

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Ekspresi wajah memegang peran yang krusial dalam interaksi yang dilakukan oleh manusia, baik secara lisan maupun non-lisan (Dewi et al., 2024). Banyak orang menghadapi kesulitan dalam mengekspresikan perasaan mereka secara jelas, yang sering kali menyebabkan kesalahpahaman atau konflik. Melalui ekspresi wajah, manusia dapat mendukung dan memperkuat emosi yang ingin disampaikan, baik itu bahagia, sedih, marah, takut, atau jijik. Ekspresi-emosi ini ditunjukkan melalui perubahan bagian otot wajah dan tampilan visual wajah secara keseluruhan. Di era digital, pengenalan ekspresi wajah dengan bantuan *artificial intelligence* (AI) menjadi sangat bermanfaat untuk membantu memahami dan menganalisa emosi seseorang.

Dalam proses pengenalan ekspresi wajah, komputer menghadapi tantangan yang lebih besar dibandingkan manusia. Pemilihan fitur yang tepat berperan penting dalam mencapai akurasi optimal dalam tugas klasifikasi ekspresi wajah. Berdasarkan model komunikasi Mehrabian (Ramzani Shahrestani et al., 2024), lebih dari setengah makna yang disampaikan manusia dalam sebuah interaksi datang dari ekspresi wajah (55%), diikuti oleh suara (38%) dan bahasa (7%). Terdapat dua pendekatan utama dalam ekstraksi fitur untuk mengenali ekspresi wajah, yaitu fitur geometris yang mengandalkan titik-titik *landmark* wajah dan fitur tampilan yang merepresentasikan tekstur otot wajah melalui perubahan kerutan dan lipatan (Fan et al., 2023). Penelitian oleh Eisenberg dan Alpres (Esté Jaloveckas & Granero, 2024) menunjukkan bahwa mata dan mulut memainkan peranan penting dalam identifikasi emosi, di mana eksperimen menggunakan teknologi *eye tracker* menunjukkan kedua area ini sebagai indikator emosi utama.

Pendekatan terbaru menggunakan *Convolutional Neural Networks* (CNN) dengan arsitektur *Siamese Net* (SNN) memungkinkan pemrosesan dengan dua input atau lebih secara bersamaan melalui lapisan *embedding* untuk mengukur kemiripan antara input. Metode ini mengukur jarak antara *input* dengan formula *Euclidean* untuk mengetahui seberapa dekat jarak satu input dari lainnya (Xu et al., 2022).

Selain itu *Hierarchical Feature with Three Channel Convolutional Neural Networks* (HFT-CNN) (He et al., 2023), telah diusulkan sebagai metode untuk meningkatkan pengenalan ekspresi wajah dengan melakukan ekstraksi fitur dari tiga bagian utama wajah, yaitu seluruh wajah, mata dan alis, mulut secara independent dalam *channel* yang berbeda menggunakan CNN. Meskipun HFT-CNN memberikan hasil yang menjanjikan, algoritma yang diusulkan ingin mempertahankan efisiensi dengan jumlah *layer* yang terbatas dan jumlah parameter yang relatif kecil sehingga ada kemungkinan tidak bisa mencapai akurasi yang tinggi jika dilakukan komparasi dengan arsitektur yang lebih kompleks. Di sisi lain, SNN memiliki keterbatasan dalam melakukan klasifikasi dan tidak dapat digunakan sebagai acuan untuk akurasi klasifikasi secara langsung.

Oleh karena itu, penulis mengusulkan metode *Klasifikasi Ekspresi Wajah Menggunakan Siamese Network dan HFT-CNN Pada Citra*. Penggabungan kedua metode ini diharapkan mampu meningkatkan ketepatan dan kinerja dalam ekspresi wajah, memanfaatkan kelebihan SNN dalam mengukur kemiripan dan HFT-CNN dalam ekstraksi fitur secara mendetail dari bagian wajah yang krusial.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah yang dapat diajukan adalah sebagai berikut:

1. Bagaimana implementasi arsitektur HFT-CNN dan SNN dapat digunakan dalam mengenali dan mengklasifikasi ekspresi wajah ?
2. Seberapa tinggi tingkat ketepatan yang dapat dicapai dalam pengklasifikasian ekspresi wajah dengan menggunakan metode gabungan HFT-CNN dan SNN ?

## **1.3 Tujuan Penelitian**

Merujuk pada rumusan masalah yang telah dijelaskan, penelitian ini bertujuan untuk:

1. Mengimplementasikan arsitektur HFT-CNN dan SNN untuk mengenali dan mengklasifikasikan ekspresi wajah secara efektif.
2. Mengukur tingkat ketepatan klasifikasi ekspresi wajah yang dihasilkan dari penerapan metode gabungan HFT-CNN dan SNN.

## 1.4 Batasan dan Asumsi Penelitian

Berdasarkan tujuan penelitian yang dibahas, berikut adalah batasan dan asumsi dari penelitian yang dilakukan:

### Batasan Penelitian

1. *Dataset* ekspresi wajah: Penelitian ini hanya menggunakan *dataset* citra ekspresi wajah yang diambil dari internet (*Bing repository*) dengan tujuh kelas emosi, yaitu *happy* (bahagia), *sad* (sedih), *angry* (marah), *fear* (takut), *disgust* (jijik), *neutral* (netral), *surprise* (terkejut). Ciri ekspresi yang digunakan, baik yang dikeluarkan secara alami maupun buatan (diperagakan) tetap dijadikan sebagai acuan dalam pembuatan data dan deskripsi bentuk emosi sesuai dengan penelitian sebelumnya (He et al., 2023).
2. Arsitektur model: Penelitian difokuskan pada arsitektur SNN yang digabungkan dengan HFT-CNN sebagai *backbone* untuk melakukan klasifikasi. Arsitektur lain tidak digunakan dalam penelitian ini.

### Asumsi Penelitian

1. Keberhasilan ekstraksi fitur: Diasumsikan bahwa arsitektur HFT- CNN yang digunakan mampu melakukan ekstraksi fitur penting dari area wajah (seluruh wajah, mata dan alis, mulut) dengan baik.
2. Keterandalan SNN: SNN diasumsikan mampu membedakan kemiripan antar ekspresi dengan menggunakan jarak *Euclidean* secara akurat.

## 1.5 Manfaat Penelitian

Melalui penelitian ini diharapkan dapat diperoleh manfaat berikut:

### 1. Manfaat Teoritis

Melalui penelitian ini, diharapkan dapat memberikan kontribusi dalam pengembangan metode pengenalan ekspresi wajah, khususnya melalui penerapan arsitektur HFT-CNN yang mampu mengekstraksi fitur wajah secara detail.

### 2. Manfaat Praktis

Mendukung pengembangan sistem pengenalan *mikro ekspresi* untuk keperluan deteksi kebohongan dalam konteks psikologi.

Memberikan kontribusi pada bidang interaksi manusia-komputer (*Human-*

*Computer Interaction*), khususnya dalam pengembangan teknologi *Augmented Reality* (AR) untuk pembuatan *avatar* yang mampu menangkap bentuk wajah secara *real-time*, serta pengembangan *virtual assistant* yang mampu memahami emosi seseorang.

## **1.6 Sistematika Penulisan**

Penulisan penelitian ini terdiri dari lima bab yang dirancang secara terstruktur untuk memastikan pembahasan terfokus pada inti permasalahan. Setiap bab saling berkaitan dan mendukung satu sama lain. Adapun susunan penulisan penelitian tugas akhir ini adalah sebagai berikut:

### **BAB 1 PENDAHULUAN**

Bab ini menyajikan gambaran umum mengenai latar belakang penelitian, rumusan masalah, batasan ruang lingkup penelitian, manfaat yang diharapkan dari penelitian, serta struktur sistematika penulisan yang digunakan.

### **BAB 2 LANDASAN TEORI**

Bab ini memaparkan teori yang mendasari penelitian termasuk konsep pengenalan ekspresi wajah, arsitektur HFT-CNN, SNN, serta pendekatan lainnya. Selain itu juga dibahas mengenai metrik yang digunakan untuk mengukur keakuratan klasifikasi.

### **BAB 3 METODOLOGI PENELITIAN**

Bab ini menguraikan tahapan-tahapan yang dilakukan dalam penelitian, mulai dari studi literatur, pengumpulan data, implementasi metode kombinasi HFT-CNN dan SNN, hingga tahap pengujian dan evaluasi hasil penelitian.

### **BAB 4 HASIL DAN PEMBAHASAN**

Bab ini memaparkan hasil implementasi model, analisis pengujian, serta pembahasan terkait performa dan akurasi model terhadap klasifikasi ekspresi wajah.

### **BAB 5 KESIMPULAN DAN SARAN**

Bab ini berisi kesimpulan dari hasil penelitian serta saran untuk pengembangan atau penelitian lanjutan di masa depan.