

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pada 14 Maret 2020, Pemerintah Indonesia mewajibkan penggunaan masker di fasilitas publik sebagai respons terhadap pandemi COVID-19. Kebijakan serupa juga diberlakukan di banyak negara, sehingga jutaan individu beraktivitas di ruang publik dengan sebagian wajah tertutup. Kondisi ini berdampak pada sistem keamanan yang bergantung pada teknologi pengenalan wajah otomatis. Ketika sebagian besar fitur wajah tertutup, performa sistem pengenalan wajah mengalami penurunan yang signifikan (Ngan et al., 2020). Permasalahan ini tidak sepenuhnya berakhir setelah pandemi, Hingga saat ini, penggunaan masker, cadar, niqab, maupun helm tetap menjadi hal yang umum dalam berbagai konteks budaya, keagamaan, dan keselamatan. Hal tersebut menjadikan keterbatasan pengenalan wajah sebagai tantangan berkelanjutan.

Urgensi yang serupa juga terlihat pada kasus keamanan fasilitas berbasis gender. Pada Maret 2024, seorang pria di Makassar berhasil memasuki area khusus jemaah perempuan di Masjid Jannatul Firdaus dengan menyamar menggunakan cadar dan melakukan aksi tersebut beberapa kali sebelum akhirnya terungkap (Detik News, 2024). Identitas pelaku baru diketahui setelah jemaah mencurigai perilakunya dan melapor kepada petugas. Peristiwa ini menunjukkan bahwa pengawasan visual manual tidak selalu memadai sebagai mekanisme pengamanan, sehingga diperlukan sistem identifikasi otomatis yang tetap dapat berfungsi sebagai pencegahan awal meskipun wajah seseorang tertutup. Sejumlah metode telah

dikembangkan untuk klasifikasi jenis kelamin, seperti pengenalan wajah (*face recognition*) (Saragih et al., 2019), analisis suara (Patil & Rao, 2021), dan pemindaian sidik jari (Malathi et al., 2021). Namun, metode berbasis wajah memiliki keterbatasan dalam kondisi tertentu, seperti ketika subjek mengenakan masker, cadar, atau helm, yang menyebabkan sistem klasifikasi *gender* berbasis wajah menjadi kurang efektif (Zhang et al., 2022). Dalam situasi ini, citra mata manusia menjadi fitur utama yang tetap terlihat dan dapat digunakan untuk klasifikasi Jenis Kelamin (Kuehlkamp et al., 2021). Penerapan *deep learning* pada citra mata juga terus berkembang untuk berbagai kebutuhan analisis medis, seperti klasifikasi tingkat keparahan *diabetic retinopathy* menggunakan arsitektur ResNet50 yang dipadukan dengan teknik *image augmentation*, sehingga menunjukkan bahwa citra mata memiliki karakteristik visual yang kaya untuk dipelajari oleh model pembelajaran mendalam (Maysanjaya et al., 2025).

Sejumlah penelitian telah dilakukan dalam klasifikasi jenis kelamin berbasis citra mata dengan teknik pembelajaran mesin dan pembelajaran mendalam.. Chalvin Kurniawan (2022) dalam penelitiannya “Perbandingan Metode *K-Nearest Neighbor* dan *Naïve Bayes* untuk Klasifikasi *Gender* Berdasarkan Mata”, menggunakan dataset sebanyak 15.059 gambar mata, dengan hasil *accuracy* 68,81% untuk *K-Nearest Neighbor* (KNN) dan 62,58% untuk *Naïve Bayes*. Hasil ini menunjukkan bahwa metode berbasis *machine learning* ini masih kurang efektif.

Sementara itu, penelitian oleh Nur Aini et al. (2022) menggunakan metode *Convolutional Neural Network* (CNN), *InceptionV3*, dan *MobileNet* dalam klasifikasi *gender* berbasis citra mata, di mana *MobileNet* menunjukkan *accuracy* tertinggi sebesar 93,90%. Meskipun CNN telah menunjukkan performa yang cukup

baik, metode ini masih memiliki beberapa keterbatasan, terutama dalam menangani variasi spasial yang kompleks. (Dosovitskiy et al., 2020).

Perkembangan teknologi *deep learning*, *Vision Transformer* (ViT) muncul sebagai alternatif yang mampu memproses informasi spasial global dengan efektif melalui mekanisme *self-attention* (Dosovitskiy et al., 2020). Studi yang dilakukan oleh Touvron et al. (2021) menunjukkan bahwa *Vision Transformer* (ViT), mampu mencapai performa kompetitif pada dataset sedang hingga besar, dan menunjukkan hasil yang sangat baik dalam berbagai tugas klasifikasi citra. Sampai saat ini belum ada penelitian yang secara spesifik meneliti *Vision Transformer* untuk klasifikasi jenis kelamin berbasis citra mata manusia. Hal ini menjadi celah penelitian yang penting untuk dieksplorasi lebih lanjut dalam upaya meningkatkan *accuracy* dan generalisasi sistem klasifikasi.

1.2. Identifikasi Masalah

Metode klasifikasi Jenis Kelamin berbasis wajah memiliki keterbatasan dalam kondisi dimana subjek mengenakan masker, cadar, atau helm, sehingga biometrik citra mata menjadi alternatif yang lebih baik. Namun, penelitian Chalvin Kurniawan (2022) menunjukkan bahwa metode *K-Nearest Neighbor* (KNN) hanya mencapai *accuracy* 68,81%, dan *Naïve Bayes* 62,58%, yang mengindikasikan bahwa metode ini masih kurang efektif dalam klasifikasi untuk digunakan.

Deep learning berbasis *Convolutional Neural Network* (CNN) telah meningkatkan *accuracy*, dengan *MobileNet* mencapai 93,90% (Nur Aini et al., 2022). Namun, CNN masih memiliki keterbatasan dalam menangani variasi spasial global kompleks.

Vision Transformer (ViT) telah terbukti sangat baik dalam beberapa tugas klasifikasi citra (Dosovitskiy et al., 2020) mekanisme *self-attention* yang baik dalam menangkap hubungan spasial secara global sekaligus. Namun, belum ada penelitian yang meneliti ViT untuk klasifikasi Jenis Kelamin berbasis citra mata manusia. Sehingga diperlukan penelitian yang berguna untuk mengeksplorasi tingkat efektivitas ViT dalam meningkatkan *accuracy* klasifikasi Jenis Kelamin berbasis citra mata manusia.

1.3. Batasan Masalah

Dari pemaparan permasalahan tersebut dapat ditentukan batasan masalah sebagai berikut.

- a. Berdasarkan rumusan masalah, penyelesaian dari permasalahan yang ada yakni terfokus pada mengklasifikasi *Jenis Kelamin* berbasis pengolahan citra mata.
- b. Tingkat *accuracy* dan efisiensi akan diambil dari pengujian proses klasifikasi Jenis Kelamin berbasis citra mata dengan menggunakan metode *Transformer* vision berdasarkan evaluasi *Confusion Matrix* dan fungsi *time.time()*.
- c. Penelitian ini tidak menganalisis pengaruh usia, atau penggunaan kacamata, yang bisa memengaruhi hasil klasifikasi.
- d. Penelitian ini menggunakan varian model *pretrained Vision Transformer Base-16*.

1.4. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, penelitian ini memiliki rumusan masalah sebagai berikut.

- a. Bagaimana implementasi *Vision Transformer* dalam klasifikasi Jenis Kelamin berbasis citra mata?
- b. Seberapa akurat dan efisien metode *Vision Transformer* dalam klasifikasi Jenis Kelamin berbasis citra mata?

1.5. Tujuan Penelitian

Dari uraian penulisan rumusan masalah yang ada dapat diambil beberapa tujuan dari penelitian yakni:

- a. Penelitian ini bertujuan untuk mengimplementasikan model klasifikasi Jenis Kelamin berbasis *Vision Transformer* dalam sebuah model berbasis pengolahan citra.
- b. Penelitian ini bertujuan untuk mengevaluasi *accuracy* dan efisiensi model *Vision Transformer* dalam klasifikasi Jenis Kelamin berbasis citra mata manusia.

1.6 Manfaat Penelitian

Dari uraian penulisan tujuan dari rumusan masalah yang ada dapat diambil beberapa tujuan dari penelitian yakni:

- a. Manfaat Akademis
 - 1) Penelitian ini diharapkan mampu memberikan kontribusi terhadap pengembangan kajian pada bidang pengolahan citra digital dan kecerdasan

buatan, khususnya yang berkaitan dengan pemanfaatan Vision Transformer dalam proses klasifikasi jenis kelamin.

- 2) Temuan yang dihasilkan dari penelitian ini diharapkan dapat dimanfaatkan sebagai salah satu acuan bagi penelitian selanjutnya dalam mengeksplorasi maupun mengadaptasi pendekatan Vision Transformer pada permasalahan yang berbeda atau pada dataset dengan karakteristik yang lebih beragam dan kompleks..

b. Manfaat Praktis

- 1) Implementasi metode *Vision Transformer* diharapkan dapat meningkatkan *accuracy* dan efisiensi dalam proses klasifikasi Jenis Kelamin
- 2) Dengan adanya sistem klasifikasi Jenis Kelamin berbasis citra mata, teknologi ini dapat diterapkan dalam berbagai aplikasi praktis seperti keamanan, dan interaksi manusia-komputer yang maju.

