

BAB I

PENDAHULUAN

1.1 Latar Belakang

Wajah merupakan alat pendukung dalam berkomunikasi secara langsung. Wajah dikategorikan sebagai salah satu biometrik manusia berdasarkan perilaku seseorang. Penggunaan wajah sudah banyak diaplikasikan di kehidupan sehari-hari. Wajah sering dimanfaatkan dalam memberikan suatu informasi penting dari seseorang. Informasi yang penting tersebut terlihat pada ekspresi wajah dari seseorang. Penelitian tentang pengenalan wajah sejak tahun 1960 hingga sekarang masih terus dilakukan suatu penyempurnaan.

Ekspresi memiliki peran yang mendasar dalam berkomunikasi *non-verbal* untuk menyampaikan keadaan emosi seseorang kepada orang lain secara alami baik yang positif maupun yang negatif. Ekspresi wajah merupakan hal penting dan kunci utama dalam mencapai suatu keberhasilan berkomunikasi [1]. Walaupun, manusia dapat mengenali ekspresi dengan baik, tetapi pada sejumlah penelitian masih dilakukan tentang analisis pengenalan ekspresi pada bidang *computer vision*, *machine learning* maupun *consumer research*.

Dengan berkembangnya suatu ilmu teknologi pengolahan citra digital (*Digital Image Processing*) semenjak adanya teknologi komputer yang dapat mengolah citra digital dengan algoritma tertentu dimulai pada tahun 1960 hingga saat ini. Teknologi pengolahan citra digital sudah banyak diterapkan di berbagai bidang ilmu, mulai dari bidang pendidikan hingga bidang kesehatan. Berkembangnya teknologi pengolahan citra yang diterapkan ke berbagai bidang ini membuat semakin banyaknya masalah-masalah yang dapat diselesaikan. Bukan

hanya teknologi dari pengolahan citra digital saja yang berkembang, tetapi saat ini metode-metode *Deep Learning* khususnya *Recognition* juga melatarbelakangi teknologi pengolahan citra digital. Salah satunya adalah pengaplikasian metode *Recognition* yang terdapat di bidang pendidikan yaitu pada sebuah sekolah di China dilakukan pemasangan kamera yang diletakkan di atas papan tulis untuk mengidentifikasi emosi siswa dan menentukan apakah siswa fokus belajar atau tidak pada saat pembelajaran di kelas yang dilihat dari ekspresi wajah.

Saat ini, *Deep Learning* merupakan salah satu teknologi yang sedang berkembang dalam bidang *Machine Learning* [2], dikarenakan kapabilitasnya yang signifikan dalam memodelkan berbagai data yang kompleks seperti citra dan suara. Salah satu metode *Deep Learning* yang saat ini memiliki hasil paling signifikan dalam pengenalan citra digital adalah algoritma *Convolutional Neural Network* (CNN) yang dimana algoritma tersebut dapat mengolah data dalam jumlah yang besar.

Penelitian tentang pengenalan ekspresi hingga saat ini masih dilakukan. Beberapa penelitian yang berkaitan dengan pengenalan ekspresi diantaranya melakukan penelitian dengan menggunakan metode *Convolutional Neural Network* (CNN) [3] dengan arsitektur yang mereka ajukan yaitu terdapat 4 bagian dengan menggunakan dataset *Extended Cohn-Kanade* (CKP) dan *MMI Facial Expression Database*. Pada bagian pertama terdapat 1 *convolution* dengan filter 64, kemudian 1 *Pooling*, dan terdapat 1 LRN. Pada bagian kedua dan ketiga terdapat 2 *feature extraction* yang diberi nama *FeatEx* (Parallel Feature Extraction Block) yang terinspirasi dari arsitektur *GoogLeNet* [4]. Selanjutnya bagian terakhir terdapat sebuah *Classifier*. Pada penelitian tersebut menghasilkan nilai akurasi

sebesar 99,6% untuk dataset CKP dan 98,63% untuk dataset MMI. Berikutnya pada penelitian mengenai pengenalan ekspresi dengan menggunakan metode *Convolutional Neural Network* (CNN) [5] dengan 5 arsitektur yang mereka ajukan. Arsitektur pertama dengan nama BKStart yang menggunakan 3 *convolutional layer*, kemudian arsitektur yang kedua hingga kelima dengan nama BKVGG8, BKVGG10, BKVGG12, BKVGG14 dengan desain arsitektur yang hampir mirip dengan arsitektur VGG. Masing-masing arsitektur tersebut menggunakan 8, 10, 12, 14 *convolutional layer*. Konfigurasi yang digunakan untuk melakukan *training* adalah menggunakan 1400 epoch, *batch size* sebesar 256, nilai *dropout* sebesar 0.5 dan untuk optimasi menggunakan *mini-batch gradient descent* dengan *momentum* sebesar 0.9. Konfigurasi yang digunakan untuk melakukan *testing* adalah menggunakan metode pertama yaitu *tencrop* dan metode kedua yaitu *eighteen* augmentasi. Metode *testing* kedua menghasilkan nilai akurasi tertinggi, yaitu terdapat pada arsitektur BKVGG12 dengan nilai akurasi *public test* 71,0% dan nilai akurasi *private test* 71,9% dalam pengenalan atau indentifikasi 7 kelas ekspresi wajah menggunakan dataset FER2013.

Berdasarkan uraian di atas, peneliti yang dilakukan oleh [3], [5], [6] melakukan penelitian tentang pengenalan ekspresi dengan berbagai macam arsitektur *Convolutional Neural Network* (CNN) yang dibuat untuk mendapatkan suatu model yang memiliki akurasi tinggi dalam pengenalan atau indentifikasi 7 kelas ekspresi wajah. Oleh karena itu, penulis akan melakukan eksplorasi arsitektur *Convolutional Neural Network* (CNN) yang tepat dalam pengenalan atau indentifikasi 7 kelas ekspresi wajah dengan menggunakan objek bahan penelitian

dari dataset FER2013 [7]. Penelitian ini akan diimplementasikan dengan menggunakan data uji wajah orang Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka dapat dirumuskan masalah sebagai berikut:

1. Bagaimana perbedaan pengaruh dari beberapa arsitektur *Convolutional Neural Network* (CNN) terhadap tingkat akurasi ?
2. Apa faktor yang mempengaruhi tingkat akurasi pada arsitektur *Convolutional Neural Network* (CNN) ?
3. Apa faktor yang mempengaruhi dalam hal waktu komputasi pada arsitektur *Convolutional Neural Network* (CNN) ?
4. Bagaimana kinerja dari beberapa arsitektur *Convolutional Neural Network* (CNN) dalam mengidentifikasi ekspresi wajah ?
5. Bagaimana kinerja metode *Convolutional Neural Network* (CNN) dalam mengidentifikasi ekspresi wajah dari Indonesia ?

1.3 Batasan Masalah

Agar penelitian ini dapat fokus sesuai pembahasan dalam penelitian maka penulis memberikan batasan masalah agar tidak menyimpang dari tujuan penelitian. Adapun batasan masalahnya, sebagai berikut:

1. Data yang digunakan pada penelitian ini adalah data berupa citra ekspresi wajah yang bersumber dari *Kaggle* yaitu FER2013 [7] yang berjumlah sebanyak 35,887 citra. *Dataset* FER2013 dibagi menjadi 3 bagian yaitu data *training*, *validation*, dan *testing* sebanyak masing-

masing 28,708, 3,589, dan 3,589 citra. Semua citra berwarna *grayscale* dan memiliki resolusi 48x48 *pixel*.

2. Data uji yang digunakan pada penelitian ini adalah data berupa citra ekspresi wajah dari orang Indonesia. Data ini berjumlah sebanyak 100 citra.
3. Jenis ekspresi yang dijadikan objek penelitian ini terdapat 7 kelas ekspresi yaitu ekspresi *Angry, Disgust, Fear, Happy, Sad, Surprise, Neutral*.
4. Variabel ketetapan yang digunakan dalam melakukan proses *training* arsitektur *Convolutional Neural Network* (CNN) adalah menggunakan konfigurasi sebagai berikut.

Tabel 1.1 Variabel Ketetapan

No.	Nama Variabel	Value
1.	Dropout	0.5
2.	Callback	0.9
3.	Optimizer	SGD
4.	Epoch	300
5.	Activation	ReLU
6.	Batch Size	128

1.4 Maksud dan Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Mengimplementasikan algoritma *Convolutional Neural Network* (CNN) untuk pengenalan atau identifikasi ekspresi wajah.

2. Mengetahui arsitektur *Convolutional Neural Network (CNN)* yang baik untuk menentukan tingkat akurasi dalam pengenalan atau identifikasi ekspresi wajah.
3. Mengetahui faktor-faktor yang mempengaruhi tingkat akurasi pada arsitektur *Convolutional Neural Network (CNN)*.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Mengetahui arsitektur *Convolutional Neural Network (CNN)* yang baik untuk pengenalan atau identifikasi ekspresi wajah.
2. Dapat menjadi referensi bagi peneliti yang akan melakukan penelitian dalam bidang *Deep Learning* dengan menggunakan algoritma *Convolutional Neural Network (CNN)* terutama bagi Mahasiswa Universitas Amikom Yogyakarta.
3. Mempermudah pengenalan atau identifikasi ekspresi wajah yang digunakan dalam bidang-bidang terkait.

1.6 Metode Penelitian

Metode yang digunakan dalam penelitian ini meliputi metode pengumpulan data, metode desain eksperimen, metode evaluasi dan metode pengujian sistem.

1.6.1 Metode Pengumpulan Data

Pada tahap metode pengumpulan data, dilakukan pengumpulan data berupa dataset FER2013 [7] yang akan digunakan dalam penelitian ini. Dataset tersebut di download melalui website *Kaggle*. Dataset FER2013 berupa file

berformat .csv dengan memiliki 3 kolom yaitu kolom *emotion*, *pixel* dan *usage*. Dataset yang berupa file format .csv akan dilakukan pengolahan data dengan melakukan pemisahan dataset menjadi 3 bagian yaitu data *training*, *validation*, dan *testing*.

Pada proses pengumpulan data pengujian, dilakukan pengumpulan data uji dari video, khususnya orang Indonesia yang dilakukan dengan cara mengambil bagian wajah yang memperlihatkan ekspresi wajah dengan cara *screenshot* pada bagian wajah tersebut. Agar dalam pengujian pengenalan atau identifikasi ekspresi wajah dapat dilakukan dengan ekspresi yang tidak disengaja maka pengambilan citra diambil dari video yang terdapat pada *film* maupun *youtube*.

1.6.1.1 Studi Literatur

Pada penelitian ini dilakukan studi literatur untuk pengumpulan data dan mempelajari informasi yang diperoleh dari buku, jurnal, skripsi, tutorial, dan berbagai informasi yang berkaitan dengan penelitian. Informasi yang berkaitan tersebut antara lain, proses pengolahan citra, *Deep Learning*, *Recognition*, *Machine Learning*, *Supervised Learning*, struktur dari *Convolutional Neural Network* (CNN), serta *library* yang akan digunakan dalam membantu penelitian ini.

1.6.2 Desain Eksperimen

Pada tahap metode desain eksperimen, dilakukan perancangan arsitektur *Convolutional Neural Network* (CNN), modifikasi parameter dan kedalaman arsitektur *Convolutional Neural Network* (CNN), pengaruh jumlah parameter, kedalaman *layer*, *learning rate*, serta metode *eval*

terhadap akurasi dan waktu *training* dari model *Convolutional Neural Network* (CNN). Dalam proses *training* arsitektur *Convolutional Neural Network* (CNN) dilakukan pembagian data *training*, *validation*, dan *testing*. Proses perancangan ini dilakukan berdasarkan hasil dari analisis studi literatur yang telah diperoleh.

1.6.3 Evaluasi

Pada tahap metode evaluasi, dilakukan evaluasi hasil akurasi yaitu berupa akurasi *training*, *validation* dan *testing* terhadap model yang didapatkan dari implementasi arsitektur *Convolutional Neural Network* (CNN). Evaluasi ini dilakukan untuk melihat faktor-faktor yang dapat mempengaruhi peningkatan akurasi.

1.6.4 Metode Pengujian Sistem

Pada tahap metode pengujian sistem, memastikan program yang dibuat berjalan sesuai apa yang diharapkan dengan melakukan pengamatan pengaruh jumlah parameter, kedalaman *layer*, learning rate dan data augmentasi terhadap akurasi dan waktu *training* dari model *Convolutional Neural Network* (CNN). Melakukan pengujian pengenalan atau identifikasi ekspresi wajah dengan menggunakan data uji orang Indonesia dari model yang didapatkan dari implementasi arsitektur *Convolutional Neural Network* (CNN).

1.7 Sistematika Penulisan

Pada bagian ini dituliskan urutan dan sistematika penulisan yang dilakukan. Sistematika penulisan ini dijelaskan dengan ringkas sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini berisikan dari latar belakang, rumusan masalah, maksud dan tujuan penelitian, metode penelitian, dan sistematika penulisan.

BAB II LANDASAN TEORI

Pada bab ini berisikan tentang studi literatur mengenai *Deep Learning*, *Recognition*, *Machine Learning*, *Supervised Learning*, dan *Computer Vision* dalam klasifikasi citra menggunakan *Convolutional Neural Network* (CNN). Teori-teori pendukung yang diperlukan untuk menyelesaikan permasalahan yang berkaitan dengan pengenalan atau identifikasi ekspresi wajah menggunakan *Convolutional Neural Network* (CNN).

BAB III METODOLOGI PENELITIAN

Pada bab ini berisikan tentang metode pengambilan dataset, pengambilan data uji orang Indonesia, proses *training* arsitektur *Convolutional Neural Network* (CNN), pengukuran tingkat akurasi terhadap model yang didapatkan dari implementasi arsitektur *Convolutional Neural Network* (CNN), pengujian data uji terhadap model yang didapatkan dari implementasi arsitektur *Convolutional Neural Network* (CNN), pengamatan *overfitting* dan parameter-parameter *Convolutional Neural Network* (CNN) yang digunakan.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini berisikan tentang pembahasan implementasi algoritma *Convolutional Neural Network* (CNN) dalam pengenalan atau identifikasi ekspresi wajah, membahas pengamatan *overfitting* dan hal yang

mempengaruhi tingkat akurasi model, serta pengujian data uji ekspresi orang Indonesia dari model yang didapatkan dari implementasi arsitektur *Convolutional Neural Network (CNN)*.

BAB V KESIMPULAN

Pada bab ini berisikan tentang kesimpulan dan saran dalam penelitian ini.

