

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi kecerdasan buatan khususnya dalam bidang computer vision telah berkembang secara signifikan seperti pada system pengenalan ekspresi wajah (Facial Emotion Recognition/FER). Seperti yang kita ketahui bahwa ekspresi wajah merupakan salah satu media utama manusia berkomunikasi (Agung et al., 2024). Oleh karena itu, sistem FER memiliki peran yang penting dalam berbagai bidang seperti Pendidikan, Kesehatan mental (Psikologis), interaksi manusia dan computer hingga sistem keamanan yang canggih.

Pendekatan modern pada FER pada umumnya menggunakan Convolutional Neural Network (CNN) karena kemampuannya dalam melakukan ekstraksi fitur secara otomatis dan secara end to end (Agung et al., 2024; Gunawan et al., 2025). Namun, performa CNN sangat dipengaruhi oleh kualitas citra yang akan di deteksi. CNN diketahui sensitive terhadap distorsi gambar dan juga noise pada gambar yang dapat menurunkan akurasi klasifikasi secara signifikan (Ziyadinov & Tereshonok, 2023).

Permasalahan ini menunjukkan bahwa preprocessing citra memiliki peran krusial dalam meningkatkan robust pada model deep learning. salah satu pendekatan yang dapat digunakan adalah menggunakan filtering. Low Pass Filter (LPF) diketahui mampu mereduksi komponen frekuensi tinggi atau noise pada gambar (Nguyen-Tat et al., 2025; Ziyadinov & Tereshonok, 2023). sebaliknya High Pass Filter (HPF) menekankan komponen frekuensi tinggi untuk menajamkan detail dan tepi pada gambar (Barral, 2022), sementara itu Band Pass Filter (BPF) adalah filtering seimbang antara detail dan reduksi noise (Feng et al., 2021).

Meskipun berbagai penelitian telah membahas robustness CNN terhadap noise dan adversarial distortion, Sebagian besar penelitian tersebut hanya berfokus pada peningkatan robustness saja, bukan pada evaluasi sistematis pengaruh masing masing jenis filter frekuensi terhadap performa FER (Pratama & Azhar, 2025; Raj

& Demirkol, 2025; Ziyadinov & Tereshonok, 2023) selain itu, pendekatan filtering juga sering digunakan untuk tujuan peningkatan kualitas visual seperti pada klasifikasi menggunakan bilateral filter. (Awarayi et al., 2024)

Bilateral filter sendiri merupakan teknik filtering yang mampu mengurangi noise pada gambar hampir mirip seperti Low Pass Filter (LPF) (Awarayi et al., 2024), oleh karena itu bilateral filter biasanya digunakan sebagai baseline filter dalam sistem klasifikasi citra medis. Oleh sebab itu, pada penelitian ini bilateral filter digunakan sebagai baseline pembanding untuk menilai efektivitas dari ketiga filter utama (LPF,HPF, dan BPF)

Selain aspek filtering, tahapan deteksi dan alignment wajah juga berperan penting dalam FER. Seperti Multi Task Cascaded Convolutional Neural Network (MTCNN) dikenal sebagai metode yang dikhususkan untuk mendeteksi wajah secara efektif dan mampu mendeteksi landmark wajah secara akurat (P. Raj, 2024). Namun dalam penelitian ini MTCNN digunakan sebagai tahap awal pendeteksian wajah sebelum nantinya akan di klasifikasikan menggunakan model atau metode InceptionV3.

Jadi pada tahap klasifikasi, penelitian ini menggunakan InceptionV3 dengan pendekatan transfer learning. Arsitektur InceptionV3 dikenal sangat efisien dan memiliki performa yang mumpuni dalam klasifikasi citra (Gunawan et al., 2025; Muttaqin & Widodo, 2025; Pratama & Azhar, 2025). Studi sebelumnya juga menunjukkan bahwa Fine Tuning InceptionV3 pada dataset Emognition mampu menghasilkan akurasi tinggi dalam pengenalan ekspresi wajah (Agung et al., 2024; Saganowski et al., 2022).

Dataset yang digunakan adalah Emognition Dataset, Emognition ini juga tidak hanya mencakup emosi dasar, tetapi juga empat emosi positif spesifik seperti enthusiasm, awe, dan liking (Saganowski et al., 2022). Emosi positif ini penting karena masing masing memiliki karakteristik ekspresi yang berbeda (Saganowski et al., 2022). Penelitian oleh (Agung et al., 2024) menunjukkan bahwa dataset ini masih relative jarang dieksplorasi untuk perkembangan dari model FER yang berbasis citra.

Selain mengevaluasi performa klasifikasi penelitian ini juga menilai dampak filtering terhadap kualitas citra menggunakan tiga point utama yaitu Mean Squared Error (MSE), Peak Signal Noise ratio (PSNR), dan juga Structural Similarity Index Measure (SSIM). MSE merupakan metrik penghitungan yang digunakan untuk mengukur rata rata kesalahan antara citra asli dengan citra hasil pemrosesan, sedangkan PSNR menghitung rasio antara sinyal maksimum dan noise dalam skala desibel sebagai indikator, lalu untuk SSIM digunakan untuk mengevaluasi kesamaan struktur, luminasi dan kontras pada citra (Sara)

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana dampak penerapan filter (Bilateral Filter, Low Pass Filter, High Pass Filter, dan Band Pass Filter) terhadap kualitas gambar wajah yang diukur menggunakan metrik PSNR, MSE, dan SSIM?
2. Bagaimana pengaruh filtering domain frekuensi terhadap tingkat kepercayaan (confidence score) dan akurasi deteksi emosi pada sistem pengenalan ekspresi wajah menggunakan arsitektur MTCNN dan InceptionV3?
3. Tipe filter manakah yang memberikan keseimbangan optimal antara reduksi noise dan preservasi fitur untuk pengenalan ekspresi wajah pada dataset Emognition?

1.3 Tujuan Penelitian

Adapun tujuan yang ingin dicapai melalui penelitian ini adalah:

1. Menganalisis tentang dampak tiga jenis filter utama (LPF, HPF, BPF) dan juga satu baseline filter (Bilateral filter) terhadap performa FER menggunakan dataset Emognition
2. Untuk meminimalisir terjadinya deployment disaster atau kesalahan terjadinya diagnosis pada sistem FER
3. Untuk mengetahui apakah dengan preprocessing gambar ini dapat meningkatkan performa dari FER atau tidak untuk efisiensi daripada harus menggunakan model yang lebih besar dan mahal

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat Akademis:
 - a. Memberikan kontribusi terhadap pengembangan literatur tentang preprocessing gambar dalam sistem pengenalan ekspresi wajah, khususnya dalam konteks filtering domain frekuensi.
 - b. Menyediakan kerangka evaluasi sistematis untuk menilai dampak filtering terhadap kinerja model deep learning dalam pengenalan emosi, yang dapat dijadikan acuan metodologis bagi penelitian lanjutan.
 - c. Memberikan wawasan empiris tentang pengenalan emosi bernuansa (nuanced emotions) seperti amusement, awe, enthusiasm, dan liking yang masih jarang diteliti dalam literatur FER.
2. Manfaat Praktis:
 - a. Menyediakan panduan praktis bagi pengembang sistem FER dalam memilih teknik preprocessing yang optimal untuk meningkatkan akurasi dan kepercayaan prediksi.
 - b. Meningkatkan reliabilitas sistem pengenalan ekspresi wajah di dunia nyata dengan mendemonstrasikan pentingnya pemilihan filter yang tepat untuk berbagai kondisi kualitas gambar.
 - c. Memberikan landasan bagi pengembangan aplikasi berbasis emotion analysis yang lebih robust, seperti sistem pemantauan kesehatan mental, evaluasi pengalaman pengguna, dan human-computer interaction.

1.5 Batasan Masalah

Agar penelitian terarah dan sesuai dengan tujuan, batasan masalah dalam penelitian ini dirumuskan sebagai berikut:

1. Dataset yang digunakan adalah Emognition Wearable Dataset 2020 yang terdiri dari 3.200 gambar ekspresi wajah dengan sepuluh kategori emosi yang mencakup 437 gambar amusement, 290 gambar anger, 438 gambar awe, 290 gambar disgust, 291 gambar enthusiasm, 290 gambar fear, 290 gambar liking, 438 gambar neutral, dan 436 gambar sadness).

2. Arsitektur yang digunakan terbatas pada Multi-Task Cascaded Convolutional Networks (MTCNN) untuk deteksi wajah dan InceptionV3 pre-trained pada ImageNet untuk klasifikasi emosi.
3. Teknik filtering yang dievaluasi terbatas pada empat tipe filter: Bilateral filter, Low Pass Filter (LPF), High-Pass Filter (HPF), dan Band-Pass Filter (BPF) dengan parameter kernel Gaussian yang telah ditentukan.
4. Metrik evaluasi kualitas gambar yang digunakan terbatas pada Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), dan Structural Similarity Index (SSIM).
5. Evaluasi kinerja model difokuskan pada metrik akurasi klasifikasi (precision, recall, F1-score) dan tingkat kepercayaan prediksi (confidence score)..
6. Pelatihan model menggunakan strategi transfer learning tiga fase dengan total 105 epoch, dengan parameter yang telah ditentukan untuk setiap fase (frozen layers, unfrozen layers, dan full model training)..

