV dengan pendekatan thresholding multi-zona untuk mendeteksi area asap, kabut, dan polutan visual lainnya, dilanjutkan dengan ekstraksi fitur berupa rasio area polusi, entropi tekstur, mean intensitas, standar deviasi, dan saturasi rata-rata yang kemudian dikombinasikan menjadi skor komposit sebagai dasar klasifikasi ke dalam tiga kategori yaitu Rendah, Sedang, dan Tinggi. Secara paralel, sistem juga menganalisis citra display sensor CO₂ melalui deteksi rasio warna indikator hijau, kuning, dan merah untuk menghasilkan estimasi konsentrasi gas CO₂ di udara. Validasi silang antara kedua metode dilakukan terhadap 100 pasang citra untuk mengukur tingkat konsistensi hasil klasifikasi. Hasil evaluasi menggunakan confusion matrix menunjukkan akurasi keseluruhan sistem sebesar 70,00% dengan Macro Precision 37,90%, Macro Recall 45,93%, dan Macro F1-Score 36,70%. Kelas Tinggi memiliki kinerja terbaik dengan F1-Score 81,53%, sedangkan kelas Sedang mencapai F1-Score 28,57% dan kelas Rendah belum berhasil terdeteksi akibat ketidakseimbangan distribusi data. Hasil validasi silang menunjukkan tingkat kesepakatan antara segmentasi citra dan sensor CO₂ sebesar 37%, yang dikategorikan perlu perbaikan, dan mencerminkan perbedaan dimensi pengukuran antara kenampakan visual polusi dan konsentrasi gas secara langsung. Sistem yang dibangun telah dilengkapi dengan antarmuka grafis interaktif, visualisasi confusion matrix, dan fitur ekspor laporan Excel multi-sheet yang mencakup detail hasil, evaluasi kinerja, analisis sensor, serta kesimpulan dan rekomendasi otomatis.

ABSTRACT

This study develops an air pollution level estimation system based on digital image processing, integrated with validation using a CO₂ sensor, and implemented as a Python-based desktop application. The proposed method employs image segmentation in the HSV color space using a multi-zone thresholding approach to detect smoke, haze, and other visual pollutants. Feature extraction is then performed by calculating the pollution area ratio, texture entropy, mean intensity, standard deviation, and average saturation. These features are combined into a composite score that serves as the basis for classifying air pollution levels into three categories: Low, Moderate, and High. In parallel, the system analyzes images of the CO₂ sensor display by detecting the color ratio of green, yellow, and red indicators to estimate atmospheric CO₂ concentration levels. Cross-validation between the two methods was conducted using 100 paired images to evaluate the consistency of classification results. Evaluation results based on a confusion matrix indicate an overall system accuracy of 70.00%, with a Macro Precision of 37.90%, Macro Recall of 45.93%, and Macro F1-Score of 36.70%. The High category demonstrated the best performance, achieving an F1-Score of 81.53%, while the Moderate category obtained an F1-Score of 28.57%. The Low category could not be effectively detected due to data distribution imbalance. Cross-validation results showed a 37% agreement rate between image segmentation and CO₂ sensor measurements, indicating that further improvements are required. This result also reflects the fundamental difference between visual pollution indicators and direct gas concentration measurements. The developed system is equipped with an interactive graphical user interface, confusion matrix visualization, and a multi-sheet Excel report export feature that includes detailed results, performance evaluation, sensor analysis, as well as automatically generated conclusions and recommendations.

PENDAHULUAN

Polusi udara merupakan fenomena nyata yang semakin sering dijumpai di berbagai wilayah perkotaan, khususnya pada area publik seperti jalan raya, persimpangan lalu lintas, terminal, dan taman kota. Peningkatan jumlah kendaraan bermotor, aktivitas industri, serta kepadatan penduduk menyebabkan kualitas udara di lingkungan perkotaan mengalami penurunan secara signifikan (Sriliani, Nusa Sebayang, 2024). Kondisi ini seringkali dapat diamati secara visual melalui keberadaan asap, kabut, atau menurunnya jarak pandang di suatu wilayah, yang menunjukkan adanya pencemaran udara pada tingkat tertentu.

Permasalahan yang terjadi adalah keterbatasan sistem pemantauan kualitas udara yang mampu menjangkau seluruh area publik secara merata dan berkelanjutan. Alat pemantau kualitas udara konvensional umumnya hanya dipasang pada titik-titik tertentu, sehingga data yang dihasilkan belum sepenuhnya merepresentasikan kondisi udara di seluruh wilayah. Selain itu, tidak semua daerah memiliki infrastruktur pemantauan yang memadai akibat keterbatasan biaya, sumber daya, dan pemeliharaan perangkat sensor.

Dampak dari permasalahan tersebut cukup serius, terutama terhadap kesehatan masyarakat dan kualitas lingkungan. Paparan polusi udara dalam jangka pendek maupun panjang dapat menyebabkan gangguan pernapasan, penyakit kardiovaskular, serta menurunkan kualitas hidup masyarakat. Di sisi lain, kurangnya informasi kondisi udara secara real-time di area publik juga menyulitkan pemerintah dan masyarakat dalam mengambil tindakan preventif untuk mengurangi risiko kesehatan dan dampak lingkungan yang lebih luas.

Berdasarkan tinjauan terhadap penelitian-penelitian sebelumnya, masih terdapat celah penelitian dalam pemanfaatan data visual sebagai alternatif atau pelengkap sistem pemantauan kualitas udara. Sebagian besar penelitian lebih berfokus pada data sensor numerik, sementara potensi citra digital yang tersedia dari kamera pengawas atau kamera lalu lintas belum dimanfaatkan secara optimal untuk mendeteksi tingkat polusi udara. Hal ini menunjukkan adanya peluang untuk mengembangkan pendekatan berbasis pengolahan citra dalam pemantauan kualitas udara.

Oleh karena itu, penelitian ini mengusulkan pendekatan untuk mengestimasi tingkat polusi udara pada area publik menggunakan teknik segmentasi citra. Pendekatan ini memanfaatkan citra digital lingkungan untuk mengidentifikasi karakteristik visual yang berkaitan dengan keberadaan kabut, asap, atau penurunan visibilitas yang menjadi indikator pencemaran udara. Hasil analisis citra kemudian digunakan untuk mengestimasi tingkat polusi udara ke dalam kategori tertentu, yaitu rendah, sedang, dan tinggi. Untuk memastikan keandalan metode yang diusulkan, hasil estimasi yang diperoleh dari citra akan divalidasi dengan data sensor kualitas udara yang diperoleh dari platform pemantauan kualitas udara seperti data Air Quality Index (AQI) atau konsentrasi partikulat PM_{2.5}. Dengan demikian, pendekatan berbasis pengolahan citra ini diharapkan dapat menjadi alternatif metode pemantauan kualitas udara yang lebih fleksibel dan ekonomis.

Metode segmentasi citra dipilih karena memiliki kemampuan untuk memisahkan area yang terindikasi polusi dari latar belakang citra berdasarkan perbedaan karakteristik warna dan intensitas. Selain itu, penggunaan fitur visual yang diekstraksi dari hasil segmentasi memungkinkan proses klasifikasi dilakukan secara sederhana dan efisien.

Diharapkan penelitian ini dapat memberikan kontribusi dalam bentuk alternatif sistem pemantauan kualitas udara yang lebih ekonomis, fleksibel, dan mudah diterapkan. Selain itu, hasil penelitian ini diharapkan mampu menjadi dasar pengembangan sistem pemantauan lingkungan berbasis visi komputer serta memberikan referensi akademik bagi penelitian selanjutnya dalam bidang pengolahan citra digital dan pemantauan kualitas udara.

LANDASAN TEORI

Pengolahan Citra Digital

Pengolahan citra digital didefinisikan sebagai proses memanipulasi dan menganalisis data visual dua dimensi menggunakan komputer untuk menghasilkan representasi yang lebih informatif atau untuk mengekstrak fitur-fitur tertentu yang tidak langsung terlihat oleh mata manusia (Hardini *et al.*, 2023). Rangkaian operasi dalam pengolahan citra biasanya mengalir secara berurutan, dimulai dari akuisisi (penangkapan citra melalui kamera atau sensor), pre-processing (perbaikan kualitas seperti peningkatan kontras, reduksi derau, dan normalisasi cahaya), segmentasi (pemisahan objek dari latar belakang), ekstraksi ciri (pengambilan parameter numerik seperti tekstur, bentuk, atau warna), hingga interpretasi

atau klasifikasi. Tujuan akhir dari keseluruhan proses ini adalah agar sistem komputer mampu "memahami" isi citra secara otomatis, atau setidaknya menyajikan informasi visual dalam bentuk yang lebih mudah dianalisis oleh manusia (Liaw *et al.*, 2020).

Preprocessing

Preprocessing merupakan tahap awal dan krusial dalam alur pengolahan citra digital yang bertujuan untuk memperbaiki kualitas citra mentah sebelum memasuki proses lanjutan seperti segmentasi, ekstraksi ciri, atau klasifikasi. Citra yang diakuisisi dari berbagai sumber baik kamera digital, smartphone, maupun citra satelit seringkali mengandung berbagai bentuk gangguan atau ketidaksempurnaan. Gangguan ini dapat berupa derau (noise) akibat keterbatasan sensor, pencahayaan yang tidak merata, kontras yang terlalu rendah atau terlalu tinggi, distorsi geometris akibat sudut pengambilan gambar, hingga artefak kompresi. Jika tidak ditangani dengan baik, gangguan-gangguan tersebut akan menurunkan akurasi interpretasi pada tahap-tahap selanjutnya (Rizki Akbar Pratama, Barry Ceasar, 2025).

Polusi Udara

Polusi udara merupakan kondisi di mana terdapat kontaminan atau zat-zat berbahaya di dalam atmosfer yang dapat membahayakan kesehatan makhluk hidup, merusak lingkungan, serta menurunkan kualitas kehidupan manusia. Menurut *World Health Organization* (WHO), polusi udara didefinisikan sebagai pencemaran lingkungan dalam ruangan maupun luar ruangan oleh agen kimia, fisik, atau biologis yang mengubah karakteristik alami atmosfer. Sumber polusi udara dapat berasal dari aktivitas manusia (*anthropogenic*) seperti emisi kendaraan bermotor, pembakaran bahan bakar fosil, aktivitas industri, dan pembakaran biomassa, maupun sumber alami seperti letusan gunung berapi, kebakaran hutan, dan debu vulkanik (Najwa *et al.*, 2025).

Gambar 1 Contoh Polusi Udara



Sensor Kualitas Udara

Sensor kualitas udara merupakan perangkat elektronik yang dirancang untuk mendeteksi, mengukur, dan memantau konsentrasi berbagai polutan atau parameter fisikokimia yang terdapat di udara ambien. Perkembangan teknologi sensor kualitas udara telah mengalami kemajuan yang signifikan, mulai dari perangkat stasioner berukuran besar dengan akurasi tinggi yang digunakan pada stasiun pemantauan udara resmi, hingga sensor berukuran kecil (low-cost sensor) yang dapat diintegrasikan ke dalam sistem pemantauan portabel dan jaringan IoT (Internet of Things). Sensor kualitas udara bekerja berdasarkan berbagai prinsip pengukuran fisika dan kimia, di antaranya deteksi optis menggunakan hamburan cahaya laser, reaksi elektrokimia, spektroskopi inframerah, serta metode gravimetri. Pemilihan jenis sensor disesuaikan dengan polutan target yang ingin diukur, rentang konsentrasi yang diharapkan, tingkat akurasi yang diperlukan, serta kondisi lingkungan tempat sensor dioperasikan (Hendi Budianto, 2024).

Segmentasi Citra

Segmentasi citra merupakan salah satu tahapan dalam pengolahan citra digital (digital image processing) yang bertujuan untuk mempartisi atau membagi suatu citra menjadi beberapa region atau segmen yang memiliki karakteristik serupa, seperti warna, intensitas, tekstur, atau batas tepi objek. Proses segmentasi mengubah representasi citra menjadi sesuatu yang lebih bermakna dan lebih mudah dianalisis (Gibran, 2025). Hasil segmentasi berupa sekumpulan region yang secara kolektif mencakup seluruh area citra, di mana setiap piksel dalam satu region memiliki kemiripan atribut visual yang tinggi dibandingkan piksel di region berbeda.

Dalam konteks estimasi polusi udara, segmentasi citra digunakan untuk memisahkan elemen-elemen visual dalam gambar seperti langit, vegetasi, bangunan, jalan, dan asap, sehingga analisis dapat difokuskan pada region yang paling relevan terhadap indikasi keberadaan polutan di atmosfer.

Phyton

Python merupakan salah satu bahasa pemrograman tingkat tinggi (high-level programming language) yang banyak digunakan dalam pengembangan perangkat lunak, analisis data, kecerdasan buatan, serta pengolahan citra digital. Bahasa pemrograman ini pertama kali dikembangkan oleh Guido van Rossum pada tahun 1991 dan dirancang dengan tujuan agar mudah dipelajari serta memiliki sintaks yang sederhana dan mudah dibaca. Python bersifat open source, sehingga dapat digunakan secara bebas oleh siapa saja dan terus dikembangkan oleh komunitas pengembang di seluruh dunia (Buwono *et al.*, 2025).

METODE PENELITIAN

Metode Penelitian

Metode yang digunakan meliputi segmentasi citra pada ruang warna HSV dengan pendekatan thresholding multi-zona untuk mendeteksi area asap, kabut, dan polutan visual lainnya, dilanjutkan dengan ekstraksi fitur berupa rasio area polusi, entropi tekstur, mean intensitas, standar deviasi, dan saturasi rata-rata yang kemudian dikombinasikan menjadi skor komposit sebagai dasar klasifikasi ke dalam tiga kategori yaitu Rendah, Sedang, dan Tinggi. Secara paralel, sistem juga menganalisis citra display sensor CO₂ melalui deteksi rasio warna indikator hijau, kuning, dan merah untuk menghasilkan estimasi konsentrasi gas CO₂ di udara. Validasi silang antara kedua metode dilakukan terhadap 100 pasang citra untuk mengukur tingkat konsistensi hasil klasifikasi.

Perancangan Program

Sistem estimasi tingkat polusi udara yang dirancang dalam penelitian ini merupakan aplikasi berbasis pengolahan citra digital yang mengintegrasikan dua pendekatan analisis, yaitu segmentasi citra lingkungan dan validasi data sensor kualitas udara berbasis CO₂. Sistem dibangun menggunakan bahasa pemrograman Python dengan memanfaatkan pustaka OpenCV untuk pemrosesan citra, Tkinter sebagai kerangka antarmuka grafis, serta NumPy dan Matplotlib untuk visualisasi data.

Secara keseluruhan, sistem terdiri atas tujuh modul utama yang disusun secara sekuensial mengikuti alur metode penelitian, yaitu: akuisisi citra, prapemrosesan citra, segmentasi area polusi, ekstraksi fitur, klasifikasi tingkat polusi, evaluasi kinerja model, dan validasi menggunakan data sensor CO₂. Setiap modul diimplementasikan sebagai tab tersendiri di dalam antarmuka pengguna berbasis notebook, sehingga pengguna dapat mengikuti setiap tahapan proses secara terstruktur dan mudah dipahami.

HASIL DAN PEMBAHASAN

Hasil

Hasil evaluasi menggunakan confusion matrix menunjukkan akurasi keseluruhan sistem sebesar 70,00% dengan Macro Precision 37,90%, Macro Recall 45,93%, dan Macro F1-Score 36,70%. Kelas Tinggi memiliki kinerja terbaik dengan F1-Score 81,53%, sedangkan kelas Sedang mencapai F1-Score 28,57% dan kelas Rendah belum berhasil terdeteksi akibat ketidakseimbangan distribusi data. Hasil validasi silang menunjukkan tingkat kesepakatan antara segmentasi citra dan sensor CO₂ sebesar 37%, yang dikategorikan perlu perbaikan, dan mencerminkan perbedaan dimensi pengukuran antara kenampakan visual polusi dan konsentrasi gas secara langsung. Sistem yang dibangun telah dilengkapi dengan antarmuka grafis interaktif, visualisasi confusion matrix, dan fitur ekspor laporan Excel multi-sheet yang mencakup detail hasil, evaluasi kinerja, analisis sensor, serta kesimpulan dan rekomendasi otomatis.

Implementasi

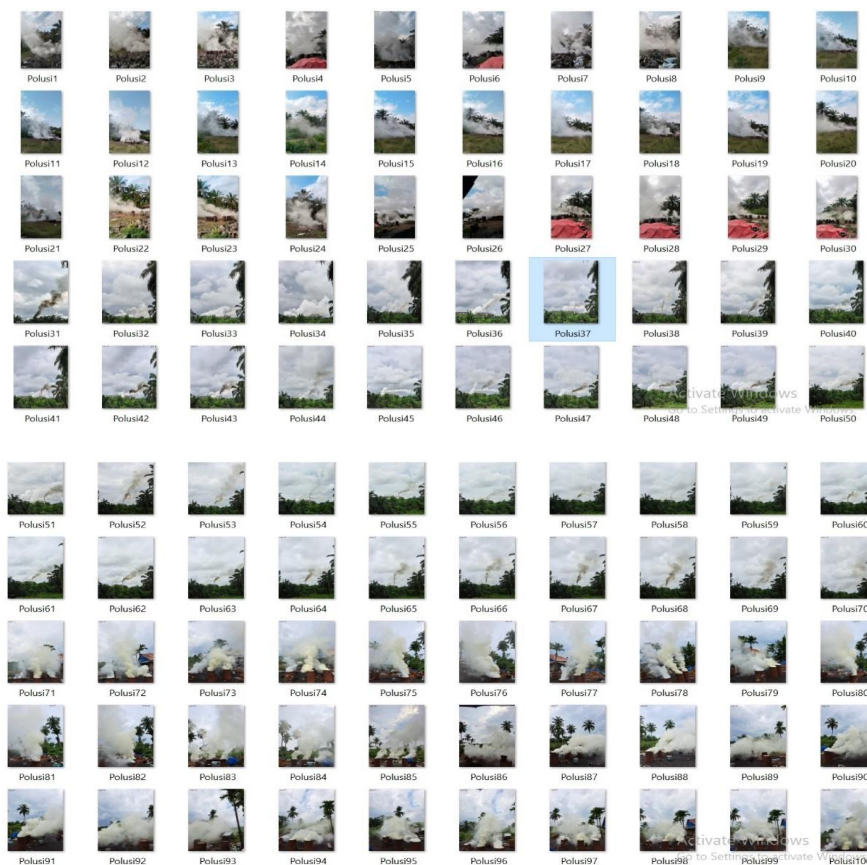
Implementasi sistem estimasi tingkat polusi udara pada area publik dilakukan secara bertahap mengikuti alur metode pelaksanaan eksperimen yang telah dirancang pada bab sebelumnya. Seluruh tahapan implementasi dikembangkan menggunakan bahasa pemrograman Python dengan memanfaatkan pustaka OpenCV untuk pengolahan citra, NumPy untuk komputasi numerik, serta scikit-learn untuk pembangunan model klasifikasi. Lingkungan pengembangan yang digunakan adalah Visual

Studio Code dengan dukungan ekstensi Python dan Jupyter Notebook yang memungkinkan setiap tahapan diuji secara interaktif dan inkremental sebelum diintegrasikan ke dalam sistem secara keseluruhan.

Persiapan Dataset Citra Polusi Udara

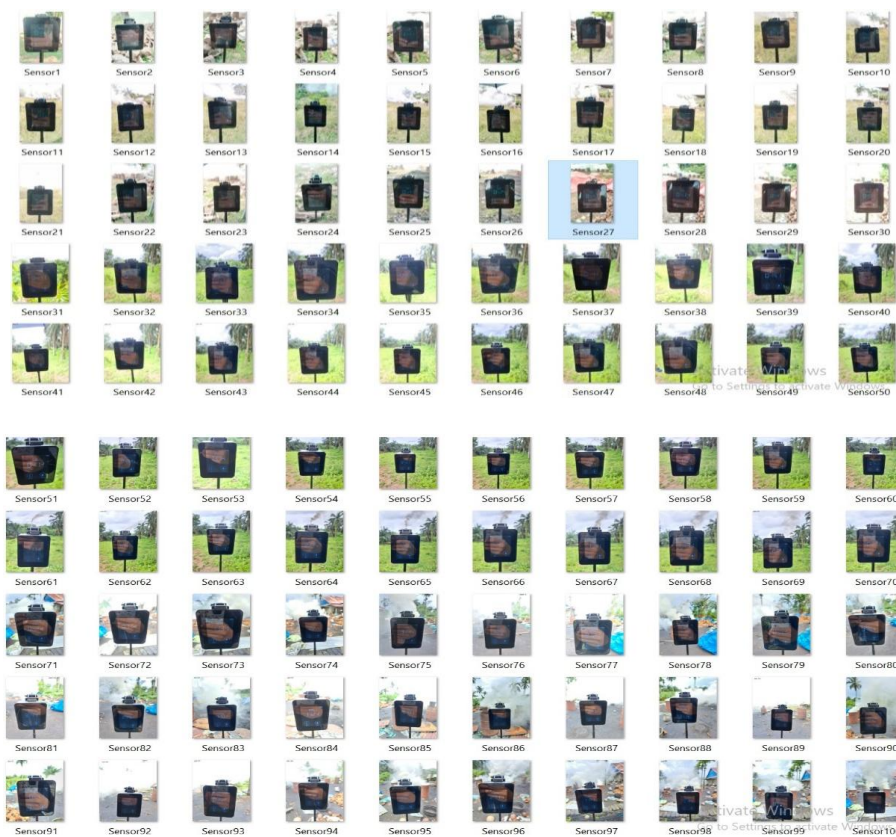
Tahap pertama implementasi adalah akuisisi citra yang merupakan proses pengumpulan dataset citra area publik sebagai bahan utama penelitian. Citra dikumpulkan dari berbagai lokasi area publik yang memiliki variasi tingkat polusi udara yang berbeda, mencakup kawasan area sekitar industri, dan kawasan terbuka yang relatif bersih dari polutan, sebagian dataset citra polusi udara tersebut diambil melalui kamera handphone dikota Manna Bengkulu Selatan. Dataset yang berhasil dikumpulkan terdiri dari 100 citra dalam format JPEG yang kemudian dimuat ke dalam sistem menggunakan fungsi pembacaan citra pada pustaka OpenCV. Setiap citra dalam dataset diberikan label ground truth secara manual berdasarkan kondisi polusi udara yang teramati secara visual dan disesuaikan dengan data referensi sensor kualitas udara, dengan tiga kategori label yaitu Rendah dengan kode 0, Sedang dengan kode 1, dan Tinggi dengan kode 2, sehingga diperoleh dataset berlabel yang siap digunakan sebagai data latih dan data uji dalam proses klasifikasi.

Gambar 2 Dataset Citra Polusi Udara



Dataset tersebut dikumpulkan sebanyak 100 citra sampel untuk diuji hasil segmentasi deteksi polusi udara. Data citra diberi nama polusi1 sampai dengan polusi 100. Kemudian untuk melihat hasil proses segmentasi polusi udara dengan perbandingan hasil analisa dari alat sensor CO₂, maka pada saat melakukan proses pengambilan citra polusi udara peneliti juga langsung memasang alat deteksi karbondioksida diarea sekitar polusi menggunakan alat detektor. Sampel citra dengan alat ini juga dikumpulkan masing masing sebanyak 100 sampel. Hasil citra sensor sebagai berikut.

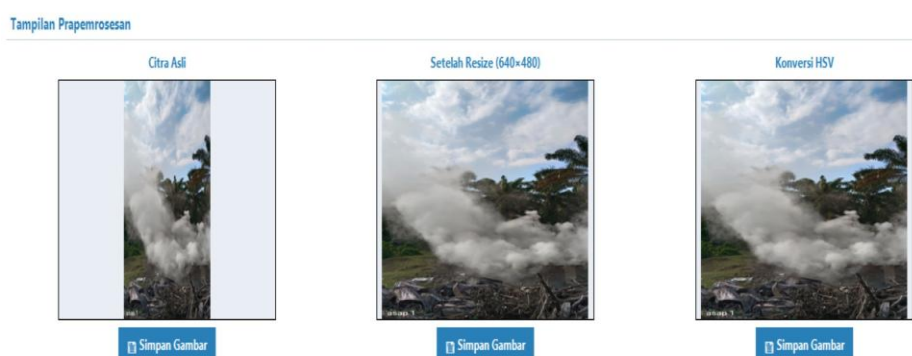
Gambar 3 Dataset Citra Polusi Udara Dengan Sensor



Tahap Preprocessing

Setelah dataset citra berhasil dikumpulkan dan dilabeli, tahap berikutnya adalah prapemrosesan citra yang bertujuan untuk menyeragamkan format dan karakteristik seluruh citra sebelum memasuki tahap segmentasi. Proses prapemrosesan diawali dengan operasi resize yang mengubah ukuran seluruh citra menjadi dimensi yang seragam yaitu 640×480 piksel, sehingga seluruh citra memiliki resolusi yang konsisten dan proses komputasi pada tahap selanjutnya dapat berjalan secara efisien tanpa pengaruh perbedaan ukuran proses resize ini dilakukan langsung pada system sehingga pada saat input citra dan melakukan proses seluruh dataset yang ada langsung otomatis diubah ukurannya kedalam 640×480 . Selanjutnya dilakukan konversi ruang warna dari RGB ke HSV menggunakan fungsi `cv2.cvtColor()` pada OpenCV, karena ruang warna HSV memisahkan informasi warna pada kanal Hue dari informasi kecerahan pada kanal Value, sehingga analisis visual citra menjadi lebih robust terhadap variasi kondisi pencahayaan yang sering terjadi pada citra yang diambil di lingkungan luar ruangan pada waktu yang berbeda-beda.

Gambar 4 Tahap Preprocessing



Segmentasi Citra

Tahap segmentasi citra merupakan inti dari sistem yang dikembangkan, di mana setiap citra yang telah melalui prapemrosesan dipartisi untuk memisahkan region yang mengandung indikasi visual polusi udara dari elemen-elemen lain dalam citra. Implementasi segmentasi dilakukan menggunakan metode thresholding pada ruang warna HSV, di mana nilai ambang batas ditetapkan secara adaptif berdasarkan distribusi nilai piksel pada kanal Saturation dan Value yang secara langsung mencerminkan kondisi kejernihan atmosfer dalam citra. Piksel-piksel yang memiliki nilai saturasi rendah dan kecerahan tinggi diidentifikasi sebagai region langit atau area berpolusi yang menjadi fokus analisis, sedangkan piksel yang merepresentasikan objek padat seperti bangunan, vegetasi, dan kendaraan dipisahkan dari region analisis melalui operasi masking. Hasil segmentasi kemudian disempurnakan menggunakan operasi morfologi berupa dilasi dan erosi untuk menghilangkan noise piksel yang terisolasi dan menutup celah kecil pada region yang tersegmentasi, sehingga diperoleh area polusi yang bersih dan representatif untuk proses ekstraksi fitur pada tahap selanjutnya.

Gambar 5 Tahap Segmentasi Citra



Gambar tersebut menampilkan empat bagian utama untuk menganalisis polusi. Pertama, "Citra Asli" menunjukkan gambar asli tanpa modifikasi. Kedua, "Mask Segmentasi" menyajikan area yang teridentifikasi sebagai polusi dalam bentuk topeng biner. Ketiga, "Area Polusi" menampilkan wilayah polusi yang telah dipisahkan dari latar belakang. Terakhir, "Overlay Polusi" menggabungkan area polusi dengan citra asli untuk menunjukkan lokasi polusi secara visual. Setiap bagian dilengkapi tombol "Simpan" untuk menyimpan hasil.

Tahap Ekstraksi Fitur

Setelah proses segmentasi menghasilkan region area polusi yang telah terisolasi, tahap ekstraksi fitur dilakukan untuk mengubah informasi visual dari region tersebut menjadi representasi numerik yang dapat diproses oleh model klasifikasi. Tiga fitur utama diekstraksi dari setiap citra yang telah tersegmentasi, yaitu mean intensitas, standar deviasi, dan entropi citra. Mean intensitas dihitung sebagai rata-rata nilai piksel pada region tersegmentasi yang merepresentasikan tingkat kecerahan rata-rata area atmosfer dalam citra, di mana nilai mean intensitas yang rendah umumnya mengindikasikan keberadaan partikulat polutan yang menyebabkan peredupan visual. Standar deviasi dihitung untuk mengukur tingkat variasi atau heterogenitas nilai piksel pada region tersegmentasi, karena kondisi udara berpolusi cenderung menghasilkan distribusi piksel yang lebih homogen dengan variasi rendah akibat efek pengaburan yang merata. Entropi citra dihitung menggunakan formula entropi Shannon untuk mengukur kompleksitas tekstur pada region tersegmentasi, di mana nilai entropi yang rendah menunjukkan tekstur yang seragam dan berkorelasi dengan keberadaan kabut polusi yang menyebabkan hilangnya detail tekstur pada citra. Ketiga nilai fitur tersebut kemudian digabungkan menjadi satu vektor fitur untuk setiap citra yang selanjutnya digunakan sebagai masukan pada tahap klasifikasi.

Tahap Klasifikasi Tingkat Polusi

Tahap klasifikasi bertujuan untuk memetakan vektor fitur yang telah diekstraksi ke dalam salah satu dari tiga kategori tingkat polusi udara yaitu Rendah, Sedang, dan Tinggi. Implementasi klasifikasi menggunakan algoritma Random Forest Classifier dari pustaka scikit-learn yang dipilih karena kemampuannya menangani hubungan non-linear antara fitur citra dengan label tingkat polusi serta robustness terhadap overfitting melalui mekanisme ensemble dari banyak pohon keputusan.

Tahap Evaluasi Kinerja Model

Evaluasi kinerja model menggunakan confusion matrix beserta metrik-metrik evaluasi turunannya untuk mengukur kemampuan model dalam mengklasifikasikan tingkat polusi udara secara akurat pada seluruh kategori. Confusion matrix dibangun dengan membandingkan label prediksi model terhadap label ground truth pada data uji, menghasilkan matriks berukuran 3×3 yang menunjukkan distribusi prediksi benar dan salah untuk setiap kombinasi kelas aktual dan kelas prediksi. Dari confusion matrix tersebut dihitung nilai akurasi keseluruhan sebagai proporsi prediksi benar terhadap total sampel uji, nilai presisi per kelas yang mengukur ketepatan prediksi positif model, serta nilai recall per kelas yang mengukur kemampuan model mendeteksi seluruh sampel dari setiap kategori polusi. Selain itu dihitung pula F1-Score sebagai rata-rata harmonik antara presisi dan recall yang memberikan gambaran keseimbangan kinerja model secara menyeluruh. Hasil evaluasi ini menjadi dasar pengambilan keputusan apakah model telah mencapai tingkat akurasi yang memadai untuk dilanjutkan ke tahap validasi sensor, atau perlu dilakukan perbaikan pada tahap klasifikasi maupun ekstraksi fitur melalui iterasi pengembangan lebih lanjut.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan terhadap sistem estimasi tingkat polusi udara menggunakan metode segmentasi citra dan validasi sensor CO₂, maka dapat ditarik beberapa kesimpulan sebagai berikut.

1. Sistem segmentasi citra yang dikembangkan berhasil mengklasifikasikan tingkat polusi udara ke dalam tiga kategori yaitu Rendah, Sedang, dan Tinggi dengan menggunakan pendekatan analisis ruang warna HSV. Proses segmentasi dilakukan melalui pembangunan masker multi-zona yang mampu mendeteksi asap terang berakurasi rendah, area gelap pekat, sekaligus mengecualikan piksel langit biru dan vegetasi hijau yang bukan merupakan indikator polusi. Fitur-fitur yang diekstraksi meliputi rasio area polusi, entropi tekstur, standar deviasi intensitas, dan saturasi rata-rata yang kemudian digabungkan menjadi skor komposit sebagai dasar klasifikasi.
2. Evaluasi kinerja sistem menggunakan confusion matrix terhadap 100 data citra, sistem memperoleh akurasi keseluruhan sebesar 70,00%, Precision 37,90%, Recall 45,93%, F1 Score 36,70%. Hasil ini menunjukkan bahwa sistem mampu mengklasifikasikan sebagian besar citra dengan benar, meskipun terdapat ketimpangan kinerja yang cukup signifikan antar kelas.
3. Sistem analisis sensor CO₂ berhasil membaca indikator warna pada display alat sensor kualitas udara dengan menghitung proporsi piksel hijau, kuning, dan merah pada area crop display. Pendekatan ini menghasilkan klasifikasi yang merepresentasikan estimasi konsentrasi CO₂ di udara, yaitu kondisi aman pada dominasi warna hijau, kondisi sedang pada dominasi warna kuning, dan kondisi berbahaya pada dominasi warna merah.
4. Hasil validasi silang antara kedua metode terhadap 100 pasang citra menunjukkan tingkat kesepakatan sebesar 37%, yang masuk dalam kategori Perlu Perbaikan. Dari total 100 pasang, hanya 37 pasang yang menghasilkan klasifikasi konsisten antara segmentasi citra dan sensor CO₂.
5. Secara keseluruhan sistem yang dibangun telah berhasil mengintegrasikan dua pendekatan pengukuran kualitas udara yang berbeda dalam satu platform terpadu berbasis Python, dilengkapi dengan antarmuka grafis, fitur ekspor laporan Excel multi-sheet, visualisasi confusion matrix, dan modul validasi silang otomatis yang memudahkan analisis komparatif.

Saran

Berdasarkan keterbatasan yang ditemukan selama penelitian, beberapa saran diberikan untuk pengembangan ke depan:

1. Untuk meningkatkan akurasi segmentasi khususnya pada kelas Rendah dan Sedang yang performanya masih rendah, disarankan dilakukan kalibrasi ulang threshold HSV berdasarkan karakteristik dataset lokal yang lebih representatif. Nilai threshold yang digunakan saat ini bersifat umum dan belum dioptimalkan secara spesifik untuk kondisi pencahayaan dan jenis polutan yang dominan di wilayah penelitian. Penggunaan teknik pembelajaran mesin seperti Random Forest atau Support Vector Machine untuk tahap klasifikasi dapat dipertimbangkan sebagai alternatif pengganti klasifikasi berbasis skor komposit manual agar batas antar kelas dapat dipelajari secara adaptif dari data.

2. Dalam rangka meningkatkan tingkat kesepakatan validasi silang yang saat ini hanya 42%, disarankan penggunaan sensor multiparameter yang tidak hanya mengukur CO₂ tetapi juga parameter polutan lain seperti PM2.5, PM10, CO, dan NO₂. Parameter-parameter tersebut lebih berkorelasi langsung dengan kenampakan visual polusi yang ditangkap oleh kamera, sehingga diharapkan dapat meningkatkan konsistensi hasil antara kedua metode pengukuran.
3. Untuk pengembangan sistem lebih lanjut, antarmuka aplikasi dapat dikembangkan menjadi berbasis web atau mobile agar dapat diakses secara real-time dan diintegrasikan langsung dengan perangkat sensor IoT. Penambahan fitur notifikasi otomatis ketika tingkat polusi melebihi ambang batas tertentu juga akan meningkatkan nilai praktis sistem bagi pemangku kepentingan di bidang pemantauan kualitas lingkungan udara.

DAFTAR PUSTAKA

- Aakanksha, Arushi Seth, S.S. (2022) "Semantic Segmentation : A Systematic Analysis From State-of-the-Art Techniques to Advance Deep Networks," *Journal of Information Technology Research*, 15(1), pp. 1–28. Available at: <https://doi.org/10.4018/JITR.299388>.
- Abidah, D.Y., Hadi Wijoyo, S. and Rahman, K. (2025) "Pengaruh Platform Visual Studio Code Terhadap Hasil Belajar Siswa pada Mata Pelajaran Pemrograman Dasar Kelas X Jurusan Teknik Komputer dan Jaringan SMKN 3 Malang," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 3, pp. 1–10, Mar., *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 9(3), pp. 2548–964. Available at: <http://j-ptiik.ub.ac.id>.
- Adawiyah Ritonga and Yahfizham Yahfizham (2023) "Studi Literatur Perbandingan Bahasa Pemrograman C++ dan Bahasa Pemrograman Python pada Algoritma Pemrograman," *Jurnal Teknik Informatika dan Teknologi Informasi*, 3(3), pp. 56–63. Available at: <https://doi.org/10.55606/jutiti.v3i3.2863>.
- Adib Ulinuha El Majid and Reflan Nuari (2025) "Performance Comparison Of BERT Metrics and Classical Machine Learning Models (SVM, Naive Bayes) for Sentiment Analysis," *INOVTEK Polbeng - Seri Informatika*, 10(2), pp. 741–752. Available at: <https://doi.org/10.35314/wmh3rg23>.
- Affandi, R. et al. (2026) "ANALISI EFEKTIVITAS RUANG WARNA HSV DAN LAB DALAM PENGELOMPOKAN KONDISI DAUN TEH MENGGUNAKAN K-MEANS UNTUK PENENTUAN," *Djtechno: Jurnal Teknologi Informasi*, 7(1), pp. 123–133. Available at: <https://doi.org/10.46576/djtechno>.
- Ashillah, S. et al. (2025) "Implementasi Algoritma Pengolahan Citra Digital untuk Perbaikan Kualitas Gambar," *Jurnal SISKOM-KB (Sistem Komputer dan Kecerdasan Buatan)*, 8(3), pp. 274–280. Available at: <https://doi.org/10.47970/siskom-kb.v8i3.823>.
- Azis, H. et al. (2024) "Analisis Kinerja Algoritma Pembelajaran Mesin Ensemble Pada Dataset Multi Kelas Citra Jaffe," *Networking Engineering Research Operation*, 9(2), pp. 107–118. Available at: <https://doi.org/10.21107/nero.v9i2.27872>.
- Azizah, Z.J. and Lestari, K.S. (2026) "Gambaran umum kualitas udara berdasarkan indeks standar pemantauan udara di kabupaten sidoarjo," *Jurnal Kesehatan Tambusai*, 7(1), pp. 1877–1884. Available at: <https://doi.org/https://doi.org/10.31004/jkt.v7i1.56764>.
- Baran, A. and Zender-Świercz, E. (2025) "Concentration of PM2.5 and PM10 Particulate Matter in Various Indoor Environments: A Literature Review," *Atmosphere*, 17(1), p. 45. Available at: <https://doi.org/10.3390/atmos17010045>.
- Buwono, H.K. et al. (2025) "Koding QB64 dan python sebagai bahasa komunikasi teknik sipil dalam memahami alur pikir," *Konstruksia*, 16(2), pp. 146–156. Available at: <https://doi.org/https://doi.org/10.24853/jk.16.2.146-156>.
- Colléaux, Y. et al. (2025) "Air Pollution Monitoring Using Cost-Effective Devices Enhanced by Machine Learning," *Sensors*, 25(5), pp. 12–16. Available at: <https://doi.org/10.3390/s25051423>.
- Cynthia, G. and Qoiriah, A. (2024) "Klasifikasi Tingkat Kesiapan Kerja Mahasiswa Universitas Negeri Surabaya Menggunakan Algoritma C5.0," *Journal of Informatics and Computer Science (JINACS)*, 6(02), pp. 383–396. Available at: <https://doi.org/10.26740/jinacs.v6n02.p383-396>.
- Damen, D. et al. (2021) "Rescaling Egocentric Vision," *International Journal of Computer Vision (IJCV)* [Preprint]. Available at: <https://doi.org/10.5523/bris.2g1n6qdydwa9u22shpxqzp0t8m>.
- Gibran, M.K. (2025) "Penerapan Teknik Binarisasi Dan Segmentasi," *Jurnal Minfo Polgan*, 14, pp. 623–630.

- Guo, Y., Wu, X. and Shu, X. (2022) "Data Acquisition and Preparation for Dual-Reference Deep Learning of Image Super-Resolution," *IEEE Transactions on Image Processing*, 31(8), pp. 4393–4404. Available at: <https://doi.org/10.1109/TIP.2022.3184819>.
- Hardini, M. et al. (2023) "Image-based Air Quality Prediction using Convolutional Neural Networks and Machine Learning," *APT/ISI Transactions on Technopreneurship*, 5(1SP), pp. 109–123. Available at: <https://doi.org/10.34306/att.v5i1Sp.337>.
- Hendi Budianto, B.S. (2024) "Perancangan Sistem Monitoring Kualitas Udara dalam Ruangan Berbasis Internet of Things," *Jurnal Listrik, Instrumentasi dan Elektronika Terapan (JuLIET)*, 5(1), pp. 9–17. Available at: <https://doi.org/https://doi.org/10.22146/juliet.v5i1.87423>.
- Humaidi, M.R. and Maulani, A. (2023) "Klasifikasi Naïve Bayes Dan Confusion Matrix Pada Pengguna Aplikasi E-Commerce Di Play Store," *Jurnal Ilmiah Informatika*, 8(2), pp. 132–139. Available at: <https://doi.org/https://doi.org/10.35316/jimi.v8i2.132-139>.
- Inal Achyar, M., Yasir, A. and Satria, W. (2023) "Pengolahan Citra Berwarna Dalam Mendeteksi Asap Menggunakan Metode Segmentasi Foreground Dan Color Models," *SISKOMTI: Jurnal Sistem Informasi Komputer dan Teknologi Informasi*, 5(2), pp. 1–10. Available at: <https://doi.org/10.54342/fmy91476>.
- Kurniawan, A. (2017) "PENGUKURAN PARAMETER KUALITAS UDARA," *Teknosains*, 7(1), pp. 1–13. Available at: <https://doi.org/10.22146/teknosains.34658>.
- Liaw, J.J. et al. (2020) "PM2.5 concentration estimation based on image processing schemes and simple linear regression," *Sensors (Switzerland)*, 20(8), pp. 1–13. Available at: <https://doi.org/10.3390/s20082423>.
- Lin, T.C. et al. (2023) "Unmasking air quality: A novel image-based approach to align public perception with pollution levels," *Environment International*, 181(October), p. 108289. Available at: <https://doi.org/10.1016/j.envint.2023.108289>.
- Liu, J. et al. (2024) "A transformer boosted UNet for smoke segmentation in complex backgrounds in multispectral LandSat imagery," *Remote Sensing Applications: Society and Environment*, 36, pp. 1–17. Available at: <https://doi.org/10.1016/j.rsase.2024.101283>.
- Liu, X. et al. (2024) "A Deep-Learning Approach to Detect and Classify Heavy-Duty Trucks in Satellite Images," *IEEE Transactions on Intelligent Transportation Systems*, 25(10), pp. 13323–13338. Available at: <https://doi.org/10.1109/TITS.2024.3431452>.
- Masnur et al. (2022) "Menghitung Kendaraan Di Parkiran Dengan Metode Support Vector Machine (SVM)," *Jurnal Ilmiah Teknik Unida*, 3(2), pp. 140–148. Available at: <https://doi.org/10.55616/jitu.v3i2.369>.
- Murani, N.A. et al. (2025) "Analisis Penggunaan Visual Studio Code Sebagai Editor Pemrograman Terbaik Untuk Pemula Pada Mahasiswa Informatika Universitas Nurul Huda," *Jurnal Rekayasa Informatika (Jrit)*, 2(2), pp. 1–10. Available at: <https://ejournal.pei.ac.id/index.php/JRIT/index>.
- Najwa, N. et al. (2025) "Pencemaran Lingkungan : Pemahaman dan Penanggulangan Pencemaran," *Deposisi*, 3(2), pp. 263–269.
- Pinem, T.H. and Putra, Z.P. (2025) "Evaluasi Kinerja Algoritma Klasifikasi Deep Learning dalam Prediksi Diabetes," *Jurnal Ilmiah FIFO*, 17(1), p. 17. Available at: <https://doi.org/10.22441/fifo.2025.v17i1.003>.
- Ramadas, M. and Abraham, A. (2023) "Segmentation on remote sensing imagery for atmospheric air pollution using divergent differential evolution algorithm," *Neural Computing and Applications*, 35(5), pp. 3977–3990. Available at: <https://doi.org/10.1007/s00521-022-07922-x>.
- Rizki Akbar Pratama, Barry Ceasar, S. (2025) "Penerapan Learning Vector Quantization Dalam Pengolahan Citra Digital Untuk Deteksi Penyakit Kulit," *Jurnal Computer Science and Information Technology (CoSciTech)*, 6(2), pp. 148–157. Available at: <https://doi.org/https://doi.org/10.37859/coscitech.v6i2.9270>.
- Simanullang, L. et al. (2013) "Penentuan Kondisi Udara(Lingkungan) Menggunakan Citra Inframerah," *ELECTRICIAN - Jurnal Rekayasa dan Teknologi Elektro Penentuan*, 7(2), pp. 78–85.
- Sohrabi, Z. and Maleki, J. (2025) "Fusing satellite imagery and ground-based observations for PM 2.5 air pollution modeling in Iran using a deep learning approach," *Scientific Report*, pp. 1–16. Available at: <https://doi.org/https://doi.org/10.1038/s41598-025-05332-2>.