



**METODE CONVOLUTIONAL NEURAL NETWORK
ARSITEKTUR RESNET DAN EANN-NET UNTUK
MENDETEKSI EKSPRESI WAJAH**

SEMINAR BIDANG KAJIAN

**DINI TRIASANTI
99216023**

**PROGRAM DOKTOR TEKNOLOGI
INFORMASI UNIVERSITAS GUNADARMA
SEPTEMBER 2022**

Daftar Isi

Daftar Isi.....	i
Daftar Gambar.....	ii
Daftar Tabel	ii
Bab 1.....	3
Pendahuluan.....	3
1.1 Latar Belakang.....	3
1.2 Batasan	6
1.3 Tujuan.....	6
1.4 Kontribusi.....	7
Bab 2	8
Tinjauan Pustaka.....	8
2.1 Ekspresi Wajah.....	8
2.2 Arsitektur CNN	9
2.3 Residual Network (ResNet)	9
2.4 Tinjauan 1	10
2.5 Tinjauan 2	10
2.6 Tinjauan 3.....	12
2.7 Tinjauan 4	13
2.8 Tinjauan 5	13
2.9 Tinjauan 6	14
2.10 Tinjauan 7.....	14
2.11 Tinjauan 8.....	15
2.12 Tinjauan 9.....	15
2.13 Tinjauan 10.....	16
2.14 Tinjauan 11.....	16
2.15 Tinjauan 12.....	18
2.16 Perbandingan Tinjauan.....	19
Bab 3.....	23
Metodologi.....	23
3.1 Motivasi.....	23
3.2 Framework Riset.....	24
3.3 Pendekatan.....	26
3.4 Rencana Kerja	28
Bibliografi.....	29

Daftar Gambar

Gambar 1. Arsitektur CNN	9
Gambar 2. Arsitektur ResNet : Blok Residual (He & Sun, 2015).....	9
Gambar 3. Contoh gambar dari database JAFFE	12
Gambar 4. Contoh gambar dari database CK.....	12
Gambar 5. Contoh gambar dari 10 kelas beranotasi dalam CIFAR-10	17
Gambar 6. Kerangka Dasar Penelitian.....	25
Gambar 7. Alur Proses Data	25
Gambar 8. ResNet-50	26
Gambar 9. Struktur Jaringan EANN	27
Gambar 10. Struktur EANN-Net yang dihasilkan oleh EANN.....	27

Daftar Tabel

Tabel 1. Perbandingan Tinjauan.....	19
Tabel 2. Algoritma EANN.....	26
Tabel 3. Rencana Kerja	28

Bab 1

Pendahuluan

1.1 Latar Belakang

Ekspresi wajah manusia sangat penting dalam komunikasi sosial. Biasanya komunikasi melibatkan verbal dan nonverbal. Komunikasi non-verbal diungkapkan melalui ekspresi wajah. Ekspresi wajah adalah sinyal halus dari komunikasi yang lebih besar. Komunikasi non-verbal berarti komunikasi antara manusia melalui kontak mata, gerak tubuh, ekspresi wajah, bahasa tubuh, dan parabahasa (Revina & Emmanuel, 2021).

Kontak mata adalah fase penting komunikasi yang memberikan campuran ide. Kontak mata mengontrol kontribusi, diskusi, dan membuat hubungan dengan orang lain. Ekspresi wajah meliputi senyum, sedih, marah, jijik, terkejut, dan takut. Senyum di wajah manusia menunjukkan kebahagiaan mereka dan itu mengekspresikan mata dengan bentuk melengkung. Ekspresi sedih adalah perasaan longgar yang biasanya diekspresikan sebagai naiknya alis miring dan cemberut. Kemarahan di wajah manusia terkait dengan kondisi yang tidak menyenangkan dan menjengkelkan. Ekspresi marah diekspresikan dengan alis yang dikerutkan, kelopak mata yang ramping dan diregangkan. Ekspresi jijik diekspresikan dengan alis yang ditarik ke bawah dan hidung yang berkerut. Kejutan atau ekspresi kaget diekspresikan ketika beberapa hal yang tidak terduga terjadi. Ini diekspresikan dengan pelebaran mata dan mulut menganga dan ekspresi ini mudah diidentifikasi. Ekspresi ketakutan berhubungan dengan ekspresi terkejut yang diekspresikan sebagai alis yang semakin miring (Revina & Emmanuel, 2021).

Pengenalan ekspresi wajah / FER memiliki tahapan penting yaitu ekstraksi ciri dan klasifikasi. Ekstraksi fitur mencakup dua jenis dan mereka berbasis geometris dan berbasis penampilan. Klasifikasi juga merupakan salah satu proses penting di mana ekspresi yang disebutkan di atas seperti senyum, sedih, marah, jijik, terkejut, dan takut dikategorikan. Ekstraksi fitur berbasis geometri terdiri dari mata, mulut, hidung, alis, komponen wajah lainnya dan ekstraksi fitur berbasis penampilan terdiri dari bagian wajah yang tepat (Zhao & Zhang, 2016).

Pengenalan ekspresi wajah memainkan peran utama dalam banyak aplikasi sistem otomatis seperti robotika, pendidikan, kecerdasan buatan, dan keamanan. Mengenali ekspresi wajah secara akurat memang menantang. Pendekatan untuk memecahkan masalah FER (Facial Expression Recognition) dapat dikategorikan menjadi 1) Gambar tunggal statis dan 2) Gambar urutan. Secara tradisional, teknik yang berbeda seperti Multi-layer Perceptron Model, k-Nearest Neighbours, Support Vector Machines digunakan oleh para peneliti untuk memecahkan FER. Metode ini mengekstrak fitur seperti Local Binary Patterns, Eigenfaces, Face-landmark features, dan Texture features. Di antara semua metode ini, Neural Networks telah mendapatkan popularitas yang sangat tinggi dan mereka banyak digunakan untuk FER. Baru-baru ini, CNN (Convolutional

Neural Networks) telah mendapatkan popularitas di bidang pembelajaran mendalam karena arsitektur kasual dan kemampuan mereka untuk memberikan hasil yang baik tanpa persyaratan ekstraksi fitur manual dari data gambar mentah. Makalah ini berfokus pada survei berbagai teknik pengenalan ekspresi wajah berdasarkan CNN mulai dari metode canggih yang disarankan oleh peneliti yang berbeda dan juga menunjukkan langkah-langkah yang diperlukan untuk penggunaan CNN untuk FER serta mencakup analisis pendekatan berbasis CNN dan masalah yang membutuhkan perhatian saat memilih CNN untuk menyelesaikan FER. Pengenalan ekspresi wajah menggunakan deep learning melalui tahap augmentasi dan pemrosesan dan klasifikasi (Suma et al., 2018).

Saat ini, analisis gerakan tubuh manusia untuk pengenalan emosi sangat penting untuk komunikasi sosial. Metode komunikasi non-verbal seperti gerakan tubuh, ekspresi wajah, gerak tubuh dan gerakan mata digunakan dalam beberapa aplikasi. Diantaranya pengenalan emosi dari gerakan tubuh memiliki keunggulan mengenali emosi seseorang dari kamera apapun melihat dan juga mengenali emosi, jika orang terlalu jauh dari kamera. Gerakan tubuh dapat dengan kuat menyampaikan keadaan emosional daripada penelitian lain. Dalam makalah ini, keadaan emosi dikenali pola gerak seluruh tubuh menggunakan arsitektur jaringan syaraf tiruan feedforward deep convolution dengan parameter yang berbeda. Sistem yang diusulkan dapat dievaluasi dengan dataset emosi (University of .YORK) dengan 15 jenis emosi dan dataset corpus GEMEP dengan 5 emosi. Hasil eksperimen menunjukkan akurasi pengenalan yang lebih baik dari sistem yang diusulkan (Santhoshkumar & Geetha, 2019).

Saat ini, ada kebutuhan di masyarakat untuk menggunakan metode digital dan siberetik untuk diagnostik jarak jauh dari kondisi mental seseorang. Sistem digital atau cyberphysical dalam proses pengenalan jarak jauh dari keadaan mental pasien, dimungkinkan untuk meningkatkan efektivitas mengenali keadaan psiko-emosional pasien yang sebenarnya, baik dalam mode konsultasi maupun dalam pengamatan dinamis sistem. Untuk penilaian jarak jauh keadaan mental, metode banyak digunakan berdasarkan pengenalan aktivitas wajah, yang secara refleks dikaitkan dengan tampilan keadaan psiko-emosional. Dalam penelitian ini mengusulkan pendekatan baru untuk menciptakan pengamatan metodologi tersebut (Lebedev et al., 2020).

Penelitian lain mengatasi masalah deteksi keadaan emosional dari ekspresi wajah. Pendekatan yang diusulkan secara bersamaan mendeteksi wajah dan memprediksi kategori emosi yang berbeda dan nilai valensi berkelanjutan dari gambar masukan mentah. Peneliti melatih dan mengevaluasi pendekatan tersebut pada 3 kumpulan data yang berbeda, membandingkan pendekatan tersebut dengan pendekatan canggih lainnya dan melakukan evaluasi lintas basis data. Dengan cara ini ditemukan bahwa pendekatan tersebut dapat digeneralisasi dengan baik dan cocok untuk aplikasi waktu nyata (Handrich et al., 2020).

Karya ini dikhususkan untuk optimalisasi metode pengenalan tujuh emosi dasar (gembira, sedih, takut, marah, terkejut, jijik, dan netral) dalam hal ekspresi wajah manusia. Pendekatan yang ada dari konstruksi sistem pengenalan emosi dianalisis berdasarkan ekspresi wajah manusia, dan difokuskan pada keuntungan menggunakan skema berdasarkan jaringan saraf. Karya ini mengusulkan metode membangun sistem pengenalan emosi berdasarkan jaringan saraf, yang mencakup algoritma yang dioptimalkan untuk menghasilkan sampel pelatihan dan pengujian, serta menentukan jumlah lapisan jaringan saraf yang rasional. (Ivanova & Borzunov, 2020).

Pengenalan emosi secara *real-time* telah menjadi bidang penelitian aktif selama beberapa dekade terakhir. Karya ini bertujuan untuk mengklasifikasikan ekspresi emosi penyandang disabilitas fisik (tuli, bisu, dan terbaring di tempat tidur) dan anak autisme berdasarkan penanda wajah dan sinyal electroencephalograph (EEG) menggunakan pengklasifikasi convolutional neural network (CNN) dan long short-term memory (LSTM). Enam emosi wajah (senang, sedih, marah, takut, jijik, dan terkejut) dikumpulkan menggunakan sepuluh penanda virtual. 55 mahasiswa sarjana (35 laki-laki dan 20 perempuan) dengan usia rata-rata 22,9 tahun secara sukarela berpartisipasi dalam percobaan untuk pengenalan emosi wajah. Sembilan belas mahasiswa sarjana mengajukan diri untuk mengumpulkan sinyal EEG. Awalnya, fitur seperti Haar digunakan untuk deteksi wajah dan mata. Kemudian, penanda virtual ditempatkan pada lokasi yang ditentukan pada wajah subjek berdasarkan sistem pengkodean tindakan wajah menggunakan pendekatan model matematis, dan penanda tersebut dilacak menggunakan algoritma aliran optik Lucas-Kandc. Jarak antara pusat wajah subjek dan setiap posisi penanda digunakan sebagai fitur untuk klasifikasi ekspresi wajah. Fitur jarak ini divalidasi secara statistik menggunakan analisis varians satu arah dengan tingkat signifikansi $p < 0,01$. Selain itu, empat belas sinyal yang dikumpulkan dari saluran pembaca sinyal EEG (EPOCH) digunakan sebagai fitur untuk klasifikasi emosional menggunakan sinyal EEG. Terakhir, fitur divalidasi silang menggunakan validasi silang lima kali lipat dan diberikan kepada pengklasifikasi LSTM dan CNN. Hasil mencapai tingkat pengenalan maksimum 99,81% menggunakan CNN untuk deteksi emosi menggunakan landmark wajah. Namun, tingkat pengenalan maksimum yang dicapai dengan menggunakan pengklasifikasi LSTM adalah 87,25% untuk deteksi emosi menggunakan sinyal EEG (Hassoun et al., 2020).

Deteksi emosi manusia adalah salah satu masalah utama dalam visi komputer. Emosi manusia terdiri dari beberapa sub-emosi yang sulit untuk diklasifikasikan ke dalam kelas tertentu. Penelitian ini telah mencoba mengklasifikasikan emosi ke dalam 6 kategori dasar (senang, sedih, jijik, takut, terkejut, marah) dan emosi manusia yang netral. Karya ini telah mengusulkan kerangka pembelajaran mendalam yang terdiri dari CNN, ResNet dan blok *attention* yang memberikan persepsi visual ke jaringan. Model yang diusulkan memiliki penerapan yang lebih besar dalam deteksi emosi wajah kehidupan nyata. Model yang diusulkan telah mencapai hasil yang memuaskan dan telah menunjukkan hasil yang efektif pada dataset FER (Gupta et al., 2020).

Penelitian berikutnya mengusulkan algoritma evolusioner baru untuk membuat arsitektur jaringan saraf secara otomatis. Algoritma yang diusulkan menggunakan blok bangunan dasar dari jaringan canggih yang ada untuk menginisialisasi populasi. Selama proses evolusi, algoritma menghasilkan jaringan baru yang dibangun dengan memanfaatkan blok yang berbeda dari jaringan yang berbeda. Selain itu, skema pengkodean, operator genetik, dan fungsi telah dirancang dengan baik. Akibatnya, arsitektur jaringan yang dihasilkan, yang memiliki sifat heterogen yang kuat, cenderung mengintegrasikan keunggulan jaringan canggih yang ada. Hasil eksperimen menunjukkan bahwa arsitektur jaringan yang dibuat oleh algoritma yang diusulkan dapat mencapai akurasi klasifikasi yang lebih tinggi daripada arsitektur jaringan referensi dan yang dikembangkan oleh metode desain topologi evolusioner *state-of-the-art* lainnya, menunjukkan efektivitas algoritma yang diusulkan pada secara otomatis membuat arsitektur jaringan. Mengingat kinerja ResNet yang tinggi, penelitian ini berusaha untuk mempromosikan efisiensi EA dengan bantuan ResNet. Secara konkret, mengusulkan algoritma evolusioner baru untuk secara otomatis membuat arsitektur jaringan saraf, yang disebut EANN untuk jangka pendek. Algoritma yang diusulkan menggunakan blok bangunan dasar yang diturunkan dari struktur ResNet untuk menginisialisasi populasi. Sebagai hasil dari mengasimilasi keunggulan ResNet dan

EA, EANN yang diusulkan yaitu EANN-Net mampu menemukan arsitektur jaringan yang efektif dengan sejumlah evaluasi fungsi yang terbatas. (Chen et al., 2019).

Metode klasifikasi citra khususnya untuk pendeteksian ekspresi wajah masih terus dikembangkan. Mulai dari menggunakan metode *Deep Learning* sampai pada penggunaan metode *Convolutional Neural Network* dengan berbagai macam variasi arsitektur. Pada penggunaan metode terdahulu memerlukan waktu lama dalam pemrosesan, sedangkan pada metode yang menggunakan CNN dengan variasi arsitektur tertentu waktu pemrosesannya jauh lebih cepat sehingga meminimalisir waktu pemrosesan.

Berdasarkan pada penelitian yang telah dilakukan sebelumnya oleh para peneliti maka dapat disimpulkan bahwa teknologi yang sudah ada mampu mengenali ekspresi wajah manusia. Oleh karena itu, pada penelitian ini akan mencoba melakukan pengembangan metode *Convolutional Neural Network* arsitektur ResNet dan EA Neural Network yang digunakan untuk mengidentifikasi ekspresi wajah. Pemilihan algoritma ini dikarenakan arsitektur ResNet dan EANN-Net memiliki performa yang lebih baik untuk mempelajari data citra dibandingkan dengan metode konvensional lainnya. Sehingga dapat memperoleh pra-pemrosesan yang jauh lebih cepat dan mendapatkan hasil yang akurat.

1.2 Batasan

Batasan masalah pada penelitian ini adalah :

- Pengembangan metode Convolutional Neural Network (CNN) yang menggunakan arsitektur ResNet dan EANN-Net.
- Penerapan metode yang digunakan untuk klasifikasi objek citra ke dalam aplikasi deteksi ekspresi wajah seperti netral, sedih, senang, marah, terkejut, takut, dan jijik.
- Pengambilan data yang dilakukan secara mandiri untuk digabungkan dengan dataset yang sudah tersedia misalnya menggunakan dataset tertentu yaitu FER 2013.
- Penerapan model arsitektur ResNet dan EANN-Net yang diusulkan untuk mendeteksi ekspresi wajah.
- Pengevaluasian kinerja aplikasi terhadap model yang diusulkan.

1.3 Tujuan

Tujuan pada penelitian ini adalah :

- Mengembangkan model Convolutional Neural Network (CNN) menggunakan arsitektur ResNet dan turunannya yaitu EANN-Net yang dapat menghasilkan durasi waktu komputasi yang lebih cepat dan akurasi yang baik.
- Menerapkan metode yang digunakan oleh klasifikasi citra ke dalam deteksi ekspresi wajah.
- Mengimplementasikan model ResNet dan EANN-Net yang diusulkan untuk deteksi ekspresi wajah.
- Mengetahui model yang diusulkan dapat menghasilkan model baru yang diterapkan pada deteksi ekspresi wajah.
- Melakukan evaluasi terhadap kinerja dari model arsitektur ResNet dan EANN-Net yang diusulkan.

1.4 Kontribusi

Adapun kontribusi yang dapat diberikan oleh penelitian ini adalah:

Kontribusi keilmuan dalam bidang teknologi informasi dalam bentuk pengembangan model arsitektur ResNet dan EANN-Net untuk pengenalan ekspresi wajah, dengan durasi waktu komputasi yang lebih cepat dan akurasi yang baik.

Kontribusi keilmuan dalam bidang psikologi berupa penggunaan metode deep learning yang berfungsi sebagai identifikasi / deteksi ekspresi wajah dengan jumlah data yang besar, dapat dilakukan dengan mudah dan efisien.

Kontribusi bagi pendidikan, sistem otomatis HMI (Human Machine Interface), kesehatan dan keamanan. Pengenalan ekspresi wajah sedang diterapkan di banyak sektor industri dan pemerintah seperti dalam aplikasi real-time yaitu pengawasan kondisi pengemudi/pengendara, medis, interaksi robotika, bagian forensik, mendeteksi penipuan. Pengenalan ekspresi wajah juga dapat mengidentifikasi penyerang potensial berdasarkan analisis emosi, analisis efektivitas promosi, menentukan efektivitas metodologi untuk menjual barang di toko, dan pengujian reaksi skala besar terhadap produk (semantik wajah).

Bab 2

Tinjauan Pustaka

2.1 Ekspresi Wajah

Umumnya, wajah menawarkan tiga jenis sinyal yang berbeda seperti sinyal statis, lambat dan cepat. Sinyal statis adalah warna kulit yang mencakup beberapa aspek pigmentasi kulit wajah yang bertahan lama, endapan berminyak, bentuk wajah, susunan tulang, tulang rawan dan bentuk, lokasi dan ukuran fitur wajah seperti alis, mata, hidung, mulut. Sinyal lambatnya adalah kerutan permanen yang meliputi perubahan penampilan wajah seperti tonus otot dan perubahan tekstur kulit yang terjadi perlahan seiring waktu.

Sinyal cepatnya adalah menaikkan alis yang meliputi gerakan otot wajah, perubahan tampilan wajah yang tidak permanen, kerutan yang tidak permanen, serta perubahan letak dan bentuk fitur wajah. Kilatan di wajah ini bertahan selama beberapa detik. Ketiga sinyal ini diubah dengan opsi individual sementara sangat sulit untuk mengubah sinyal statis dan lambat. Juga, wajah adalah sistem multi-pesan dan bukan hanya sistem multi-sinyal. Pesan ditransmisikan melalui wajah yang meliputi emosi, perasaan posisi, usia, kualitas, kecerdasan, daya tarik dan hampir pasti substansi lainnya juga. (Paul Ekman, 2003)

Ekspresi wajah adalah cara yang paling efisien dan alami untuk menyampaikan emosi dan niat. Ekspresi wajah mewakili intensitas dan keadaan emosional orang tersebut. Ada banyak ekspresi yang mungkin, tetapi hanya tujuh ekspresi yang merupakan ekspresi prototipe, yaitu: Kemarahan, Kesedihan, Jijik, Kebahagiaan, Ketakutan, Kejutan, dan Netral (Suma et al., 2018).

Analisis emosional adalah campuran yang menarik dari psikologi dan teknologi. Salah satu pendekatan utama untuk mengenali emosi adalah analisis ekspresi wajah. Emosi dasar dianggap tetap secara biologis, bawaan dan, sebagai akibatnya, universal untuk semua orang, serta untuk banyak hewan. Emosi kompleks adalah kumpulan emosi dasar, atau emosi aneh. Masalah utamanya adalah menentukan emosi mana yang mendasar dan mana yang kompleks. Orang dapat mengenali sinyal-sinyal ini, meskipun ditampilkan secara halus, saat memproses informasi yang diterima oleh telinga dan mata. Berdasarkan studi psikologis yang menunjukkan bahwa informasi visual mengubah persepsi ucapan, dapat diasumsikan bahwa persepsi emosi manusia mengikuti tren yang sama.

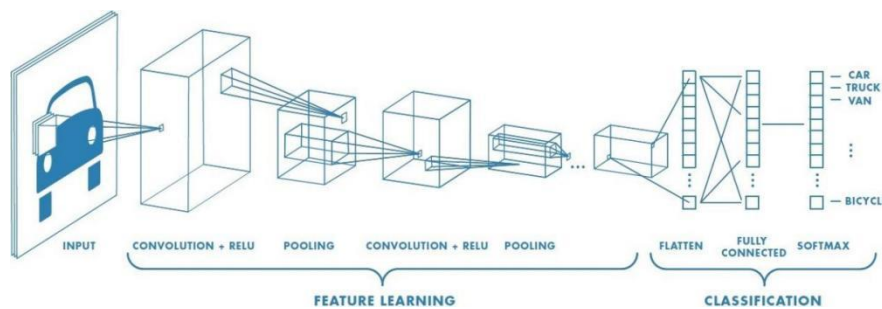
Fungsi deteksi wajah menganalisis ekspresi wajah dalam video dan foto, mengungkapkan ekspresi mikro yang mendefinisikan emosi umum, seperti kegembiraan, kesedihan, kemarahan, ketakutan, kejutan, penghinaan dan jijik. Kebahagiaan adalah ekspresi emosional paling umum yang dimanifestasikan seseorang. Dalam fisiologi, senyum adalah ekspresi wajah yang terbentuk ketika otot-otot ini ditekuk, terutama di kedua ujung mulut. Ekspresi sedih menyampaikan pesan terkait kehilangan, ketidaknyamanan, rasa sakit, ketidakberdayaan. Kemarahan adalah respons alami terhadap ancaman tertentu. Akibatnya, agresi terkadang merupakan respons yang cocok untuk kemarahan. Ekspresi ketakutan menyampaikan informasi tentang bahaya yang akan segera terjadi, ancaman yang akan segera terjadi, kecenderungan untuk melarikan diri atau ekspresi terkejut

hampir selalu terjadi sebagai respons terhadap peristiwa yang tidak terduga, dan mereka mengirimkan pesan tentang sesuatu yang tidak terduga, tiba-tiba, baru atau mengejutkan (Ivanova & Borzunov, 2020).

2.2 Arsitektur CNN

Convolutional neural network (CNN) adalah tipe khusus dari neural network untuk memproses data yang mempunyai topologi jala atau *grid-like topology*. Nama convolutional neural network mengindikasikan bahwa jaringan tersebut menggunakan operasi matematika yang disebut konvolusi yang merupakan sebuah operasi linear. Dan merupakan jaringan saraf yang menggunakan konvolusi sebagai pengganti perkalian matriks umum, setidaknya satu dari convolutional layer tersebut.

Secara umum arsitektur dari convolutional neural network dibagi menjadi dua bagian besar, Feature Extraction atau Feature Learning dan Classification. Selain itu, arsitektur Convolutional Neural Network mempunyai tiga tipe lapisan utama yaitu: convolutional layer, pooling layer dan fully connected layer. Fully-connected layer yang terakhir disebut sebagai output layer. Variasi arsitektur Convolutional Neural Network untuk klasifikasi gambar/citra seperti LeNet, AlexNet, VGG, GoogLeNet / Inception, ResNet dan banyak lagi.

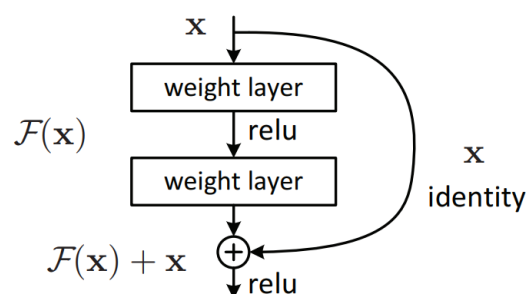


Gambar 1. Arsitektur CNN

(Sumber : <https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html>)

2.3 Residual Network (ResNet)

Residual Network Residual Network merupakan salah satu jenis arsitektur Convolutional Neural Network yang dapat mengatasi masalah vanish gradient pada arsitektur Residual Network. ResNet menggunakan sistem melewati bagian komputasi perkalian pada model sehingga sampai ke bagian layer pertama sehingga bisa menghindari kehilangan informasi penting dari gambar (He & Sun, 2015).



Gambar 2. Arsitektur ResNet : Blok Residual

2.4 Tinjauan 1

Pengenalan emosi melalui deteksi ekspresi wajah adalah salah satu bidang studi penting untuk interaksi manusia-komputer. Untuk mendeteksi ekspresi wajah satu sistem perlu menemukan berbagai variabilitas wajah manusia seperti warna, postur, ekspresi, orientasi, dll. Untuk mendeteksi ekspresi wajah manusia terlebih dahulu diperlukan untuk mendeteksi fitur wajah yang berbeda seperti gerakan mata, hidung, bibir, dll. dan kemudian mengklasifikasikannya dibandingkan dengan data terlatih menggunakan pengklasifikasi yang sesuai untuk pengenalan ekspresi. Dalam penelitian ini, sistem pengenalan ekspresi wajah manusia dimodelkan menggunakan pendekatan eigenface. Metode yang diusulkan menggunakan model warna HSV (Hue-Saturation-Value) untuk mendeteksi wajah dalam sebuah gambar. PCA telah digunakan untuk mengurangi dimensi tinggi dari eigenspace dan kemudian dengan memproyeksikan gambar uji pada eigenspace dan menghitung jarak Euclidean antara gambar uji dan rata-rata eigenfaces dari dataset pelatihan ekspresi diklasifikasikan. Sebuah dataset generik digunakan untuk tujuan pelatihan. Gambar skala abu-abu dari wajah digunakan oleh sistem untuk mengklasifikasikan lima emosi dasar seperti terkejut, sedih, takut, marah dan bahagia. Pelatihan PCA digunakan untuk menghasilkan wajah eigen. Sebuah database dibuat untuk formasi eigenface yang terdiri dari 30 gambar wajah yang mewakili ekspresi seperti terkejut, marah, takut, sedih dan bahagia. Sebuah $M \times N$ ruang dimensi dibuat menggunakan matriks kovarians dari 30 gambar ini yang kemudian direduksi menjadi K dimensi menggunakan PCA.

Makalah ini mengusulkan model pengenalan ekspresi wajah manusia berdasarkan pendekatan eigenface di mana berbagai emosi dikenali dengan menghitung jarak Euclidean antara gambar uji input dan rata-rata eigenfaces dari dataset pelatihan. Dataset pelatihan terdiri dari gambar orang yang berbeda dan ketika diuji memberikan hasil yang memuaskan tetapi ada kemiripan antara Kesedihan dan Ketakutan sampai batas tertentu yang dapat dianggap sebagai pekerjaan masa depan dan dapat ditingkatkan dengan pelatihan yang lebih ekstensif. Bidang penelitian dalam pengenalan ekspresi adalah area yang dapat dieksplorasi dan ditingkatkan lebih lanjut (De et al., 2015).

2.5 Tinjauan 2

Selama dekade terakhir, peningkatan perhatian telah diarahkan pada studi FER. Makalah ini menyajikan ulasan tentang FER yang menangani ekstraksi fitur wajah dan klasifikasi ekspresi wajah, yang merupakan dua langkah terpenting dalam sistem FER. Untuk metode ekstraksi fitur wajah, metode berbasis fitur geometris dan metode berbasis penampilan, yang digunakan untuk gambar statis, ditinjau terlebih dahulu. Kemudian, metode ekstraksi fitur wajah ini untuk urutan gambar dinamis, termasuk aliran optik dan pelacakan titik fitur, juga diselidiki. Untuk klasifikasi ekspresi wajah, enam metode klasifikasi tipikal termasuk HMM, ANN, BN, KNN, SVM dan SRC, juga disurvei. Selain itu, penelitian ini melakukan eksperimen FER pada database JAFFE dan database Cohn-Kanade Kedua dimana database ini berisi tujuh ekspresi wajah, yaitu marah, senang, sedih, netral, terkejut, jijik dan takut, dan menyajikan studi perbandingan metode klasifikasi yang berbeda berdasarkan fitur LBP yang diekstraksi. Hasil percobaan menunjukkan bahwa SRC mengungguli metode lain yang digunakan seperti KNN, SVM, HMM, naive-Bayes, dan ANN. Kesimpulan dan pekerjaan masa depan akhirnya dibahas di bagian terakhir dari survei ini.

Meskipun upaya ekstensif telah dikhususkan untuk FER dan banyak keberhasilan baru-baru ini telah dicapai, seperti yang disebutkan di atas, banyak pertanyaan masih terbuka. Menurut pendapat kami, beberapa poin berikut harus dipertimbangkan di masa depan.

(1) Bagaimana cara manusia mengidentifikasi ekspresi wajah dengan benar?

Sejauh ini, penelitian psikologis dan medis tentang persepsi dan kognisi manusia telah berlangsung lama, tetapi masih ambigu bagaimana manusia mengidentifikasi ekspresi wajah. Jenis parameter apa yang dapat digunakan oleh manusia dan bagaimana mereka diproses?

(2) Bagaimana kita mengidentifikasi ekspresi wajah dalam pemandangan dunia nyata?

Karena deformasi wajah yang halus, gerakan kepala yang sering, dan pengukuran gerakan wajah yang ambigu dan tidak pasti, mengidentifikasi ekspresi wajah spontan dalam pemandangan dunia nyata jauh lebih sulit daripada yang dipelajari secara luas oleh FER hingga saat ini. Mengingat fakta bahwa ekspresi wajah spontan dapat dicirikan oleh gerakan kepala yang kaku dan gerakan otot wajah yang tidak kaku, penelitian terbaru [58] dalam pemodelan pengenalan gerakan kepala spontan dan pengenalan unit aksi dalam ekspresi wajah spontan merupakan perkembangan yang menarik. Mereka mengembangkan model aksi wajah probabilistik terpadu [58] untuk secara bersamaan dan koheren mewakili gerakan wajah kaku dan tidak kaku, serta ketergantungan spatiotemporal mereka. Bagaimana mengelaborasi model aksi wajah probabilistik terpadu diperlukan untuk pekerjaan seperti itu, yang masih menjadi pertanyaan penelitian. Selain itu, banyak pekerjaan sebelumnya yang terlibat dalam masalah privasi karena kumpulan data standar yang ada adalah kumpulan data berbasis pose. Untuk mengatasi masalah privasi, karya terbaru [59] adalah arah penting, di mana sistem FER berbasis kamera kedalaman menggunakan skema multilayer dikembangkan.

(3) Bagaimana kita secara otomatis mempelajari fitur wajah yang lebih efektif untuk pengenalan ekspresi wajah?

Perlu ditunjukkan bahwa metode ekstraksi fitur yang dirancang dengan tangan di atas biasanya bergantung pada operasi manual dengan data berlabel. Dengan kata lain, metode ini diawasi. Selain itu, fitur yang dirancang dengan tangan seperti representasi wavelet LBP dan Gabor ini mampu menangkap informasi tingkat rendah dari gambar wajah, kecuali untuk representasi gambar wajah tingkat tinggi. Dalam beberapa tahun terakhir, pembelajaran mendalam [60-63], sebagai teori pembelajaran mesin yang baru-baru ini muncul, didasarkan pada arsitektur hierarkis pemrosesan informasi dalam sistem persepsi visual primata, dan telah menunjukkan bagaimana hierarki fitur dapat dipelajari secara langsung dari data asli tanpa pengawasan. Cara menggunakan teknik pembelajaran mendalam untuk mempelajari fitur wajah yang lebih efektif secara otomatis merupakan arahan penting bagi FER.

(4) Bagaimana kita dapat mengintegrasikan analisis ekspresi wajah dengan modalitas lain?

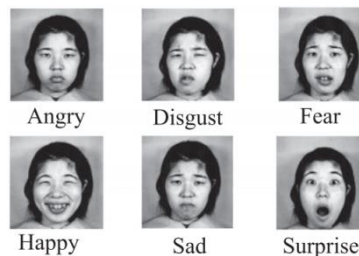
Emosi mentransfer informasi psikologis manusia, karena emosi disampaikan oleh berbagai perubahan fisiologis, seperti perubahan detak jantung, kecepatan, derajat keringat, tekanan darah, dll. Emosi juga diungkapkan dengan ucapan afektif, ekspresi wajah, gerak tubuh, dan sebagainya. Untuk mempromosikan kinerja pengenalan emosi, mengintegrasikan beberapa modalitas afektif, seperti informasi bicara, wajah, fisiologis dan leksikal, adalah subjek yang sangat aktif [64-67] dalam beberapa tahun terakhir. Namun demikian, bagaimana mengintegrasikan secara efektif modalitas ekspresi emosi yang heterogen untuk lebih meningkatkan kinerja pengenalan emosi multimodal masih menjadi pertanyaan terbuka. (Zhao & Zhang, 2016)

2.6 Tinjauan 3

Pada penelitian Revina dan Emmanuel tahun 2021, telah menjelaskan tentang survei teknik Face Expression Recognition (FER) yang mencakup tiga tahap utama seperti preprocessing, ekstraksi fitur dan klasifikasi. Hanya teknik FER berbasis gambar yang dipilih untuk tinjauan literatur dan teknik FER berbasis video tidak dipilih. Sebagian besar sistem FER memenuhi masalah variasi dalam pencahayaan, variasi pose, variasi pencahayaan, variasi warna kulit.

Analisis algoritma pada 3 tahap : preprocessing, ekstraksi fitur dan klasifikasi. Berbagai metode preprocessing yang digunakan adalah face detection, image enhancement, normalization, Gabor filter, localization, face acquisition, down sampling, histogram equalization, face region detection, face alignment, ROI segmentation and resizing. Berbagai metode ekstraksi fitur yang digunakan dalam tabel ini adalah LEM, Model berbasis aksi, filter Gabor, LBP-TOP, GASM, Berbasis patch, wavelet GL, LBP, VTB, Momen, PCA, ICA, LCT, HOG, Piramida yang dapat dikendalikan, DCT, SWLDA, WLD, SDM, WPLBP, haar like features, LDN, LDTP, DWT, K-ELBP, 2DPCA, dan eigenfaces. Classifier yang digunakan pada tabel ini adalah ID3 decision tree, LVQ, SVM, KNN, HMM, MFFNN, OSLEM, Bayesian neural network, HCRF, pairwise, CART, Euclidean distance, CNN.

Deskripsi database : JAFFE, CK, CK+, MMI, MUG, TFEID, AR, Yale, KDEF (Karolinska Directed Emotional Faces), Real-time dan dataset sendiri. Sebagian besar percobaan menggunakan database JAFFE. JAFFE menyimpan sepuluh ekspresi wanita Jepang dengan tujuh ekspresi wajah dan total 213 gambar. Setiap gambar dalam database JAFFE berisi resolusi 256 - 256 piksel.



Gambar 3. Contoh gambar dari database JAFFE

Database CK juga memiliki tujuh ekspresi tetapi berisi 132 subjek yang berpose dengan natural dan tersenyum. Ini berisi total 486 urutan gambar dengan resolusi 640 - 490 piksel gambar abu-abu.



Gambar 4. Contoh gambar dari database CK

Makalah survei ini berguna bagi pengembang perangkat lunak untuk mengembangkan

algoritma berdasarkan akurasi dan kompleksitas serta membantu untuk implementasi perangkat keras dengan biaya rendah tergantung pada kebutuhan mereka. Survei ini membandingkan algoritma berdasarkan preprocessing, ekstraksi fitur, klasifikasi dan kontribusi utama. Analisis kinerja dilakukan berdasarkan database, tingkat kompleksitas, akurasi pengenalan dan kontribusi utama. Survei ini membahas properti seperti ketersediaan preprocessing dan ekstraksi fitur dan jumlah ekspresi. Kekuatan algoritma, keuntungan dibahas secara rinci untuk mencapai tujuan survei ini. Metode segmentasi ROI digunakan untuk preprocessing dan memberikan akurasi tertinggi 99%. Menurut ekstraksi fitur, GF memiliki kompleksitas yang lebih kecil yang memberikan akurasi selalu antara 82,5% dan 99%. Akurasi pengenalan tertinggi 99% disediakan oleh pengklasifikasi SVM dan mengenali beberapa ekspresi seperti jijik, sedih, senyum, terkejut, marah, takut, netral secara efektif. Dalam 2D FER, sebagian besar database JAFFE dan CK digunakan untuk kinerja yang efisien daripada database lainnya (Revina & Emmanuel, 2021).

2.7 Tinjauan 4

Pendekatan untuk memecahkan masalah FER (Facial Expression Recognition) dapat dikategorikan menjadi 1) Gambar tunggal statis dan 2) Gambar urutan. Secara tradisional, teknik yang berbeda seperti Multi-layer Perceptron Model, k-Nearest Neighbours, Support Vector Machines digunakan oleh para peneliti untuk memecahkan FER. Metode ini mengekstrak fitur seperti Local Binary Patterns, Eigenfaces, Face-landmark features, dan Texture features. Di antara semua metode ini, Neural Networks telah mendapatkan popularitas yang sangat tinggi dan mereka banyak digunakan untuk FER. CNN (Convolutional Neural Networks) telah mendapatkan popularitas di bidang pembelajaran mendalam karena arsitektur kasual dan kemampuan mereka untuk memberikan hasil yang baik tanpa persyaratan ekstraksi fitur manual dari data gambar mentah. Makalah ini berfokus pada survei berbagai teknik pengenalan ekspresi wajah berdasarkan CNN. Pengenalan ekspresi wajah menggunakan deep learning melalui tahap augmentasi dan pemrosesan dan klasifikasi. Dataset yang banyak digunakan adalah FER-13, JAFFE, dan CK+. Data telah menampilkan gambar dari tujuh ekspresi prototypic yaitu: Happy, Sad, Disgust, Fear, Surprise, Angry dan Neutral.

Hasil yang terkait dengan arsitektur CNN : CNN yang ringkas dengan jumlah lapisan dan hyperparameter yang lebih sedikit lebih disukai untuk gambar wajah bagian depan, CNN yang kompleks dengan lebih banyak jumlah lapisan dan parameter lebih disukai untuk wajah spontan dan CNN moderat dengan jumlah lapisan dan parameter yang seimbang dengan pra-pemrosesan yang tepat lebih disukai untuk gambar yang terhalang. Disimpulkan bahwa metode berbasis ekstraksi ciri tradisional memakan waktu untuk memilih fitur yang sesuai untuk pembelajaran model. CNN secara otomatis mempelajari fitur-fitur tersebut secara efisien dan CNN dapat menjadi pilihan tepat untuk skenario dunia nyata (Suma et al., 2018).

2.8 Tinjauan 5

Pengenalan emosi dari gerakan tubuh memiliki keunggulan mengenali emosi seseorang dari kamera apapun melihat dan juga mengenali emosi, jika orang terlalu jauh dari kamera. Gerakan tubuh dapat dengan kuat menyampaikan keadaan emosional daripada penelitian lain. Dalam makalah ini, keadaan emosi dikenali pola gerak seluruh tubuh menggunakan arsitektur jaringan syaraf tiruan feedforward deep convolution dengan parameter yang berbeda. Sistem yang diusulkan dapat dievaluasi dengan dataset emosi (University of .YORK) dengan 15 jenis

emosi dan dataset corpus GEMEP dengan 5 emosi. Dalam makalah ini, diusulkan FDCNN untuk memprediksi emosi manusia dari gerakan tubuh pada urutan frame. Model ini mewakili fitur deep convolutional untuk mengekstrak informasi saliency pada berbagai skala. Kinerja model ini lebih baik daripada model dasar. Pekerjaan di masa depan bertujuan untuk mengembangkan aplikasi penelitian untuk mengenali emosi anak-anak dengan gangguan spektrum autisme (ASD) (Santhoshkumar & Geetha, 2019).

2.9 Tinjauan 6

Saat ini, ada kebutuhan di masyarakat untuk menggunakan metode digital dan siberetik untuk diagnostik jarak jauh dari kondisi mental seseorang. Sistem digital atau cyberphysical dalam proses pengenalan jarak jauh dari keadaan mental pasien, dimungkinkan untuk meningkatkan efektivitas mengenali keadaan psiko-emosional pasien yang sebenarnya, baik dalam mode konsultasi maupun dalam pengamatan dinamis sistem. Untuk penilaian jarak jauh keadaan mental, metode banyak digunakan berdasarkan pengenalan aktivitas wajah, yang secara refleks dikaitkan dengan tampilan keadaan psiko-emosional. Metodologi yang diusulkan oleh penulis dalam karya ini untuk menilai keadaan psiko-emosional seseorang tampaknya sangat berguna dalam menilai kesiapan fungsional orang dalam profesi yang bertanggung jawab dan berbahaya. Selain itu, pendekatan ini akan memberikan kontribusi yang signifikan untuk pemantauan jarak jauh pasien dengan penyakit mental, yang sangat penting dalam konteks epidemi dan karantina paksa (Lebedev et al., 2020).

2.10 Tinjauan 7

Penelitian lain mengatasi masalah deteksi keadaan emosional dari ekspresi wajah. Pendekatan yang diusulkan secara bersamaan mendeteksi wajah dan memprediksi kategori emosi yang berbeda dan nilai valensi berkelanjutan dari gambar masukan mentah. Peneliti melatih dan mengevaluasi pendekatan tersebut pada 3 kumpulan data yang berbeda, membandingkan pendekatan tersebut dengan pendekatan canggih lainnya dan melakukan evaluasi lintas basis data. Dengan cara ini ditemukan bahwa pendekatan tersebut dapat digeneralisasi dengan baik dan cocok untuk aplikasi waktu nyata. Oleh karena itu, penelitian ini mengusulkan pendekatan yang didasarkan pada arsitektur YOLO (you-only-look-once). CNN telah dirancang untuk secara bersamaan memprediksi kotak pembatas wajah, emosi dasar, dan nilai valensi (VA). Dibandingkan dengan jaringan canggih yang sebanding, CNN kami kecil dan hanya terdiri dari sepuluh lapisan konvolusi. Akibatnya, ini cocok untuk aplikasi waktu nyata bahkan dengan banyak kamera. Memang, saat ini kami menerapkan CNN yang diusulkan untuk menganalisis ekspresi wajah untuk skenario kerjasama manusia - robot, untuk mendeteksi ketakutan atau frustrasi. Kami mengevaluasi kemampuan jaringan yang diusulkan untuk memprediksi nilai valensi (VA) dan kelas emosi diskrit pada database AffectNet. Untuk prediksi VA, evaluasi lintas basis data pada dataset Aff-Wild dan AFEW-VA juga dilakukan, yang menunjukkan bahwa pendekatan kami dapat digeneralisasi dengan baik dan menghasilkan prediksi yang andal bahkan di luar kondisi laboratorium. Meskipun sampel pelatihan AffectNet sudah menunjukkan berbagai variasi pencahayaan, pose lebih, kami selanjutnya mendapatkan variasi dengan augmentasi data, yang mengarah ke kesalahan prediksi yang sedikit berkurang. Selanjutnya, dengan menggabungkan sampel pelatihan dari dua atau tiga database, kami menunjukkan bahwa hasil – saat pengujian pada Aff-Wild dan AFEW-VA– dapat ditingkatkan dengan cara ini (Handrich et al., 2020).

2.11 Tinjauan 8

Karya ini bertujuan untuk mengklasifikasikan ekspresi emosi penyandang disabilitas fisik (tuli, bisu, dan terbaring di tempat tidur) dan anak autisme berdasarkan penanda wajah dan sinyal electroencephalograph (EEG) menggunakan pengklasifikasi convolutional neural network (CNN) dan long short-term memory (LSTM). Enam emosi wajah (senang, sedih, marah, takut, jijik, dan terkejut) dikumpulkan menggunakan sepuluh penanda virtual. 55 mahasiswa sarjana (35 laki-laki dan 20 perempuan) dengan usia rata-rata 22,9 tahun secara sukarela berpartisipasi dalam percobaan untuk pengenalan emosi wajah. Sembilan belas mahasiswa sarjana mengajukan diri untuk mengumpulkan sinyal EEG. Awalnya, fitur seperti Haar digunakan untuk deteksi wajah dan mata. Kemudian, penanda virtual ditempatkan pada lokasi yang ditentukan pada wajah subjek berdasarkan sistem pengkodean tindakan wajah menggunakan pendekatan model matematis, dan penanda tersebut dilacak menggunakan algoritma aliran optik Lucas-Kandere. Jarak antara pusat wajah subjek dan setiap posisi penanda digunakan sebagai fitur untuk klasifikasi ekspresi wajah. Fitur jarak ini divalidasi secara statistik menggunakan analisis variansi satu arah dengan tingkat signifikansi $p < 0,01$. Selain itu, empat belas sinyal yang dikumpulkan dari saluran pembaca sinyal EEG (EPOCp) digunakan sebagai fitur untuk klasifikasi emosional menggunakan sinyal EEG. Terakhir, fitur divalidasi silang menggunakan validasi silang lima kali lipat dan diberikan kepada pengklasifikasi LSTM dan CNN. Hasil mencapai tingkat pengenalan maksimum 99,81% menggunakan CNN untuk deteksi emosi menggunakan landmark wajah. Namun, tingkat pengenalan maksimum yang dicapai dengan menggunakan pengklasifikasi LSTM adalah 87,25% untuk deteksi emosi menggunakan sinyal EEG (Hassoun et al., 2020).

2.12 Tinjauan 9

Karya ini dikhususkan untuk optimalisasi metode pengenalan tujuh emosi dasar (gembira, sedih, takut, marah, terkejut, jijik, dan netral) dalam hal ekspresi wajah manusia. Pendekatan yang ada dari konstruksi sistem pengenalan emosi dianalisis berdasarkan ekspresi wajah manusia, dan difokuskan pada keuntungan menggunakan skema berdasarkan jaringan saraf. Sistem pengenalan emosi pengenalan wajah terdiri dari dua modul utama: ekstraksi fitur dan pengenalan emosi. Deteksi wajah bertujuan untuk mengetahui keberadaan suatu objek dan mengetahui posisinya dalam sistem koordinat piksel citra asli. Selain itu, memecahkan masalah deteksi memungkinkan memperoleh informasi tentang posisi relatif objek dalam gambar dan kemudian menggunakannya untuk analisis. Untuk mengatasi masalah pengenalan wajah dalam karya digunakan metode Viola-Jones, algoritma yang optimal dengan kecepatan tinggi dan akurasi pengenalan. database ORL untuk wajah yang terdiri dari 400 gambar berukuran 112 x 92. Masing-masing berisi 40 orang, masing-masing 10 gambar. Gambar diambil pada waktu yang berbeda, dengan pencahayaan dan ekspresi wajah. Wajah berada dalam posisi tegak pada tampilan depan, dengan sedikit berbelok ke kiri-kanan.

Dataset FER2013, karena berada dalam domain publik Kaggle dan ditandai sesuai dengan tujuh emosi dasar, sehingga digunakan untuk melatih model. FER2013 berisi 28.709 gambar pelatihan, 3.589 gambar uji dan 3.589 gambar uji, yang dibagi menjadi tujuh bagian, sesuai dengan tujuh emosi: marah, jijik, takut, bahagia, sedih, takut, dan netralitas. Untuk tugas mengenali emosi berdasarkan wajah manusia, kami menggunakan jaringan saraf berdasarkan GoogLeNet. GoogLeNet dikembangkan oleh Christian Szegedy dan peneliti Google. Modul baru telah ditambahkan ke jaringan saraf convolutional - Inception, yang merupakan kombinasi paralel dari filter convolutional 3x3, 5x5, 1x1. Filter 1x1 mendeteksi korelasi antar

saluran, sementara filter yang lebih besar merespons lebih banyak fitur global, dan lapisan penarik memungkinkan seseorang untuk mengurangi dimensi tanpa kehilangan informasi yang besar. Selain itu, untuk mengurangi dimensi, lapisan konvolusi 1x1 ditambahkan ke setiap filter, yang mengurangi jumlah peta fitur untuk menghindari peningkatan jumlah parameter.

Solusi desain yang ditunjukkan berhasil diimplementasikan, eksperimen komputasi mengkonfirmasi validitasnya. Hasil utama dari penelitian ini adalah model yang dikembangkan dari CNN dan program untuk mengenali emosi dengan ekspresi wajah. Model yang dibangun dari CNN disetel: parameter lapisan jaringan ditetapkan untuk memastikan fungsi kompleks perangkat lunak yang paling efisien. Proporsi pengenalan yang benar untuk sampel uji publik adalah 69%, untuk sampel uji privat 59%. Akurasi pengenalan indikator juga ditentukan untuk setiap kelompok emosi. Akurasi pengenalan tertinggi dicatat pada gambar dengan emosi "kebahagiaan" (88%), terendah pada gambar dengan emosi "kesedihan" (57%) (Ivanova & Borzunov, 2020).

2.13 Tinjauan 10

Deteksi emosi manusia adalah salah satu masalah utama dalam visi komputer. Emosi manusia terdiri dari beberapa sub-emosi yang sulit untuk diklasifikasikan ke dalam kelas tertentu. Penelitian ini telah mencoba mengklasifikasikan emosi ke dalam 6 kategori dasar (senang, sedih, jijik, takut, terkejut, marah) dan emosi manusia yang netral. Karya ini telah mengusulkan kerangka pembelajaran mendalam yang terdiri dari CNN, ResNet dan blok perhatian yang memberikan persepsi visual ke jaringan. Model yang diusulkan memiliki penerapan yang lebih besar dalam deteksi emosi wajah kehidupan nyata. Model yang diusulkan telah mencapai hasil yang memuaskan dan telah menunjukkan hasil yang efektif pada dataset FER.

Dalam makalah ini kami telah mengusulkan jaringan perhatian diri yang mendalam untuk pengenalan emosi wajah yang terdiri dari blok perhatian, koneksi residual, dan jaringan konvolusi. Setiap blok jaringan memiliki fungsi yang berbeda untuk dilakukan karena koneksi residual membantu menghilangkan masalah gradien yang hilang, jaringan konvolusi mengekstraksi fitur dan perhatian telah terbukti efektif karena memberikan persepsi visual yang lebih baik ke jaringan. Model yang diusulkan telah mengungguli jaringan berbasis CNN yang ada dengan mencapai akurasi yang lebih tinggi yaitu 85,76 pada fase pelatihan dan 64,40 pada fase pengujian. Model yang diusulkan telah menunjukkan hasil yang lebih baik pada dataset FER dan dapat digunakan dalam aplikasi waktu nyata. Selain itu, jaringan ini dapat digunakan dan diperluas dalam tugas visi komputer lainnya (Gupta et al., 2020).

2.14 Tinjauan 11

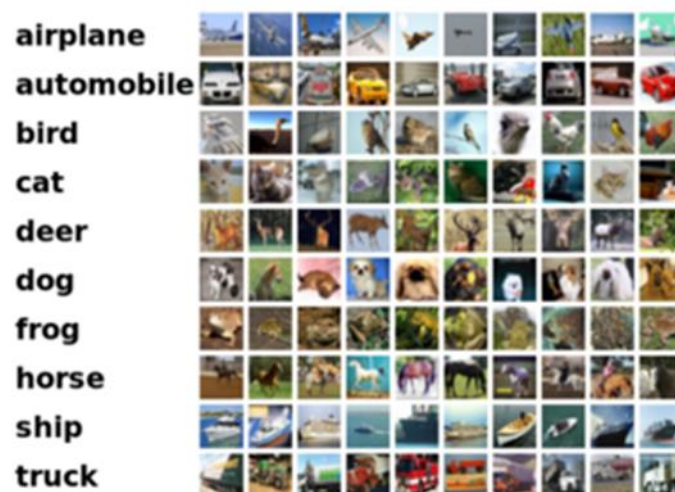
Dalam makalah ini, kami mengusulkan algoritma evolusioner baru untuk secara otomatis membuat arsitektur jaringan saraf. Algoritma yang diusulkan menggunakan blok bangunan dasar dari jaringan canggih yang ada untuk menginisialisasi populasi. Selama proses evolusi, algoritma menghasilkan individu baru (yaitu, jaringan baru) yang dibangun dengan memanfaatkan blok yang berbeda dari jaringan yang berbeda. Selain itu, skema pengkodean, operator genetik, dan fungsi kebugaran dirancang dengan baik. Akibatnya, arsitektur jaringan yang dihasilkan, yang memiliki sifat heterogen yang kuat, cenderung mengintegrasikan keunggulan jaringan canggih yang ada.

Jaringan saraf (NN), terutama jaringan saraf dalam (DNN), telah terbukti berkinerja baik pada banyak tugas kecerdasan buatan yang menantang. Dalam literatur, berbagai jenis arsitektur jaringan telah dirancang secara manual oleh para ahli yang memiliki pengetahuan domain yang kaya tentang NN, seperti GoogLeNet, Residual Neural Nets (ResNet) dan Residual Inception Nets. Di antara mereka, ResNet memperoleh hasil yang sangat baik pada masalah klasifikasi gambar.

Sebagai jenis algoritma pencarian stokastik yang terinspirasi oleh alam, EA telah terbukti secara empiris memiliki kinerja yang baik dalam memecahkan masalah optimasi dunia nyata. Namun, salah satu kelemahan utama adalah bahwa kebanyakan EA memerlukan sejumlah besar evaluasi fungsi untuk menemukan solusi yang optimal. Khususnya, saat menerapkan EA untuk mengoptimalkan parameter hiper dan/atau arsitektur sebuah NN, evaluasi satu fungsi perlu menghabiskan banyak biaya untuk melatih setiap solusi kandidat selama evolusi (terutama dalam kasus DNN). Dengan demikian, masalah bagaimana menerapkan EA untuk optimasi NN dengan cara yang efisien masih merupakan area penelitian yang kaya.

Dalam makalah ini, mengingat kinerja ResNet yang tinggi, kami berusaha untuk mempromosikan efisiensi EA dengan bantuan ResNet. Secara konkret, kami mengusulkan algoritma evolusioner baru untuk secara otomatis membuat arsitektur jaringan saraf, yang disebut EANN untuk jangka pendek. Algoritma yang diusulkan menggunakan blok bangunan dasar yang diturunkan dari struktur ResNet untuk menginisialisasi populasi. Sebagai hasil dari mengasimilasi keunggulan ResNet dan EA, EANN yang diusulkan mampu menemukan arsitektur jaringan yang efektif dengan sejumlah evaluasi fungsi yang terbatas.

Dalam situasi pengujian yang diberikan menggunakan dataset CIFAR-10 32×32 gambar beranotasi. Dataset terdiri dari subset 60000 gambar dari 80 juta dataset gambar kecil. Subset ini dibagi menjadi 50000 gambar pelatihan dan 10.000 gambar uji. Total ada 10 kelas. Gambar pelatihan terdiri dari 5000 gambar per kelas. Gambar uji berisi tepat 1000 gambar per kelas. Gambar 4 mencantumkan 10 contoh gambar dari 10 kelas beranotasi.



Gambar 5. Contoh gambar dari 10 kelas beranotasi dalam CIFAR-10

EANN-Net mencapai kesalahan $e=8.86\%$ sementara ResNet-44 mendapatkan kesalahan $e=11.64\%$. Ini berarti bahwa EANN-Net dapat mengklasifikasikan data uji dengan akurasi 2,78 poin persentase lebih baik daripada ResNet-44 yang digunakan sebagai jaringan referensi.

Dalam makalah ini, kami telah mengusulkan algoritma EANN untuk membuat arsitektur jaringan saraf secara otomatis. EANN yang diusulkan mampu menemukan arsitektur jaringan yang efektif dengan jumlah evaluasi fungsi yang terbatas. arsitektur jaringan yang dikembangkan dalam algoritma EANN berbeda secara struktural dari arsitektur *state-of-the-art* lainnya karena mereka tidak dibangun secara homogen dari pengulangan struktur dasar yang sama, tetapi secara heterogen terdiri dari blok yang berbeda. Selain itu, EANN-Net yang dihasilkan dengan properti heterogen dapat dilatih pada kumpulan data dengan iterasi pelatihan yang lebih sedikit daripada arsitektur yang sebanding. Dalam penelitian masa depan, pilihan hyperparameter pelatihan yang dibuat khusus untuk EANN-Net dan algoritma evolusioner yang lebih maju harus diselidiki untuk lebih meningkatkan hasil klasifikasi pada kumpulan data yang lebih besar (Chen et al., 2019).

2.15 Tinjauan 12

Sebagai salah satu arahan terpenting di bidang visi komputer, pengenalan emosi wajah memainkan peran penting dalam pekerjaan dan kehidupan sehari-hari orang. Pengenalan emosi manusia berdasarkan ekspresi wajah sangat penting dalam penerapan interaksi manusia-komputer yang cerdas. Namun, dalam penelitian saat ini tentang pengenalan emosi wajah, ada beberapa masalah seperti kemampuan generalisasi yang buruk dari model jaringan dan kekokohan sistem pengenalan yang rendah. Dalam konten ini, kami mengusulkan metode ekstraksi fitur menggunakan jaringan sisa dalam ResNet-50, yang menggabungkan CNN untuk pengenalan emosi wajah. Melalui simulasi eksperimental dari kumpulan data yang ditentukan.

Melalui analisis literatur, dapat ditemukan bahwa fitur emosional wajah yang diekstraksi dengan metode di atas memiliki masalah bahwa informasi emosional asli mudah hilang. Selain itu, generalisasi dan ketahanan model jaringan ini juga buruk dan akurasi pengenalan ekspresi wajah tidak tinggi. Untuk mengatasi masalah tersebut, kami mengusulkan model pengenalan emosi wajah yang ditingkatkan. Kami menggunakan ResNet-50 sebagai infrastruktur jaringan. Fitur diekstraksi oleh CNN, BN dan fungsi aktivasi ReLU digunakan untuk meningkatkan kemampuan konvergensi model. Melalui percobaan simulasi di bawah ini, dapat diverifikasi bahwa kinerja pengenalan metode pengenalan emosi wajah yang diusulkan lebih baik daripada metode yang paling canggih.

Dalam makalah ini, sistem pengenalan ekspresi wajah yang ditingkatkan dipelajari dan metode pengenalan ekspresi wajah berdasarkan jaringan sisa dalam diusulkan. Makalah ini berfokus pada proses pembelajaran pengenalan ekspresi wajah. Kami menggunakan algoritma CNN populer saat ini, dikombinasikan dengan jaringan residual ResNet-50, yang telah mencapai efek yang baik dalam tugas multi-klasifikasi. Melalui validasi kumpulan data, hasil eksperimen menunjukkan bahwa metode yang diusulkan dalam makalah ini memiliki akurasi yang baik dan efek pengenalan yang baik dalam hal akurasi pengenalan rata-rata. Dalam percobaan ini, jaringan ResNet-50 telah mencapai hasil yang baik dibandingkan dengan metode lain. Selanjutnya, kami akan mencoba menggunakan lapisan yang berbeda dari ResNet serta beberapa varian ResNet, seperti Wide Residual Network (WRN), ResNeXt dan MobileNet untuk mempelajari pengenalan emosi wajah. Dalam pekerjaan masa depan, kami akan menguji kinerja mereka.

Pada penelitian selanjutnya, kami akan fokus pada penelitian pengenalan emosi wajah dan mencoba mengumpulkan lebih banyak gambar emosional daripada dalam percobaan ini, sehingga dapat mengoptimalkan dan mengusulkan algoritma yang lebih baik untuk melatih hyperparameter jaringan saraf umpan maju multi-layer, seperti bobot dan bias. Dan kami juga

akan mencoba algoritma optimasi tersebut berdasarkan metode yang disebutkan di atas untuk meningkatkan kinerja jaringan saraf feedforward multi-layer. Kami akan terus mencari cara berdasarkan jaringan sisa yang dalam untuk meningkatkan akurasi pengenalan ekspresi wajah (Li & Lima, 2021).

2.16 Perbandingan Tinjauan

Perbandingan penelitian dengan melihat kelebihan kekurangan dari tiap-tiap tinjauan, dan pilihan mana yang digunakan atau diadaptasi terkait Pengenalan Ekspresi Wajah pada jurnal tahun 2015 sampai 2021 seperti yang terlihat pada tabel 1 sebagai berikut :

Tabel 1. Perbandingan Tinjauan

No	Peneliti / tahun / Judul	Dataset	Metode	Hasil/Keterbatasan
1.	(De et al., 2015) / A Human Facial Expression Recognition Model based on Eigen Face Approach	Database berisi 30 gambar	Pelatihan PCA digunakan untuk membangkitkan eigen faces	memberikan hasil yang memuaskan tetapi ada kemiripan antara Kesedihan dan Ketakutan sampai batas tertentu yang dapat dianggap sebagai pekerjaan masa depan dan dapat ditingkatkan dengan pelatihan yang lebih ekstensif.
2.	(Zhao & Zhang, 2016) / A Review on Facial Expression Recognition: Feature Extraction and Classification	JAFFE, CK : (tujuh ekspresi wajah, yaitu marah, senang, sedih, netral, terkejut, jijik dan takut)	Metode ekstraksi fitur wajah, mulai meninjau metode berbasis fitur geometris dan metode berbasis penampilan, yang digunakan untuk gambar statis, kemudian juga diselidiki metode ekstraksi fitur wajah ini untuk urutan gambar dinamis, termasuk aliran optik dan pelacakan titik fitur. Untuk klasifikasi ekspresi wajah : HMM, ANN, BN, KNN, SVM dan SRC	Menyajikan studi perbandingan metode klasifikasi yang berbeda berdasarkan fitur LBP yang diekstraksi. Hasil percobaan menunjukkan bahwa SRC mengungguli metode lain yang digunakan seperti KNN, SVM, HMM, naive-Bayes, dan ANN.
3.	(Revina & Emmanuel, 2021) / A Survey on Human Face Expression Recognition Techniques	JAFFE, CK, CK+, MMI, MUG, TFEID, AR, Yale, KDEF (Karolinska Directed Emotional Faces), Real-time dan dataset sendiri	Analisis algoritma pada 3 tahap : preprocessing, ekstraksi fitur dan klasifikasi. Metode preprocessing yang digunakan adalah, face detection, image enhancement, ormalization, Gabor filter, localization, face acquisition, down sampling, histogram equalization, face region detection, face alignment, ROI segmentation and resizing. Metode ekstraksi fitur yang digunakan dalam tabel ini adalah LEM, Model	Metode segmentasi ROI digunakan untuk preprocessing dan memberikan akurasi tertinggi 99%. Menurut ekstraksi fitur, GF memiliki kompleksitas yang lebih kecil yang memberikan akurasi selalu antara 82,5% dan 99%. Akurasi pengenalan tertinggi 99% disediakan oleh pengklasifikasi SVM dan mengenali beberapa ekspresi seperti jijik, sedih, senyum, terkejut, marah, takut, netral secara efektif.

No	Peneliti / tahun / Judul	Dataset	Metode	Hasil/Keterbatasan
			berbasis aksi, filter Gabor, LBP-TOP, GASM, berbasis patch, wavelet GL, LBP, VTB, Momen, PCA, ICA, LCT, HOG, Piramida yang dapat dikendalikan, DCT, SWLDA, WLD, SDM, WPLBP, haar like features, LDN, LDTP, DWT, K-ELBP, 2DPCA, dan eigenfaces. Classifier : ID3 decision tree, LVQ, SVM, KNN, HMM, MFFNN, OSLEM, Bayesian neural network, HCRF, pairwise, CART, Euclidean distance, CNN	Dalam 2Dekade FER, sebagian besar database JAFFE dan CK digunakan untuk kinerja yang efisien daripada database lainnya.
4.	(Suma et al., 2018) / Survey on Face Expression Recognition using CNN	FER-13, JAFFE, dan CK+	CNN	Efektivitas CNN untuk klasifikasi ekspresi wajah : <ul style="list-style-type: none"> • CNN yang ringkas dengan jumlah lapisan dan hyperparameter yang lebih sedikit lebih disukai untuk gambar wajah bagian depan. • CNN yang kompleks dengan lebih banyak jumlah lapisan dan parameter lebih disukai untuk wajah spontan. • CNN moderat dengan jumlah lapisan dan parameter yang seimbang dengan pra-pemrosesan yang tepat lebih disukai untuk gambar yang terhalang.
5.	(Santhoshkumar & Geetha, 2019) / Deep Learning Approach for Emotion Recognition from Human Body Movements with Feedforward Deep Convolution Neural Networks	dataset emosi (University of .YORK) dengan 15 jenis emosi dan dataset corpus GEMEP dengan 5 emosi	feedforward deep convolution neural network (FDCNN)	Kinerja model ini lebih baik daripada model dasar
6.	(Lebedev et al., 2020) /			Usulan untuk pemantauan jarak jauh keadaan mental,

No	Peneliti / tahun / Judul	Dataset	Metode	Hasil/Keterbatasan
	Remote Recognition of Human Emotions Using Deep Machine Learning of Artificial Neural Networks			metode banyak digunakan berdasarkan pengenalan aktivitas wajah, yang secara refleks dikaitkan dengan tampilan keadaan psiko-emosional
7.	(Handrich et al., 2020) / Simultaneous Prediction of Valence/Arousal and Emotions on AffectNet, Aff-Wild and AFEW-VA	AffectNet, Aff-Wild dan AFEW-VA	arsitektur YOLO	Sampel pelatihan AffectNet sudah menunjukkan hasil pose yang lebih dari berbagai variasi pencahayaan, selanjutnya mendapatkan variasi dengan augmentasi data, yang mengarah ke kesalahan prediksi yang sedikit berkurang. Lalu menggabungkan sampel pelatihan dari dua atau tiga database, saat pengujian pada Aff-Wild dan AFEW-VA hasilnya meningkat
8.	(Hassouneh et al., 2020) / Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods		Berdasarkan penanda wajah dan sinyal electroencephalograph (EEG) menggunakan pengklasifikasi convolutional neural network (CNN) dan long short-term memory (LSTM)	Hasil mencapai tingkat pengenalan maksimum 99,81% menggunakan CNN untuk deteksi emosi menggunakan landmark wajah. Namun, tingkat pengenalan maksimum yang dicapai dengan menggunakan pengklasifikasi LSTM adalah 87,25% untuk deteksi emosi menggunakan sinyal EEG
9.	(Ivanova & Borzunov, 2020) / Optimization of machine learning algorithm of emotion recognition in terms of human facial expressions	database ORL untuk wajah yang terdiri dari 400 gambar berukuran 112 x 92. Masing-masing berisi 40 orang, masing-masing 10	Metode Viola Jones untuk akurasi deteksi objek, klasifikasi emosi menggunakan GoogleNet	Model yang dibangun dari CNN disetel: parameter lapisan jaringan ditetapkan untuk memastikan fungsi kompleks perangkat lunak yang paling efisien. Proporsi pengenalan yang benar untuk sampel uji publik adalah 69%, untuk sampel uji privat 59%. Akurasi pengenalan indikator juga ditentukan untuk setiap kelompok emosi. Akurasi pengenalan tertinggi dicatat

No	Peneliti / tahun / Judul	Dataset	Metode	Hasil/Keterbatasan
		gambar, Dataset pelatihan : FER2013		pada gambar dengan emosi "kebahagiaan" (88%), terendah pada gambar dengan emosi "kesedihan" (57%).
10.	(Gupta et al., 2020) / Deep self-attention network for facial emotion recognition	Data gambar statis	Menggabungkan CNN, ResNet dan mekanisme attention	Model yang diusulkan telah menunjukkan hasil yang lebih baik pada dataset FER dan dapat digunakan dalam aplikasi waktu nyata
11.	(Chen et al., 2019) / Auto-creation of Effective Neural Network Architecture by Evolutionary Algorithm and ResNet for Image Classification	CIFAR-10	EANN dan EANN-Net	EANN-Net mencapai kesalahan e=8.86% sementara ResNet-44 mendapatkan kesalahan e=11.64%.Ini berarti bahwa EANN-Net dapat mengklasifikasikan data uji dengan akurasi 2,78 poin persentase lebih baik daripada ResNet-44 yang digunakan sebagai jaringan referensi. EANN ynag diusulkan yaitu Arsitektur EANN-Net yang efektif dengan jumlah evaluasi fungsi yang terbatas. EANN-Net yang dihasilkan dengan properti heterogen dapat dilatih pada kumpulan data dengan iterasi pelatihan yang lebih sedikit daripada arsitektur yang sebanding
12.	(Li & Lima, 2021) /Facial expression recognition via ResNet-50	Foto dari fotografer berpengalaman (kamera digital Canon) setiap subjek sepuluh kali untuk 20 subjek total 700 gambar.	ResNet-50	memiliki akurasi yang baik dan efek pengenalan yang baik dalam hal akurasi pengenalan rata-rata.

Bab 3

Metodologi

3.1 Motivasi

Ekspresi wajah berperan penting dalam komunikasi antar sesama manusia di dalam kehidupan sehari-hari. Ekspresi wajah adalah salah satu bentuk komunikasi nonverbal yang dapat menyampaikan bermacam-macam keadaan emosi seseorang kepada orang yang mengamati. Ekspresi dapat dilihat dengan memperhatikan komponen pada wajah seperti mata, alis, dahi, dan mulut. Menurut para ahli, emosi dasar manusia yang telah dikenal sejak dulu menjadi aspek penting dari perilaku manusia.

Perkembangan teknologi yang pesat sekarang ini membawa dampak yang signifikan untuk berbagai jenis teknologi dan penemuan baru, sebagai solusi untuk mengatasi berbagai permasalahan kehidupan. Hal inilah yang mendorong manusia membutuhkan sebuah sistem yang dapat digunakan untuk membantu menyelesaikan masalah dengan cepat dan efisien. Beberapa metode telah ditemukan dan dihasilkan untuk melakukan pengenalan ekspresi wajah digunakan untuk kebutuhan pendidikan, sistem otomatis HMI (Human Machine Interface), kesehatan dan keamanan. Pengenalan ekspresi wajah digunakan dalam aplikasi real-time seperti pengawasan kondisi pengemudi/pengendara, medis, interaksi robotika, bagian forensik, mendeteksi penipuan serta diterapkan di banyak sektor industri dan pemerintah. Pengenalan ekspresi wajah dapat mengidentifikasi penyerang potensial berdasarkan analisis emosinya, analisis efektivitas promosi, menentukan efektivitas metodologi untuk menjual barang di toko, dan pengujian reaksi skala besar terhadap produk (semantik, wajah). Pada bidang teknologi pengenalan ekspresi wajah dapat digunakan sebagai pembelajaran bagi robot dalam mengenali ekspresi wajah manusia.

Pengenalan ekspresi wajah (FER) saat ini menjadi topik penelitian yang sangat menarik di bidang visi komputer, pengenalan pola, kecerdasan buatan, dan telah menarik perhatian luas karena aplikasi potensialnya untuk interaksi manusia-komputer alami (HCI), analisis emosi manusia, video interaktif, pengindeksan dan pengambilan gambar dan lainnya. Mengidentifikasi emosi manusia adalah salah satu aplikasi yang paling menarik dan berguna dalam visi komputer, dapat digunakan untuk mengidentifikasi keadaan mental manusia, mendeteksi kebohongan, memahami suasana hati, merawat pasien gangguan jiwa, hingga membuat sistem pencarian lebih baik, konseling dll. Mendeteksi ekspresi dilakukan oleh manusia dalam kehidupan sehari-hari tanpa masalah dan dengan mudah, tetapi sulit untuk membuat mesin memahami hal ini.

Beberapa penelitian telah dilakukan menggunakan komputer khususnya dalam pengenalan ekspresi wajah yang sudah diteliti sejak lama. Pendekatan pengenalan ekspresi wajah dapat menggunakan pendekatan gambar ataupun video. Pada pendekatan video, ekspresi wajah dideteksi berdasarkan urutan bagian gambar. Dari berbagai penelitian banyak metode yang telah digunakan untuk melakukan proses pengenalan ekspresi wajah manusia, Pengenalan ekspresi wajah sangat berkembang pesat menggunakan metode-metode tertentu. Metode-metode tersebut adalah Naïve Bayes, Support Vector Machine (SVM), k-Nearest Neighbor (KNN), Convolutional Neural

Networks (CNN), dll. Metode-metode tersebut dapat diimplementasikan untuk pendeteksian ekspresi wajah manusia menggunakan pendekatan gambar.

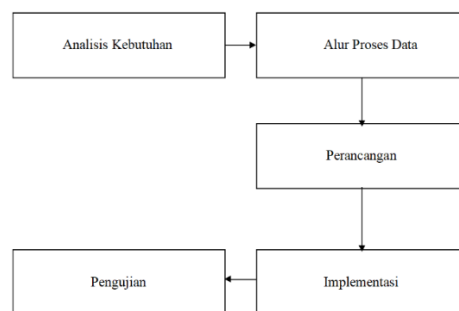
Berdasarkan tinjauan penelitian terdahulu yang sudah diuraikan diatas, menunjukkan bahwa pengenalan ekspresi wajah masih terus dilakukan dengan memanfaatkan berbagai macam teknologi dan juga metode baru agar menghasilkan keakuratan hasil pendeteksian. Sehingga hal ini memotivasi penulis untuk mengusulkan adanya pendekatan dengan pengembangan model Convolutional Neural Network arsitektur ResNet dan model turunannya yaitu EANN-Net untuk mendeteksi ekspresi wajah, sehingga dapat memperoleh pra-pemrosesan yang jauh lebih cepat dan mendapatkan hasil yang akurat.

3.2 Framework Riset

Kerangka penelitian yang terlihat pada Gambar 5 terdiri dari beberapa tahap yaitu :

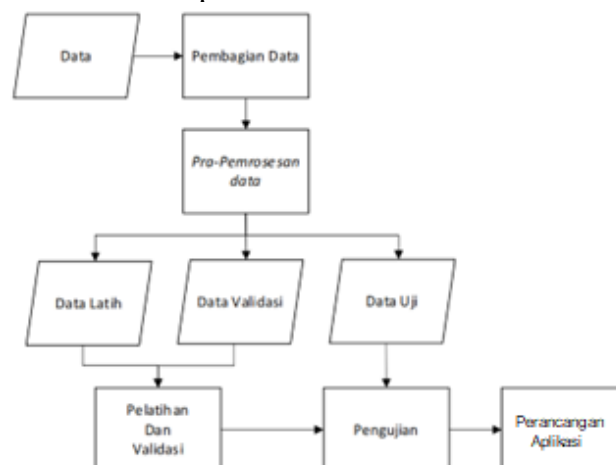
1. Tahap analisa kebutuhan dibagi menjadi kebutuhan fungsional dan non fungsional. Pada Tahap analisa kebutuhan fungsional adalah bagian penjelasan mengenai fitur yang dimasukkan ke dalam aplikasi yang dibuat. Fitur yang terdapat dalam aplikasi ini seperti fitur identifikasi ekspresi melalui pendekatan citra dengan cara mengambil foto galeri atau dari kamera *smartphone* atau *webcam* dan menampilkan hasil pengklasifikasian ekspresi wajah. Tahap analisa kebutuhan non fungsional adalah mempersiapkan perangkat keras dan perangkat lunak yang diperlukan yaitu spesifikasi laptop peneliti dan spesifikasi perangkat lunak bahasa pemrograman contohnya Python.
2. Tahap selanjutnya alur proses data pada Gambar 6 :
 - a. dimulai dari akuisisi data citra sebagai *input* yang akan diambil dari dataset yang sudah tersedia dan sering digunakan untuk penelitian contoh dataset FER 2013 atau dataset yang diambil secara mandiri untuk digabungkan dengan dataset yang sudah tersedia.
 - b. Kemudian data akan dibagi ke dalam beberapa penggunaan seperti data latih 80%, data validasi 10% dan data uji 10%.
 - c. Setelah itu dilakukan pra-pemrosesan untuk mempersiapkan data citra agar dapat digunakan dan sesuai dengan model arsitektur yang akan dibangun agar dapat meningkatkan kinerja model arsitektur tersebut. Pada pra-pemrosesan akan dilakukan proses pengubahan citra menjadi array dan melakukan augmentasi data. Proses augmentasi ini dilakukan untuk menangani masalah data yang tidak seimbang dan menghindari overfitting pada model yang hendak dibentuk.
 - d. Pembuatan model Convolutional Neural Network dalam struktur arsitektur ResNet dan EANN-Net. Penjelasan ada di subbab pendekatan.
 - e. Pelatihan model, proses ini berguna untuk mempelajari fitur-fitur dan kelas dari citra ekspresi wajah. Tujuan dari latihan ini adalah agar model dapat mengenali ciri citra menggunakan data latih dan data validasi. Data akan dilatih dengan menggunakan arsitektur yang telah dirancang sebelumnya.

- f. Proses pengujian model menggunakan data uji yang telah di pisahkan Sebelumnya. Proses pengujian dilakukan untuk menguji apakah model yang kita gunakan sudah mempelajari data yang sudah pada saat proses pelatihan model. Selain itu proses pengujian dilakukan untuk mengukur performa model dan akurasi dari model yang telah dibuat, dan juga sebagai simulasi penggunaan model pada dunia nyata. Dari hasil pengujian tersebut diperoleh nilai akurasi berdasarkan pengujian terhadap data uji. Pada pengujian model tersebut menghasilkan *confusion matrix* dengan bentuk tabel matriks yang menggambarkan kinerja model klasifikasi pada serangkaian data uji yang nilai sebenarnya diketahui. Ketepatan model dapat diukur menggunakan beberapa metode pengukuran *confusion matrix* diantaranya pengukuran *accuracy*, pengukuran *precision*, pengukuran *recall* dan pengukuran *F1-Score*. Pada dasarnya *confusion matrix* memberikan informasi perbandingan hasil klasifikasi yang dilakukan oleh model dengan klasifikasi sebenarnya.
3. Tahap perancangan aplikasi menggunakan rancangan *flowchart* untuk mengetahui alur fitur yang dimiliki aplikasi dan rancangan tampilan.
4. Tahap implementasi : pembuatan aplikasi ini menggunakan bahasa pemrograman yang dipilih dan juga menggunakan *library* tambahan untuk mengimplementasikan hasil model yang telah dibuat ke dalam aplikasi.
5. Tahap pengujian terhadap aplikasi harus dilakukan untuk mengetahui kinerja aplikasi dan mengevaluasi arsitektur yang digunakan dalam mengidentifikasi ekspresi wajah.



Gambar 6. Kerangka Dasar Penelitian

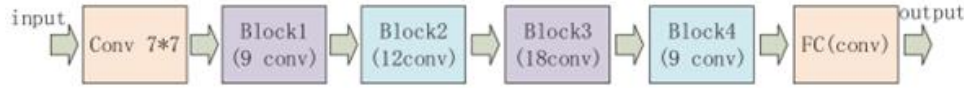
Pada alur proses data dilakukan tahap berikut :



Gambar 7. Alur Proses Data

3.3 Pendekatan

Didalam pembuatan model arsitektur CNN, mereferensi dari pendekatan arsitektur ResNet sebagai berikut :



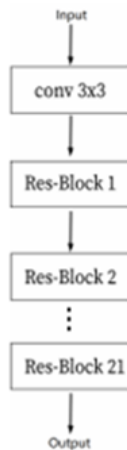
Gambar 8. ResNet-50

Gambar 7 merupakan rancangan dari arsitektur *ResNet-50*. Pada awalnya, ResNet-50 melakukan operasi konvolusi pada input, diikuti oleh 4 blok residual, dan akhirnya melakukan operasi koneksi penuh untuk mencapai tugas klasifikasi. Struktur jaringan ResNet-50 ditampilkan di Gambar 7, yang memiliki 50 operasi Conv2D. Metode ini memiliki akurasi yang baik dan efek pengenalan yang baik dalam hal akurasi pengenalan rata-rata. (Li & Lima, 2021)

Pendekatan arsitektur yang kedua adalah EANN-Net yang dihasilkan dari EANN, sebagai berikut : Algoritma 1 menjelaskan pseudocode dari EANN yang diusulkan. Hal ini terutama didasarkan pada satu set blok bangunan dasar, dan menghasilkan individu terbaik dalam populasi akhir dan arsitektur jaringan yang sesuai EANN-Net. Prosedur penting dari Algoritma 1 akan diperkenalkan secara rinci di bagian berikut.

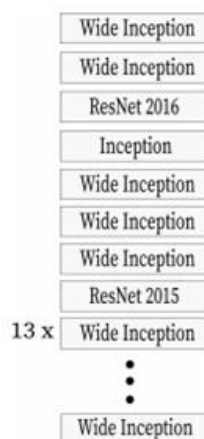
Tabel 2. Algoritma EANN

Algorithm 1 EANN
<p>Input: a set of basic building blocks, population size N, maximal generation number T.</p> <p>Output: the best individual x^{best} and its corresponding network architecture.</p> <ol style="list-style-type: none"> 1: Initialize population $P_0 = \{x^1, x^2, \dots, x^N\}$ by using the basic building blocks; 2: $Q_0 \leftarrow \emptyset$; 3: for $t \leftarrow 0$ to T do 4: Determine the best individual x^{best} in P_t; 5: if $\text{rand} < 0.5$ then 6: Produce new individual x' by mutating x^{best}; 7: Determine the best one between x^{best} and x', and copy it into Q_{t+1}; 8: else 9: Copy x^{best} into Q_{t+1}; 10: end if 11: Determine the remaining individuals in Q_{t+1} by applying roulette wheel selection to P_t; 12: Apply crossover and mutation to Q_{t+1}; 13: $P_{t+1} \leftarrow Q_{t+1}$; 14: end for 15: Determine the best individual x^{best} in P_t; 16: Obtain the network architecture (denoted as EANN-Net) by decoding x^{best}.



Gambar 9. Struktur Jaringan EANN

Semua jaringan berasal dari struktur ResNet-44 (He & Sun, 2015). Ini terdiri dari 21 blok sisa (lihat Gambar 8), didahului oleh lapisan konvolusi awal dan diikuti oleh lapisan yang terhubung penuh di akhir untuk klasifikasi.



Gambar 10. Struktur EANN-Net yang dihasilkan oleh EANN

Karena kinerja tinggi dari prosedur pengujian ini, Gambar 9 merinci struktur EANN-Net yang dihasilkan oleh EANN. Setelah lapisan konvolusi awal, dua blok residu pertama diikuti dua kali oleh struktur Inception yang diadaptasi untuk pekerjaan ini dengan jalan pintas dan lebih dalam 5×5 filter. Ini diikuti oleh lapisan ResNet-2016 dengan pra-aktivasi Lapisan konvolusi kedua di blok ini tidak dinormalisasi batch, juga tidak memiliki fungsi aktivasi oleh lapisan ReLU, karena pra-aktivasi. Sinyal yang ditambahkan di akhir blok diteruskan ke lapisan Inception berikut dengan jalan pintas tanpa nonlinier dan normalisasi. Di blok ini, kedalaman 5×5 filter tidak meningkat. Ini diikuti oleh tiga blok Inception, dengan 5×5 filter dengan kedalaman yang ditingkatkan dan blok residu (He & Sun, 2015) tanpa pra-aktivasi dan dengan lapisan ReLU, yang mengisi jalur pintas jaringan yang sebelumnya tidak terhalang. Kemudian diikuti oleh 13 blok Wide-Inception dan struktur keluaran diambil dari (He et al., 2016). Setelah satu blok ResNet tradisional, tidak ada lagi kerusakan di jalur pintasan jaringan. Dari blok kedelapan, jaringan dapat memanfaatkan arsitektur yang disajikan pada (He et al., 2016). Berfokus pada struktur EANN-Net yang ditunjukkan pada Gambar 9, penggunaan besar-

besaran blok Wide-Inception membuatnya tampak seolah-olah arsitektur ini hanyalah alasan untuk lebih banyak parameter hiper yang dapat dioptimalkan untuk akurasi klasifikasi yang lebih tinggi. Seperti yang kita ketahui, arsitektur Inception [1] dan [3] memungkinkan filter yang lebih kompleks dan lebih banyak parameter hiper untuk dioptimalkan per blok daripada ResNet-44 dari (He & Sun, 2015) yang digunakan sebagai jaringan referensi. Namun, dalam percobaan tambahan, dapat diamati bahwa kesalahan pengujian e dicapai oleh jaring homogen dari arsitektur Inception secara signifikan lebih tinggi daripada ResNet-44. Dengan demikian, menjadi jelas bahwa jumlah hyperparameter yang lebih besar tidak dapat dianggap bertanggung jawab atas peningkatan kinerja saja. Oleh karena itu, alasan untuk hasilnya harus terletak pada kombinasi struktur blok yang digunakan. (Chen et al., 2019)

3.4 Rencana Kerja

Rencana kerja penyusunan disertasi sebagai berikut :

Tabel 3. Rencana Kerja

Kegiatan	Bulan Ke-																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Studi Pustaka																				
Penyusunan Proposal																				
Pengajuan Proposal																				
Pengerjaan Penelitian																				
Publikasi Ilmiah/Seminar																				
Perbaikan dan Pengembangan																				
Penyusunan Disertasi																				
Sidang Disertasi																				

Bibliografi

- Chen, Z., Zhou, Y., & Huang, Z. (2019). Auto-creation of effective neural network architecture by evolutionary algorithm and ResNet for image classification. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics, 2019-Octob*, 3895–3900. <https://doi.org/10.1109/SMC.2019.8914267>
- De, A., Saha, A., & Pal, M. C. (2015). A human facial expression recognition model based on eigen face approach. *Procedia Computer Science*, 45(C), 282–289. <https://doi.org/10.1016/j.procs.2015.03.142>
- Gupta, A., Arunachalam, S., & Balakrishnan, R. (2020). Deep self-attention network for facial emotion recognition Deep network for facial emotion recognition. *Procedia Computer Science*, 171(2019), 1527–1534. <https://doi.org/10.1016/j.procs.2020.04.163>
- Handrich, S., Dinges, L., Al-Hamadi, A., Werner, P., & Aghbari, Z. Al. (2020). Simultaneous Prediction of Valence/Arousal and Emotions on AffectNet, Aff-Wild and AFEW-VA. *Procedia Computer Science*, 170(2019), 634–641. <https://doi.org/10.1016/j.procs.2020.03.134>
- Hassouneh, A., Mutawa, A. M., & Murugappan, M. (2020). Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods. *Informatics in Medicine Unlocked*, 20, 100372. <https://doi.org/10.1016/j.imu.2020.100372>
- He, K., & Sun, J. (2015). *Deep Residual Learning fo r Image Recognition*. 1–9. <https://arxiv.org/pdf/1512.03385.pdf>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Identity mappings in deep residual networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9908 LNCS, 630–645. https://doi.org/10.1007/978-3-319-46493-0_38
- Ivanova, E., & Borzunov, G. (2020). Optimization of machine learning algorithm of emotion

- recognition in terms of human facial expressions. *Procedia Computer Science*, 169(2019), 244–248. <https://doi.org/10.1016/j.procs.2020.02.143>
- Lebedev, G., Zhovnerchuk, E., Zhovnerchuk, I., & Moskoventko, A. (2020). Remote recognition of human emotions using deep machine learning of artificial neural networks. *Procedia Computer Science*, 176, 1517–1522. <https://doi.org/10.1016/j.procs.2020.09.162>
- Li, B., & Lima, D. (2021). Facial expression recognition via ResNet-50. *International Journal of Cognitive Computing in Engineering*, 2(January), 57–64. <https://doi.org/10.1016/j.ijcce.2021.02.002>
- Paul Ekman, W. V. F. (2003). *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues, Volume 10*. books.google.com. [https://books.google.co.id/books?hl=en&lr=&id=TukNoJDgMTUC&oi=fnd&pg=PR3&dq=Ekman,+P.,+%26+Friesen,+W.+V.+\(2003\).+Unmasking+the+Face.+A+Guide+to+Recognizing+Emotions+from+Facial+Clues.+Los+Altos,+CA:+Malor+Books.&ots=GWIpem0Z9b&sig=bWjuBLc6yDvOubYwRxRQx_z](https://books.google.co.id/books?hl=en&lr=&id=TukNoJDgMTUC&oi=fnd&pg=PR3&dq=Ekman,+P.,+%26+Friesen,+W.+V.+(2003).+Unmasking+the+Face.+A+Guide+to+Recognizing+Emotions+from+Facial+Clues.+Los+Altos,+CA:+Malor+Books.&ots=GWIpem0Z9b&sig=bWjuBLc6yDvOubYwRxRQx_z)
- Revina, I. M., & Emmanuel, W. R. S. (2021). A Survey on Human Face Expression Recognition Techniques. *Journal of King Saud University - Computer and Information Sciences*, 33(6), 619–628. <https://doi.org/10.1016/j.jksuci.2018.09.002>
- Santhoshkumar, R., & Geetha, M. K. (2019). Deep Learning Approach for Emotion Recognition from Human Deep Learning Approach for Emotion Recognition from Human Body Movements with Feedforward Deep Convolution Neural Body Movements with Feedforward Deep Convolution Neural Networks Networks. *Procedia Computer Science*, 152, 158–165. <https://doi.org/10.1016/j.procs.2019.05.038>
- Suma, K., Lakshminarayana, M., & Tech, M. (2018). Survey on Face Expression Recognition using CNN. *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, 5, 102–106.
- Zhao, X., & Zhang, S. (2016). A review on facial expression recognition: Feature extraction and classification. *IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India)*, 33(5), 505–517. <https://doi.org/10.1080/02564602.2015.1117403>

<https://github.com/xw-hu/Reading-List/blob/master/README.md>