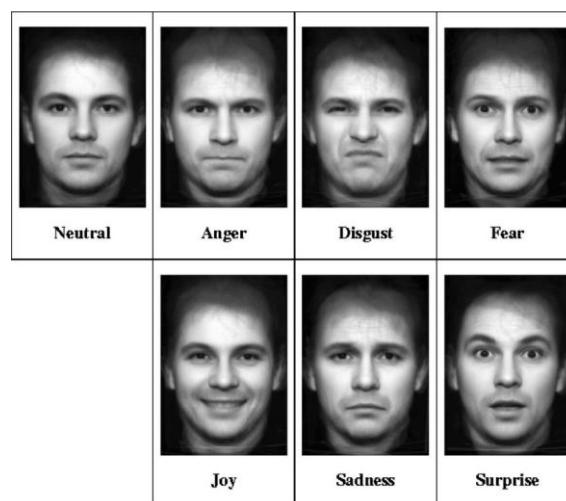


BAB II

TINJAUAN PUSTAKA

2.1 Emosi

Emosi adalah sebuah perasaan yang dapat mendorong seseorang untuk bertindak ataupun merespon dari suatu *stimulus* (Goleman, 2002). Emosi seseorang dapat diketahui dengan melihat dan mengamati ekspresi *mikro* seseorang tersebut, ekspresi kecil yang ditunjukkan oleh seseorang merupakan hal yang bersifat *universal* bagi seluruh manusia (Ekman, 2016). Ekspresi mikro merupakan Gerakan wajah singkat yang mencoba untuk menyembunyikan emosi saat sedang mengungkapkan emosi yang dialami (Amynarto et al., 2018). Emosi dapat dilihat dari perubahan pada raut wajah, seperti kerutan pada kening dan kedipan mata (L.Pt.Purnamaningsih, Ni Kt. Suarni, 2019).



Gambar 2.1 gambar ekspresi wajah manusia

Sumber : (youngontop.com)

2.2 Citra

Citra adalah representasi visual dua dimensi yang dapat dibuat melalui proses sampling dari gambar analog yang kontinu menjadi gambar diskrit. Proses sampling ini terdiri dari dua tahap, yaitu *downsampling* dan *unsampling*. *Downsampling* adalah langkah di mana nilai citra diperkecil dengan mengurangi jumlah piksel atau

resolusi spesial citra. Sedangkan *unsampling* merupakan kebalikan dengan *downsampling*, yaitu proses yang dapat menaikkan resolusi gambar atau jumlah *pixel* gambar (Nabuasa, 2019).

2.3 Deteksi Wajah

Deteksi wajah atau sistem pengenalan wajah memiliki prinsip yang hampir serupa dengan sistem pengenalan biometrik lainnya, seperti pengenalan telapak tangan dan sidik jari. Konsepnya adalah bahwa setiap individu memiliki fitur dan struktur wajah yang unik atau berbeda satu sama lain. Oleh karena itu, pengenalan wajah secara otomatis dapat dilakukan dengan membaca simetri dari wajah setiap individu. Meskipun demikian, sistem pengenalan wajah menghadapi beberapa tantangan, termasuk: (1) Variasi titik skala dan pergeseran pada wajah. (2) Perbedaan dalam penampilan wajah, seperti pose wajah, bentuk wajah, gaya rambut, makeup, kumis, janggut, aksesoris yang dipakai, dan lainnya. (3) Pengaruh pencahayaan. (4) Pengaruh usia. Oleh karena itu, hampir semua area wajah menjadi penting dalam mendeteksi emosi seseorang, terutama bagian wajah dan bibir yang dapat secara jelas menggambarkan emosi seseorang (Tri Anindia Putra, 2015).

Proses pengenalan wajah umumnya terdiri dari dua tahap, yaitu mendeteksi wajah (*pre-processing*) untuk mengekstraksi ciri-ciri wajah dan sistem pengenalan wajah (*face recognition*) (Alexander, 2013). Pengenalan wajah pada seseorang dapat menjadi input yang membantu dalam mendeteksi ekspresi wajah. Ekspresi dan emosi individu menjadi bentuk komunikasi nonverbal yang dapat mengungkapkan isi hati. Ekspresi wajah, yang melibatkan gerakan otot wajah, dapat menyampaikan berbagai perasaan kepada orang yang berkomunikasi. Contohnya, mengerutkan alis bisa menunjukkan kemarahan atau ketidaksetujuan, sementara mengangkat alis dapat mengekspresikan keheranan dan kejutan. Semua jenis emosi dan berbagai perasaan manusia tercermin dalam ekspresi wajah yang beragam (Abidin, 2011).

2.4 Python

Python merupakan bahasa pemrograman yang memungkinkan analisis data untuk melakukan perhitungan statistik yang kompleks, membuat visualisasi data, dan mengimplementasikan algoritma *machine learning*. Selain itu, Python dapat

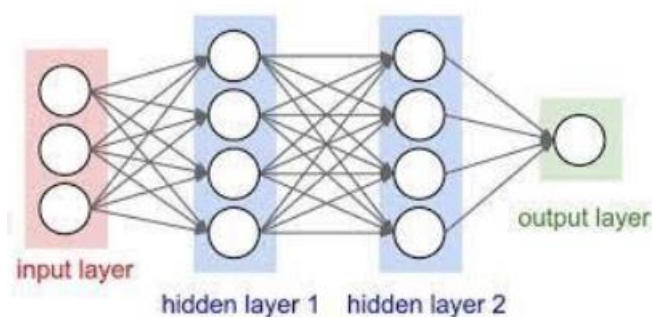
digunakan untuk memanipulasi dan menganalisis data, serta menyelesaikan berbagai tugas lain yang terkait dengan data.

2.5 OpenCV

OpenCV adalah perpustakaan (library) yang digunakan untuk memproses gambar dan video dengan kemampuan untuk mengekstrak informasi di dalamnya. *OpenCV* dapat diimplementasikan dalam berbagai bahasa pemrograman, termasuk C, C++, Java, Python, dan mendukung berbagai platform seperti Windows, Linux, Mac OS, iOS, dan Android.

2.6 Convolutional Neural Network (CNN).

CNN merupakan metode pengembangan dari multilayer perceptron (MLP) yang dirancang khusus untuk memproses data dua dimensi. Salah satu jenis *neural network* yang termasuk dalam kategori deep neural network adalah CNN, karena memiliki struktur jaringan yang mendalam dan sering digunakan untuk pengolahan citra (Eka Putra, 2016). CNN terdiri dari dua tahapan utama, yaitu *feature learning* dan *classification*. Tahap *feature learning* melibatkan *convolution layer*, *ReLU* (fungsi aktivasi), dan *pooling layer*; sementara tahap *classification* melibatkan *flatten*, *fully-connected layer*, dan *prediksi*. Meskipun cara kerjanya mirip dengan MLP, CNN dirancang khusus untuk memproses data dalam bentuk dua dimensi (Achmad et al., 2019)



Gambar 2.2 Arsitektur MLP

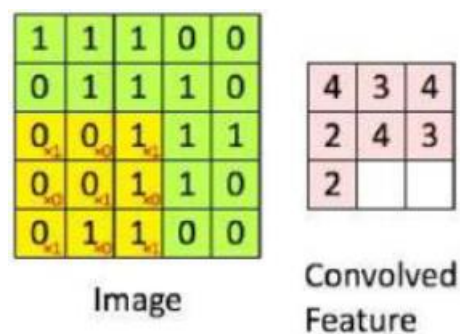
Sumber : (Eka Putra, 2016)

Ilustrasi pada Gambar 2.2 menjelaskan konsep kerja dari arsitektur MLP. Kotak merah dan biru melambangkan layer yang mengandung *neuron* yang

direpresentasikan oleh lingkaran putih. Input untuk MLP berupa data satu dimensi, yang kemudian disebarakan melalui jaringan untuk menghasilkan *output*. Hubungan antara satu *neuron* dengan *neuron* lain di dua layer yang bersebelahan dijelaskan melalui parameter bobot. Data input mengalami operasi linear menggunakan bobot di setiap layer, dan hasilnya diubah melalui operasi linear yang dikenal sebagai fungsi aktivasi. Secara kontras, pada CNN, data yang disebarakan adalah data dua dimensi, mengakibatkan penggunaan operasi linear yang berbeda dan parameter bobot yang tidak lagi berupa satu dimensi seperti pada MLP. Operasi linear pada CNN disebut operasi *konvolusi*, menandai perbedaan utama antara kedua arsitektur ini (Eka Putra, 2016).

2.6.1 *Konvolusi*

Konvolusi adalah proses kunci yang menjadi fondasi utama dalam arsitektur CNN. Proses *konvolusi* pada keluaran dari layer sebelumnya disebut *Convolution* layer. Ini dapat dijelaskan sebagai penerapan suatu fungsi pada keluaran fungsi lain secara berulang. Dalam konteks citra, operasi *konvolusi* dilakukan dengan menerapkan *kernel* ke berbagai posisi pada citra, seperti yang terlihat pada Gambar 2.3.



Gambar 2.3 operasi konvolusi

Sumber : (Eka Putra, 2016)

Pada gambar 2.3 kotak kuning merupakan *kernel* dan kotak hijau *offset* dari data citra. *Kernel* bergerak dari kiri atas menuju kanan bawah. *Konvolusi* bertujuan untuk mengekstraksi fitur dari citra input. *Konvolusi* akan menghasilkan *transformasi linear* dari data input sesuai informasi *special*

pada data. Bobot pada layer tersebut menspesifikasikan *kernel konvolusi* yang digunakan, sehingga *kernel konvolusi* dapat dilatih berdasarkan input pada CNN (Eka Putra, 2016).

2.6.2 Max polling

Pooling layer merupakan cara untuk melakukan pengurangan ukuran *matrix*. *Pooling layer* terdiri dari sebuah filter yang memiliki ukuran dan nilai (Hakim & Rainarli, 2019). Filter ini akan bergeser pada seluruh area *feature map*. Ukuran dan jumlah parameter akan berkurang setelah menggunakan *layer* ini, sehingga akan mencapai komputasi. *Pooling layer* memiliki beberapa macam tipe antara lain *average pooling*, *max pooling*, dan *Lp Pooling* (Zufar & Setiyono, 2016).

Image Matrix				Max Pool	
2	1	3	1	2	4
1	0	1	4	7	9
0	6	9	5		
7	1	4	1		

Gambar 2.4 Operasi Max pooling

Sumber : (Walters, 2019)

2.6.3 ReLu

ReLU merupakan fungsi aktivasi yang digunakan untuk normalisasi nilai yang dihasilkan oleh layer konvolusi. Fungsi *ReLU* beroperasi dengan mengonversi nilai-nilai di bawah 0 menjadi 0, menggunakan fungsi maksimum (Achmad et al., 2019). *Batch normalization* merupakan metode yang dapat diterapkan untuk menormalkan *input* dari setiap lapisan. Tujuan menggunakan *batch normalization* untuk mengatasi masalah pergeseran kovariat internal (*internal covariate shift*). Manfaat dari *batch normalization* yaitu membantu menstabilkan pelatihan dan sangat efektif

mengurangi *epoch* yang dibutuhkan untuk melatih jaringan saraf (Astuti, 2019).

2.6.4 Dropout

Dropout memiliki tujuan yaitu mencegah terjadi *overfitting* dan mempercepat proses *learning*. *Overfitting* merupakan kondisi dimana terjadi ketidaksesuaian pada proses prediksi setelah semua data melalui proses *training* dengan persentase yang baik. Cara kerja *dropout* dengan menghilangkan sementara suatu *neuron* pada *Hidden Layer* maupun *Visible Layer* yang berada didalam jaringan (Santoso & Ariyanto, 2018).

2.6.5 Flatten

Flattening adalah proses yang mengubah *matriks* menjadi vektor satu dimensi. Operasi *flattening* dilakukan pada *feature map* yang telah dihasilkan dari *layer* sebelumnya, mengubahnya menjadi vektor satu dimensi agar dapat diklasifikasikan menggunakan *fully-connected layer* dan *softmax*. Ilustrasi proses *flatten* dapat dilihat pada Gambar 2.5. Pada tahap *backpropagation*, layer ini mengembalikan vektor satu dimensi ke bentuk matriks asalnya, memungkinkan dilakukannya proses perubahan bobot *filter* lebih lanjut. (Achmad et al., 2019).



Gambar 2.5 ilustrasi flatern

Sumber : (Achmad et al., 2019)

2.6.6 *Softmax*

Softmax adalah fungsi aktivasi yang diterapkan pada *output layer*. *Output layer* memiliki sedikit perbedaan dengan *fully connected layer*, yang membedakan keduanya adalah penggunaan fungsi aktivasi *softmax* pada *layer output*, sedangkan *fully-connected layer* menggunakan fungsi aktivasi *ReLU* (Achmad et al., 2019).

2.6.7 *Fully Connected Layer*

Sebelum memasuki *fully connected layer*, *feature map* akan melalui tahap "*flatten*" atau *reshape*. Proses *flatten* menghasilkan vektor yang kemudian digunakan sebagai input bagi *fully connected layer*. *Fully connected layer* bertujuan untuk melakukan *transformasi* pada dimensi data sehingga data dapat diklasifikasikan secara *linear*. Implementasi *fully connected layer* sebaiknya dilakukan di akhir jaringan untuk menghindari kehilangan informasi spesifik pada data yang tidak dapat diubah kembali (Eka Putra, 2016).

2.7 *Confusion matrix*

Confusion matrix adalah metode yang umumnya digunakan untuk menghitung akurasi. *Confusion matrix* dapat dijelaskan melalui tabel yang menunjukkan jumlah data uji yang terklasifikasi dengan benar dan jumlah data uji yang salah terklasifikasi. Tabel dalam matrik *konvolusi* terdiri dari empat kombinasi berbeda antara nilai prediksi dan nilai aktual. (Rahman et al.,2017).

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Gambar 2.6 tabel confusion matrix

Sumber : (Stazio et al, 2019)

Gambar 2.6 dapat dijelaskan dengan konsep *True Positive* (TP), yang mengindikasikan jumlah data yang kelas aktualnya *positif* dan kelas prediksi juga *positif*. Ini dapat diartikan sebagai prediksi *positif* yang benar. *False Negative* (FN) adalah jumlah data yang kelas aktualnya *positif*, tetapi kelas prediksinya *negatif*. Ini dapat diinterpretasikan sebagai prediksi *negatif* yang salah. *False Positive* (FP) menggambarkan jumlah data yang kelas aktualnya *negatif*, tetapi kelas prediksinya *positif*. Ini dapat diartikan sebagai prediksi *positif* yang salah. Terakhir, *True Negative* (TN) adalah jumlah data yang kelas aktualnya *negatif* dan kelas prediksinya juga *negatif*. Ini dapat diinterpretasikan sebagai prediksi *negatif* yang benar. (Narkhede,2018).

A. Akurasi

Akurasi adalah parameter yang menilai sejauh mana hasil pengukuran serupa dengan nilai sebenarnya yang diukur. Tingkat akurasi mencerminkan sejauh mana nilai prediksi mendekati nilai aktual (Hendriyana & Maulana, 2020). Formula untuk perhitungan dapat ditemukan dalam persamaan. (2.1)

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN}$$

B. Presisi

Presisi adalah ukuran tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem (Hendriyana & Maulana, 2020). Rumus perhitungannya dapat ditemukan dalam suatu persamaan. (2.2)

$$Presisi = \frac{TP}{TP+FP}$$

C. *Recall*

Recall merupakan *parameter* yang mengukur tingkat keberhasilan sistem dalam menemukan kembali suatu informasi (Hendriyana & Maulana, 2020). Formula untuk menghitungnya dapat ditemukan dalam persamaan. (2.3)

$$Recall = \frac{TP}{TP+FN}$$

2.8 Penelitian sebelumnya

Penelitian sebelumnya yang dilakukan oleh Sang et al. (2017) menciptakan arsitektur model BKVGG8, BKVGG10, BKVGG12, dan BKVGG14 dengan menggabungkan BK-Start dan VFF. Mereka menggunakan *dataset* FER2013 dan mencapai akurasi sebesar 71%. Penelitian ini menunjukkan bahwa emosi dengan akurasi tertinggi adalah emosi senang.

Penelitian Nugroho et al. (2020) mencapai nilai akurasi sebesar 65%, meskipun tidak menjelaskan arsitektur model yang digunakan. Hasilnya menunjukkan bahwa emosi dengan akurasi paling tinggi dalam penelitian ini adalah emosi senang. Penelitian ini memiliki tingkat akurasi yang hampir sama dengan penelitian Septian et al., yang mencapai akurasi 64%, dengan fokus pada emosi senang.

Penelitian Ramdhani et al. (2018) mencapai akurasi sebesar 63,41%, menggunakan dataset FER2013 dan fokus pada emosi senang. Arsitektur model pada penelitian ini terdiri dari 9 lapisan.

Penelitian Achmad et al. (2019) mencapai akurasi tertinggi, yaitu 80,75%, menggunakan dataset CK+. Hasil ini menunjukkan kinerja yang lebih baik dibandingkan penelitian-penelitian sebelumnya. Dengan demikian, ringkasan penelitian sebelumnya dapat dilihat pada Tabel 2.1.

Tabel 2.1 Penelitian yang menggunakan CNN

No	Judul penelitian	penulis	Sumber	Dataset	Deteksi emosi	Variabel	Akurasi
1	Facial Expression Recognition Using Deep Convolutional Neural Network (Sang et al., 2017)	Ding Viet Sang, Nguyen Van Dat & Do Phan Thuan	2017 9th International Conference on Knowledge and Systems Engineering (KSE) (researchgate.net)	(FER2013) Grayscale 48x48 pixels (35.887 citra)	Marah,jijik, Takut,senang Sedih,terkejut, Netral	ekspresi wajah, Citra digital, persentase, persamaan	71%
2	Implementasi Deep Learning Menggunakan Convolutional Neural Network (CNN) pada Ekspresi Manusia (Nugroho et al.,2020)	Pulung Adi Nugroho, Indah Fenriana & Rudy Arijanto, M.Kom	Jurnal Algor, Universitas Buddhi Darma (jurnal.ubd.ac.id)	(FER2013) Grayscale 48x48 pixels (35.887 citra)	Marah,jijik, Takut,senang Sedih,terkejut, Netral	Klasifikasi Ekspresi Wajah, pengambilan data	65%

3	Klasifikasi Emosi Menggunakan Convolutional Neural Network (Septian et al.,n.d.)	Ripan Septian, Dede Irawan Saputra & Ridwan Ilyas	Prosiding Seminar Nasional Teknik Elektro UIN Sunan Gunung Djati Bandung (senter.ee.uinsgd.ac.id)	(FER2013) Grayscale 48x48 pixels (35.887 citra)	Marah,jijik, Takut,senang Sedih,terkejut, Netral	Emosi wajah. Pengenalan wajah, kemiripan	64,54%
4	Convolutional Neural Network Models for Facial Expression Recognition (Ramdhani et al., 2018)	Burhanudin Ramdhani, Esmeralda C. Djamal & Ridwan Ilyas	2018 International Symposium on Advanced Intelligent Informatics (SAIN)	(FER2013) Grayscale 48x48 pixels (35.887 citra)	Marah,jijik, Takut,senang Sedih,terkejut, Netral	Wajah, Validasi ekspresi. Deteksi wajah	63,41%
5	Klasifikasi Emosi Berdasarkan Ciri Wajah Menggunakan Convolutional Neural Network	Achmad Yusuf, Randy Cahya Wihandika, & Chandra Dewi	Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 2019 (ub.ac.id)	(CK+) Grayscale 48x48 pixels (981 citra)	Marah,jijik, Takut,senang Sedih,terkejut, Netral	Penjumlahan citra digital, Wajah, Ekspresi, Emosi,	80,75%

	(Achmad et al., 2019)						
6	Deteksi Emosi Menggunakan Citra Ekspresi Wajah Secara Otomatis	Faza Nur Azizi	Jurnal Universitas Islam Indonesia (UII) (uii.ac.id)	(FER2013+CK+) Grayscale 48x48 Pixels (36.868 citra)	Marah,jijik, Takut,senang Sedih,terkejut, Netral	Ekspresi, deteksi emosi, Pengenalan wajah, Jumlah citra digital	83,25%