BAB I

PENDAHULUAN

1.1 Latar Belakang

Manusia menggunakan ekspresi wajah sebagai salah satu sinyal untuk mengungkapkan perasaan dalam komunikasi sehari-hari. Kondisi fisik atau mental seseorang pada suatu waktu dapat diketahui melalui analisis ekspresi wajah. Dengan demikian, pengenalan ekspresi wajah memegang peranan penting di berbagai bidang, seperti sistem autopilot, interaksi antara manusia dan komputer, layanan medis, serta aplikasi lain yang berkaitan dengan ekspresi wajah.[1]. Contohnya pada dunia pelayanan, menurut Abi Prasidi yang mengatakan bahwa ekspresi wajah dapat mendukung komunikasi verbal agar dapat dipercaya oleh lawan bicara untuk mendukung situasi [2]. Maka itu perlu pengenalan ekspresi wajah pada bidang komputer memahami emosi dan kondisi seseorang melalui analisis ekspresi wajah mereka melalui mesin untuk meningkatkan kualitas pelayanan.

Salah satu tantangan utama dalam pengenalan ekspresi wajah adalah kemampuan untuk menemukan dan memahami area wajah yang relevan secara efektif. Dalam beberapa tahun terakhir, pendekatan tradisional seperti deteksi titik-titik kunci pada wajah dan pemodelan objek telah banyak digunakan untuk menyelesaikan permasalahan ini. Namun, dalam situasi dunia nyata, pengenalan ekspresi wajah memiliki berbagai tantangan. Faktor seperti pencahayaan dinamis, oklusi (seperti penggunaan masker atau rambut yang menutupi wajah), pose kepala yang bervariasi, serta kompleksitas latar belakang menjadikan pengenalan ekspresi wajah lebih sulit untuk dilakukan secara akurat. Dalam skenario semacam ini, metode tradisional sering kali tidak mampu memberikan hasil yang memuaskan[3].

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence*) sangat cepat, khususnya dalam ranah pembelajaran mesin (*Machine Learning*). Sebagai salah satu cabang AI, machine learning memungkinkan komputer untuk memperoleh pengetahuan dari data tanpa harus diberi instruksi secara langsung. Dengan memanfaatkan algoritma yang mampu mengenali pola-pola dalam data,

pembelajaran mesin dapat diterapkan dalam berbagai bidang seperti pengenalan suara, pengolahan teks, hingga pengenalan gambar. Pengenalan gambar dapat dilakukan melalui berbagai pendekatan, salah satunya adalah *Deep Learning* yang merupakan bagian dari *Machine Learning*. Dalam *deep learning*, pengenalan citra sering memanfaatkan algoritma *Convolutional Neural Network* (CNN).

Convolutional Neural Network (CNN) menunjukkan performa paling baik dalam tugas pengenalan citra, karena arsitekturnya meniru cara kerja visual cortex pada manusia, sehingga mampu mengolah informasi visual secara efektif. Algoritma CNN memang khusus dirancang untuk pengenalan dan klasifikasi gambar. CNN terdiri dari beberapa lapisan (*layers*) yang berfungsi untuk mengekstraksi fitur penting dari gambar dan menghasilkan nilai klasifikasi yang digunakan untuk menentukan kategori gambar tersebut. [4].

Penelitian terdahulu dilakukan oleh Zhao dkk., yang menggunakan arsitektur *Multi-Level Convolutional Neural Network* (Multi-Level CNN) untuk klasifikasi ekspresi wajah. Penelitian ini menggunakan dataset FER2013 menunjukkan bahwa pendekatan Multi-Level CNN dapat meningkatkan akurasi pengenalan ekspresi wajah dibandingkan dengan metode CNN standar dan mampu mengekstraksi fitur dari berbagai tingkat abstraksi (*low-level, mid-level, high-level*) sehingga menghasilkan representasi fitur yang kaya dan lebih akurat dalam mendeteksi ekspresi wajah. Penelitian tersebut menggunakan *Multi-level Convolutional Neural Networks* (MLCNN) mendapatkan akurasi hingga 74,09% [3].

Penelitian lain telah menunjukkan bahwa penggunaan jaringan saraf tiruan (*neural networks*) secara signifikan meningkatkan kemampuan sistem dalam memahami emosi manusia melalui ekspresi wajah. Seperti penelitian yang dilakukan oleh Jie Shao dan Yongsheng Qian dengan *Dual-branch CNN* yang menggabungkan fitur tekstur tradisional seperti *Local Binary Pattern* (LBP) dengan fitur *deep learning* melalui dua jalur jaringan paralel, memungkinkan ekstraksi informasi yang lebih kaya. Sementara itu, *Pretrained CNN* menerapkan *transfer learning* dengan memanfaatkan model CNN yang telah dilatih sebelumnya untuk menangani keterbatasan data pelatihan. Model-model ini diuji menggunakan dataset CK+, BU-3DEF, dan FER2013, serta dibandingkan dengan metode *state-*

of-the-art lainnya. Hasil penelitian menunjukkan bahwa pendekatan *Dual-branch CNN* yang menggabungkan LBP dan CNN mampu meningkatkan akurasi pengenalan ekspresi wajah dibandingkan dengan model lain pada dataset CK+ yakni sebesar 85.71%. Namun pada dataset FER2013 akurasinya sebesar 54.64% [5].

Terbukti bahwa model dengan menggunakan CNN dan LBP dapat meningkatkan akurasi. Tidak hanya itu, pada penelitian yang dilakukan oleh Ehsan Bahmanabady, Maryam Imani dan Hassan Ghassemian menerangkan bahwa CNN dengan model Multi Level CNN dan LBP dapat menjadi ekstraksi fitur saja, untuk klasifikasi menggunakan SVM. Model yang diusulkan terdiri dari tiga lapisan CNN, di mana fitur yang diekstraksi dari setiap lapisan digunakan secara aktif untuk klasifikasi. Selain itu, LBP digunakan sebagai deskriptor tekstur untuk mengekstraksi fitur diskriminatif dari struktur spasial gambar wajah. Eksperimen dilakukan pada dataset *Japanese Female Facial Expression (JAFFE)*. Metode yang diusulkan menghasilkan akurasi keseluruhan sebesar 84%[6].

Penelitian oleh Jing Li et al. memperkenalkan metode baru untuk pengenalan ekspresi wajah menggunakan arsitektur jaringan syaraf konvolusional (CNN) yang dipadukan dengan *attention mechanism*. Penelitian tersebut menggunakan ekstraksi fitur *Local Binary Patterns (LBP)* digunakan untuk menangkap informasi tekstur wajah. Hasil menunjukkan bahwa kombinasi LBP dan CNN dapat meningkatkan akurasi pengenalan ekspresi wajah, ditambah dengan *attention module* yang digunakan dalam penelitian. Pada dataset FER2013, Metode gabungan yang digunakan mendapat akurasi sebesar 75.82%, tanpa *attention module* sebesar 73.96% dan tanpa LBP mendapat akurasi sebesar 67.73 [1].

Dari beberapa penelitian terdahulu, deteksi ekspresi wajah masih perlu peningkatan. Seperti menunjukkan adanya kemungkinan *overfitting* yang menjadi tantangan. Namun, beberapa penelitian sebelumnya belum mengeksplorasi lebih jauh bagaimana model dapat robust terhadap kondisi dunia nyata, seperti variasi pose, pencahayaan, atau noise. Dengan demikian, tujuan dari penelitian ini adalah untuk meningkatkan tingkat akurasi dalam klasifikasi ekspresi wajah dan mengeksplorasi lebih jauh bagaimana model dapat akurat dan stabil terhadap

berbagai kondisi. Untuk metode yang akan digunakan adalah menggunakan kombinasi *Multi-Level Convolutional Neural Network* (Multi-Level CNN) dan *Local Binary Pattern* (LBP). Metode ini diharapkan mampu mengekstraksi fitur dari berbagai tingkat abstraksi (*low-level, mid-level, dan high-level*) dan juga fitur tekstur lokal dari LBP sehingga dapat diperoleh model yang lebih akurat, stabil, dan mampu menghasilkan performa yang lebih unggul dibandingkan dengan metode CNN konvensional.

1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan yang telah dijelaskan, terdapat beberapa rumusan masalah sebagai berikut:

1. Bagaimana cara mengekstraksi fitur ekspresi wajah dengan menggabungkan tekstur lokal dari LBP dan representasi hierarkis dari *Multi-Level CNN*?
2. Bagaimana implementasi arsitektur LBP dan *Multi-Level CNN* dapat memberikan hasil yang lebih baik dibandingkan model tradisional dalam tugas klasifikasi ekspresi wajah?
3. Sejauh mana akurasi model gabungan *Multi-Level CNN* dan LBP dalam mengenali tujuh kategori ekspresi wajah?

1.3 Tujuan Penelitian

Pada penelitian ini, memiliki beberapa tujuan atau target yang ingin dicapai. Berikut adalah beberapa tujuan permasalahan :

1. Mengembangkan arsitektur klasifikasi ekspresi wajah berbasis gabungan *Multi-Level CNN* dan LBP untuk menangkap fitur pada berbagai tingkat representasi dan tekstur lokal.
2. Membandingkan performa model gabungan *Multi-Level CNN* dan LBP dengan model CNN konvensional untuk mengevaluasi kontribusi tiap metode.
3. Mengevaluasi hasil kombinasi LBP dan *Multi-Level CNN* dalam mengklasifikasi ekspresi wajah.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi pada berbagai aspek. Adapun manfaat yang dapat diperoleh dari penelitian ini antara lain sebagai berikut:

1. Menambah wawasan di bidang kecerdasan buatan, khususnya dalam pengembangan dan implementasi kombinasi *Multi-Level CNN* dan LBP untuk analisis ekspresi wajah.
2. Menjadi acuan untuk penelitian lanjutan terkait deteksi emosi manusia dengan pendekatan berbasis deep learning.
3. Menghasilkan model klasifikasi ekspresi wajah yang dapat digunakan dalam aplikasi nyata seperti sistem interaksi manusia-mesin, monitoring emosi dalam pendidikan, atau dukungan kesehatan mental.
4. Memberikan solusi untuk pengembangan sistem yang dapat menganalisis emosi pengguna dalam kondisi variatif.

1.5 Batasan Masalah

Penelitian ini memiliki beberapa batasan masalah agar penelitian dapat fokus pada permasalahan yang dikaji dan tidak terlalu luas. Beberapa batasan masalah pada penelitian ini adalah sebagai berikut:

1. Penelitian hanya berfokus pada klasifikasi ekspresi wajah berdasarkan data statis (gambar), tanpa mempertimbangkan data temporal (video).
2. Penelitian ini tidak membahas deteksi wajah secara otomatis dari gambar beresolusi penuh, melainkan menggunakan gambar wajah yang telah dipotong berukuran 48×48 piksel.
3. Pengujian dilakukan menggunakan dataset FER2013 dan pengujian tambahan menggunakan data primer berupa citra wajah individu secara terbatas, hanya sebagai analisa pendukung kemampuan model terhadap variasi nyata.

Halaman ini sengaja dikosongkan