

Pengembangan Model Deteksi Autism Spectrum Disorder Berbasis FaceNet, VGGFace2, dan InsightFace

Marsha Falen Fransisca, Lukman Sunardi, Harma Oktafia LW, Budi Santoso, *

Fakultas Ilmu Teknik, Program Studi Informatika, Universitas Bina Insan, Lubuklinggau, Indonesia

e-mail : ¹marsyafalen306@gmail.com, ²lukmanmmci@gmail.com, ³harmaoktafialingga@gmail.com, ⁴budisantoso@univbinainsan.ac.id

Email Penulis Corresponding: marsyafalen306@gmail.com

Abstrak- Autism Spectrum Disorder (ASD) merupakan gangguan perkembangan saraf yang memengaruhi kemampuan komunikasi, interaksi sosial, serta perilaku individu. Proses diagnosis ASD secara konvensional masih bergantung pada observasi klinis yang bersifat subjektif dan membutuhkan waktu yang relatif lama. Oleh karena itu, diperlukan pendekatan berbasis kecerdasan buatan untuk mendukung deteksi dini ASD secara lebih objektif dan efisien. Penelitian ini mengusulkan model deteksi ASD berbasis analisis citra wajah menggunakan pendekatan deep learning dengan memanfaatkan arsitektur FaceNet, VGGFace2, dan InsightFace sebagai metode ekstraksi fitur wajah. Dataset yang digunakan terdiri dari 3.620 citra wajah yang terbagi ke dalam dua kelas, yaitu ASD dan non-ASD. Proses penelitian meliputi tahapan preprocessing, ekstraksi fitur, pelatihan model, serta evaluasi menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa ketiga model mampu memberikan performa klasifikasi yang baik, dengan FaceNet memperoleh akurasi tertinggi sebesar 98%, diikuti oleh InsightFace sebesar 96%, dan VGGFace2 sebesar 95%. Temuan ini menunjukkan bahwa model berbasis face embedding memiliki kemampuan yang lebih optimal dalam mengekstraksi fitur wajah untuk mendukung deteksi ASD.

Kata Kunci: Deep Learning; Autism Spectrum Disorder; FaceNet; VGGFace2; InsightFace; Citra Wajah

Abstract- Autism Spectrum Disorder (ASD) is a neurodevelopmental condition that affects communication, social interaction, and behavior. Conventional ASD diagnosis relies on clinical observation, which is time-consuming and subjective. Therefore, an automated approach using artificial intelligence is required to support early detection. This study proposes an ASD detection model based on facial image analysis using deep learning approaches, namely FaceNet, VGGFace2, and InsightFace as facial feature extraction methods. The dataset consists of 3,620 facial images categorized into ASD and non-ASD classes. The research process includes preprocessing, feature extraction, model training, and evaluation using accuracy, precision, recall, and F1-score metrics. The results indicate that all models achieved good classification performance, with FaceNet achieving the highest accuracy of 98%, followed by InsightFace with 96%, and VGGFace2 with 95%. These findings demonstrate that face embedding-based models provide superior feature extraction capabilities for ASD detection.

Keywords: Deep Learning; Autism Spectrum Disorder; FaceNet; VGGFace2; InsightFace; Facial Images

1. PENDAHULUAN

Dalam beberapa dekade terakhir, kemajuan teknologi kecerdasan buatan memberikan dampak signifikan pada berbagai sektor, khususnya di bidang computer vision. Teknologi ini memungkinkan komputer untuk secara otomatis memahami, menganalisis, dan menginterpretasikan data visual dari citra digital. Berbagai penerapan computer vision kini banyak dimanfaatkan dalam bidang kesehatan, keamanan, industri, hingga pendidikan. Salah satu pendekatan yang paling berkembang dalam area ini adalah deep learning, yang dapat mengekstraksi fitur secara otomatis tanpa memerlukan perancangan fitur manual [1], [2].

Deep learning adalah bagian dari machine learning yang memakai jaringan saraf tiruan berlapis untuk mempelajari representasi data pada beragam tingkat abstraksi [3]. Dengan mengaplikasikan algoritma backpropagation, model deep learning dapat menyesuaikan bobot secara optimal sehingga menghasilkan kinerja yang lebih baik pada berbagai tugas klasifikasi serta pengenalan pola. Salah satu arsitektur paling terkenal dalam computer vision adalah Convolutional Neural Network (CNN), yang dapat menangkap pola spasial pada gambar seperti tekstur, tepi, dan bentuk objek [4], [5], [6].

Di sektor kesehatan, penggunaan deep learning untuk menganalisis citra wajah telah terbukti memiliki potensi yang signifikan dalam membantu proses diagnosis berbagai kondisi medis, termasuk Autism Spectrum Disorder (ASD). ASD adalah gangguan perkembangan saraf yang ditandai dengan kesulitan dalam komunikasi, interaksi sosial, serta perilaku yang terbatas dan berulang. Gangguan ini biasanya muncul pada usia dini dan dapat secara signifikan memengaruhi kualitas hidup individu. Karena itu, deteksi ASD secara awal menjadi sangat penting untuk memberikan intervensi yang tepat serta meningkatkan perkembangan kognitif dan sosial pada individu yang terdiagnosis [7], [8].

Namun demikian, proses diagnosis ASD hingga saat ini masih didominasi oleh metode konvensional yang bergantung pada observasi klinis oleh tenaga ahli. Metode ini memiliki beberapa keterbatasan, antara lain membutuhkan waktu yang relatif lama, bersifat subjektif, serta bergantung pada pengalaman dan keahlian profesional. Kondisi ini menjadi tantangan tersendiri, terutama di wilayah yang memiliki keterbatasan tenaga medis. Oleh karena itu, diperlukan suatu sistem otomatis berbasis teknologi yang dapat membantu proses deteksi ASD secara lebih cepat, objektif, dan efisien [9], [10]. Seiring dengan perkembangan penelitian, beberapa studi menunjukkan bahwa individu dengan ASD memiliki karakteristik morfologi wajah tertentu yang dapat digunakan sebagai indikator dalam proses klasifikasi. Perbedaan tersebut meliputi struktur wajah, jarak antar fitur wajah, serta proporsi bagian tertentu pada wajah. Dengan memanfaatkan

teknologi pengolahan citra dan deep learning, fitur-fitur tersebut dapat diekstraksi dan digunakan sebagai dasar dalam membedakan individu ASD dan non-ASD [11], [12].

Berbagai studi terdahulu telah mengembangkan model deteksi ASD berbasis gambar wajah dengan memanfaatkan arsitektur CNN populer seperti VGG16, ResNet, MobileNet, dan AlexNet. Temuan mereka menunjukkan bahwa model deep learning dapat mencapai akurasi yang tinggi, bahkan melebihi 90% pada beberapa dataset tertentu. Hal ini menandakan bahwa pendekatan berbasis gambar wajah memiliki potensi signifikan sebagai alat bantu dalam skrining awal ASD [13], [14].

Meskipun demikian, sebagian besar penelitian masih berfokus pada penggunaan arsitektur CNN umum, sementara model khusus pengenalan wajah (face recognition) yang memiliki kemampuan representasi fitur lebih kuat belum banyak dieksplorasi dalam konteks deteksi ASD. Model-model seperti FaceNet, VGGFace2, dan InsightFace dikenal memiliki kemampuan yang sangat baik dalam mengekstraksi fitur wajah yang diskriminatif dan stabil terhadap berbagai variasi seperti pose, pencahayaan, dan ekspresi [15], [16], [17].

FaceNet adalah salah satu model pembelajaran mendalam yang dibuat untuk menghasilkan gambaran wajah dalam bentuk embedding dengan dimensi yang lebih kecil, dengan menggunakan metode triplet loss. Model ini dapat memaksimalkan jarak antara kelas yang berbeda dan mendekatkan jarak antara kelas yang sama, sehingga menghasilkan fitur yang sangat baik untuk tugas klasifikasi dan mengenali wajah. Di sisi lain, VGGFace2 adalah dataset dan model yang sudah dilatih, yang menyediakan banyak gambar wajah beragam, sehingga dapat meningkatkan kemampuan model untuk beradaptasi dengan berbagai kondisi visual [16].

Selain itu, InsightFace dengan pendekatan ArcFace menawarkan peningkatan performa melalui optimasi berbasis margin sudut (angular margin loss), yang mampu meningkatkan separasi antar kelas dalam ruang fitur. Pendekatan ini terbukti mampu menghasilkan performa state-of-the-art dalam berbagai tugas pengenalan wajah [17]. Ketiga model tersebut memiliki keunggulan masing-masing dalam menghasilkan representasi fitur wajah yang berkualitas tinggi.

Namun, penelitian yang secara khusus membandingkan performa ketiga model tersebut dalam konteks deteksi ASD masih sangat terbatas. Padahal, dengan memanfaatkan model khusus pengenalan wajah, diharapkan dapat diperoleh hasil klasifikasi yang lebih akurat dibandingkan pendekatan CNN konvensional. Oleh karena itu, penelitian ini berfokus pada pengembangan model deteksi ASD berbasis citra wajah dengan menggunakan FaceNet, VGGFace2, dan InsightFace sebagai metode ekstraksi fitur.

Penelitian ini bertujuan untuk mengembangkan model yang mampu membedakan individu dengan Autism Spectrum Disorder (ASD) dan individu non-ASD secara otomatis berdasarkan citra wajah. Selain itu, penelitian ini juga bertujuan untuk membandingkan hasil ketiga model dalam aspek akurasi, presisi, recall, dan skor F1, agar dapat menentukan model yang paling efektif dalam mendeteksi ASD. Dengan adanya penelitian ini, diharapkan bisa membantu mengembangkan sistem deteksi dini ASD yang menggunakan kecerdasan buatan, sehingga lebih objektif dan efisien.

Dengan demikian, penelitian ini tidak hanya berfokus pada peningkatan performa klasifikasi, tetapi juga berkontribusi dalam menjembatani kesenjangan antara teknologi face recognition dan aplikasi kesehatan berbasis computer vision, khususnya dalam deteksi gangguan perkembangan seperti ASD.

2. METODOLOGI PENELITIAN

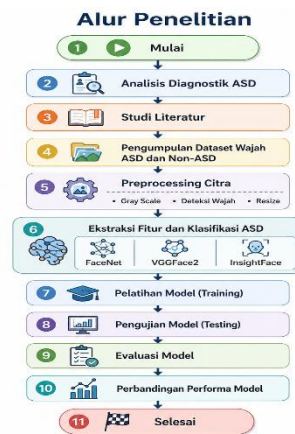
2.1 Desain dan Tahapan Penelitian

Penelitian ini mengaplikasikan pendekatan kuantitatif dengan metode eksperimental untuk merancang model klasifikasi berbasis deep learning yang mampu mengidentifikasi Autism Spectrum Disorder (ASD) melalui analisis citra wajah. Metode ini dipilih karena fokus penelitian terletak pada pengolahan data numerik berupa gambar digital yang dievaluasi secara objektif menggunakan algoritma pembelajaran mesin.

Pada bidang computer vision, proses klasifikasi gambar biasanya melibatkan tahapan-tahapan utama seperti pengumpulan data, praproses citra, ekstraksi fitur, pelatihan model, serta penilaian kinerja [1]. Secara umum, alur penelitian meliputi analisis karakteristik ASD dan studi literatur, pengumpulan dataset citra wajah ASD dan non-ASD, prapemrosesan citra, ekstraksi fitur menggunakan model FaceNet, VGGFace2, dan InsightFace, pelatihan model klasifikasi, evaluasi performa model, serta perbandingan hasil model. Alur penelitian tersebut ditampilkan pada Gambar 1.

Berdasarkan Gambar 1, penelitian dimulai dari analisis diagnostik ASD untuk memahami karakteristik Autism Spectrum Disorder sebagai dasar perancangan sistem. Tahap berikutnya adalah studi literatur untuk meninjau penelitian terdahulu, teori pendukung, serta metode yang relevan dengan deteksi ASD berbasis citra wajah. Setelah itu dilakukan pengumpulan dataset wajah ASD dan non-ASD, kemudian data diproses melalui tahap preprocessing citra yang meliputi grayscale conversion, deteksi wajah, dan resize. Citra yang telah diproses selanjutnya digunakan pada tahap ekstraksi fitur dan klasifikasi ASD menggunakan FaceNet, VGGFace2, dan InsightFace. Setelah proses ekstraksi fitur dilakukan, model dilatih menggunakan data latih dan diuji menggunakan data uji. Hasil pengujian kemudian dievaluasi menggunakan metrik accuracy, precision, recall, dan F1-score. Tahap akhir penelitian adalah membandingkan performa ketiga model untuk menentukan model terbaik dalam mendeteksi Autism Spectrum Disorder berbasis citra wajah.

Tahapan ini bertujuan untuk menghasilkan model yang mampu melakukan klasifikasi secara akurat serta memiliki kemampuan generalisasi yang baik terhadap data baru.



Gambar 1. Alur Penelitian Deteksi ASD Berbasis Citra Wajah

2.2 Dataset Penelitian

Dataset terdiri dari 3.620 citra wajah anak yang terbagi ke dalam dua kelas, yaitu kelas ASD (Autistic) sebanyak 1.887 citra dan kelas Non-ASD sebanyak 1.733 citra.



Kelas ASD

Kelas Non-ASD

Gambar 2. Citra Wajah Anak

Berdasarkan Gambar 2 ditampilkan contoh citra wajah anak pada kelas ASD dan non-ASD yang digunakan dalam penelitian. Dataset tersebut mencakup variasi pencahayaan, pose, ekspresi wajah, dan resolusi citra sehingga mampu merepresentasikan kondisi nyata saat proses pengambilan gambar. Keanekaragaman data ini penting untuk meningkatkan kemampuan generalisasi model dalam mengenali pola wajah secara lebih optimal pada berbagai kondisi [11].

Data kemudian dibagi menjadi data latih dan data uji dengan komposisi 80% untuk data latih dan 20% untuk data uji. Pembagian ini bertujuan untuk melatih model serta menguji performa model pada data yang belum pernah dilihat sebelumnya.

2.3 Prapemrosesan Citra

Tahapan prapemrosesan dilakukan untuk meningkatkan kualitas citra serta memastikan kesesuaian data dengan kebutuhan model *deep learning*. Proses ini sangat penting karena kualitas input sangat memengaruhi hasil ekstraksi fitur dan performa model [18].

Langkah-langkah prapemrosesan yang dilakukan meliputi grayscale conversion, face detection dan cropping, resizing citra, serta normalisasi nilai piksel.

Grayscale conversion dilakukan untuk menyederhanakan informasi visual dan mengurangi kompleksitas data. Selanjutnya dilakukan face detection dan cropping untuk memisahkan area wajah dari latar belakang agar model hanya berfokus pada fitur wajah sebagai objek utama. Setelah itu, seluruh citra diubah ke ukuran yang seragam melalui proses resizing agar sesuai dengan kebutuhan input model. Tahap terakhir adalah normalisasi nilai piksel untuk meningkatkan stabilitas pelatihan model dan mempercepat proses konvergensi.

Tahapan ini bertujuan untuk mengurangi noise, meningkatkan konsistensi data, serta mengoptimalkan proses pembelajaran model *deep learning* [18].

2.4 Ekstraksi Fitur Wajah

Ekstraksi fitur merupakan tahap utama dalam penelitian ini, di mana citra wajah diubah menjadi representasi numerik (feature embedding) yang dapat diproses oleh model klasifikasi. Penelitian ini menggunakan tiga model deep learning berbasis face recognition, yaitu FaceNet, VGGFace2, dan InsightFace.

FaceNet membuat representasi wajah dengan menggunakan metode triplet loss yang dapat memaksimalkan jarak antar wajah yang berbeda dan mendekatkan jarak antar wajah yang sama, sehingga menghasilkan fitur yang mampu membedakan satu wajah dengan wajah lainnya [15]. VGGFace2 merupakan model pralatih yang dirancang untuk menangani variasi pose, usia, dan pencahayaan sehingga mampu menghasilkan fitur dengan kemampuan generalisasi tinggi [16]. Sementara itu, InsightFace menggunakan pendekatan ArcFace yang mengoptimalkan angular margin untuk meningkatkan separasi antar kelas dalam ruang fitur sehingga menghasilkan performa tinggi pada pengenalan wajah [17]. Ketiga model tersebut digunakan sebagai feature extractor untuk membandingkan kemampuan masing-masing dalam mendeteksi ASD berdasarkan citra wajah.

2.5 Pelatihan Model

Setelah proses ekstraksi fitur dilakukan, embedding yang dihasilkan digunakan sebagai input untuk proses klasifikasi. Model dilatih untuk mengenali pola perbedaan antara wajah individu ASD dan non-ASD berdasarkan fitur yang telah diekstraksi.

Proses pelatihan dilakukan secara terpisah untuk setiap model agar dapat dilakukan perbandingan performa secara objektif. Model dilatih hingga mencapai kondisi good fitting, yaitu ketika model mampu mengenali pola data dengan baik tanpa mengalami overfitting maupun underfitting.

Pendekatan ini memungkinkan evaluasi yang lebih akurat terhadap kemampuan masing-masing model dalam melakukan klasifikasi.

2.6 Evaluasi Kinerja Model

Evaluasi model dilakukan menggunakan metode *confusion matrix*, yang merupakan salah satu teknik evaluasi paling umum dalam *supervised learning* [19].

Berdasarkan confusion matrix, metrik evaluasi yang digunakan meliputi accuracy untuk mengukur tingkat ketepatan model dalam klasifikasi, precision untuk mengukur ketepatan prediksi positif, recall untuk mengukur kemampuan model mendeteksi data positif, serta F1-score sebagai rata-rata harmonik antara precision dan recall.

Rumus accuracy, precision, recall, dan F1-score dinyatakan sebagai berikut :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

di mana TP (*True Positive*), TN (*True Negative*), FP (*False Positive*), dan FN (*False Negative*) merupakan komponen utama dalam evaluasi model [20].

Evaluasi ini bertujuan untuk mengetahui performa model secara menyeluruh serta mengidentifikasi kesalahan klasifikasi pada masing-masing kelas.

2.7 Perbandingan Model

Tahap akhir penelitian melibatkan perbandingan kinerja model FaceNet, VGGFace2, dan InsightFace, dengan menilai akurasi, precision, recall, serta F1-score yang diperoleh masing-masing.

Tujuan dari tahap ini adalah untuk menentukan model terbaik dalam mendeteksi ASD berdasarkan citra wajah. Model dengan performa terbaik diharapkan dapat digunakan sebagai dasar pengembangan sistem deteksi dini ASD berbasis kecerdasan buatan di bidang kesehatan.

3. HASIL DAN PEMBAHASAN

Pada bagian ini disajikan hasil eksperimen yang diperoleh dari proses pengembangan model deteksi Autism Spectrum Disorder (ASD) berbasis citra wajah, serta pembahasan terkait performa masing-masing model yang digunakan. Evaluasi dilakukan dengan membandingkan tiga pendekatan utama dalam ekstraksi fitur wajah, yaitu FaceNet, VGGFace2, dan InsightFace.

Penelitian ini bertujuan mengidentifikasi model dengan kemampuan paling baik dalam membedakan gambar wajah individu ASD dan non-ASD berdasarkan representasi fitur yang dihasilkan. Seluruh percobaan dilakukan dengan memakai dataset yang berisi 3.620 gambar wajah yang telah melalui proses pra-pemrosesan serta dibagi menjadi data pelatihan dan data pengujian. Evaluasi kinerja model dilakukan dengan metrik akurasi, presisi, recall, dan F1-score, serta dianalisis melalui confusion matrix untuk memeriksa distribusi kesalahan klasifikasi pada tiap model.

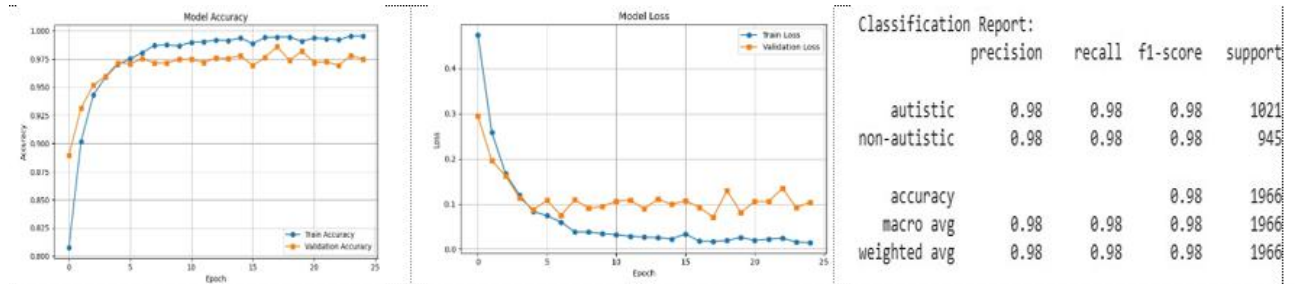
3.1 Hasil Pelatihan Model FaceNet

FaceNet merupakan model deep learning yang dirancang khusus untuk menghasilkan embedding wajah dalam ruang vektor berdimensi rendah. Model ini menggunakan pendekatan triplet loss yang memungkinkan pembelajaran jarak antar kelas secara optimal. Pada penelitian ini, FaceNet digunakan sebagai feature extractor untuk menghasilkan representasi fitur wajah dari setiap citra. Embedding yang dihasilkan kemudian digunakan dalam proses klasifikasi untuk membedakan antara kelas ASD dan non-ASD.

Hasil pelatihan menunjukkan bahwa model FaceNet mampu mencapai performa yang sangat tinggi. Berdasarkan hasil pengujian pada data uji, model ini memperoleh nilai akurasi sebesar 98%, yang merupakan nilai tertinggi dibandingkan model lainnya. Selain itu, nilai precision dan recall juga menunjukkan performa yang konsisten tinggi, yang mengindikasikan bahwa model mampu mengklasifikasikan data secara tepat serta memiliki kemampuan yang baik dalam mendeteksi kelas positif (ASD).

Tingginya performa FaceNet dapat dijelaskan oleh kemampuannya dalam menghasilkan embedding wajah yang sangat diskriminatif. Dengan memanfaatkan triplet loss, Model tersebut dapat memaksimalkan jarak antar kelas yang berbeda dan mengurangi jarak antar kelas yang sama. Hal ini membuat fitur yang dihasilkan lebih representatif dalam membedakan karakteristik wajah individu ASD dan non-ASD. Selain itu, FaceNet juga dikenal memiliki ketahanan terhadap variasi pencahayaan, pose, dan ekspresi wajah, sehingga mampu bekerja dengan baik pada dataset yang memiliki variasi kondisi citra.

Pada Gambar 3 terlihat bahwa model FaceNet mampu mencapai performa pelatihan yang stabil dengan tingkat akurasi yang tinggi dalam proses klasifikasi ASD dan non-ASD.



Gambar 3. Hasil Pelatihan FaceNet

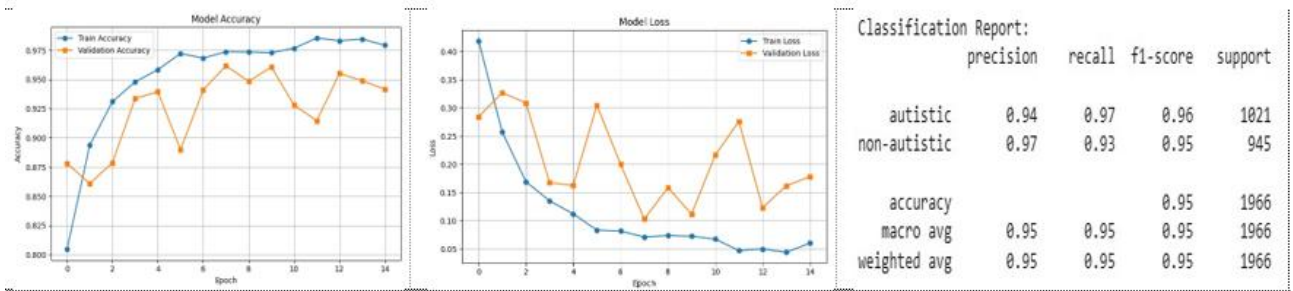
3.2 Hasil Pelatihan Model VGGFace2

Model kedua yang digunakan dalam penelitian ini adalah VGGFace2, yang merupakan model pralatih berbasis dataset wajah berskala besar. Model ini dirancang untuk menangani variasi pose, usia, dan kondisi pencahayaan yang beragam. Dalam penelitian ini, VGGFace2 digunakan sebagai feature extractor untuk menghasilkan representasi fitur wajah dari citra input. Hasil embedding kemudian digunakan dalam proses klasifikasi.

Berdasarkan hasil eksperimen, model VGGFace2 memperoleh nilai akurasi sebesar 95%. Nilai ini menunjukkan bahwa model mampu melakukan klasifikasi dengan baik, meskipun masih berada di bawah performa FaceNet. Nilai precision dan recall yang diperoleh juga menunjukkan bahwa model cukup stabil dalam melakukan klasifikasi, namun terdapat beberapa kesalahan prediksi terutama pada citra yang memiliki karakteristik wajah yang mirip antara kelas ASD dan non-ASD.

Performa VGGFace2 yang sedikit lebih rendah dibandingkan FaceNet dapat disebabkan oleh perbedaan pendekatan dalam pembelajaran fitur. VGGFace2 lebih bergantung pada arsitektur CNN konvensional yang menghasilkan fitur berdasarkan pola visual umum, sementara FaceNet secara khusus dirancang untuk menghasilkan embedding wajah yang optimal untuk tugas pengenalan wajah. Meskipun demikian, VGGFace2 tetap menunjukkan performa yang baik dan memiliki kemampuan generalisasi yang tinggi karena dilatih pada dataset yang besar dan beragam.

Pada Gambar 4 terlihat hasil pelatihan model VGGFace2 yang menunjukkan performa klasifikasi yang cukup baik dalam membedakan kelas ASD dan non-ASD.



Gambar 4. Hasil Pelatihan VGGFace2

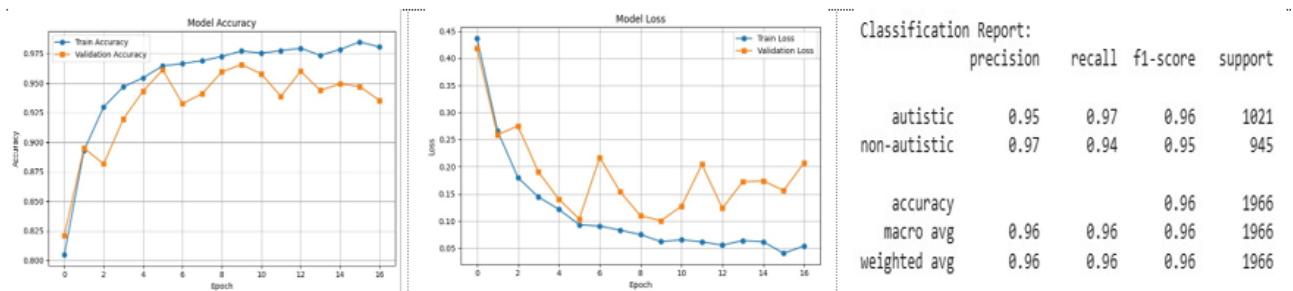
3.3 Hasil Pelatihan Model InsightFace

Model ketiga yang digunakan adalah InsightFace, yang merupakan framework face recognition modern dengan pendekatan ArcFace. Model ini menggunakan konsep angular margin loss untuk meningkatkan separasi antar kelas dalam ruang fitur. InsightFace digunakan sebagai feature extractor untuk menghasilkan embedding wajah yang kemudian digunakan dalam proses klasifikasi.

Hasil eksperimen menunjukkan bahwa model InsightFace memperoleh nilai akurasi sebesar 96%, yang berada di antara FaceNet dan VGGFace2. Nilai ini menunjukkan bahwa InsightFace memiliki performa yang sangat baik dalam mendeteksi ASD berdasarkan citra wajah. Nilai precision dan recall yang diperoleh juga menunjukkan bahwa model memiliki kemampuan yang baik dalam mengklasifikasikan data, serta cukup stabil dalam menghadapi variasi citra.

Keunggulan InsightFace terletak pada penggunaan ArcFace yang mampu meningkatkan margin antar kelas dalam ruang fitur. Hal ini membuat model lebih sensitif terhadap perbedaan kecil pada fitur wajah, sehingga mampu meningkatkan akurasi klasifikasi. Namun demikian, performa InsightFace masih sedikit di bawah FaceNet, yang menunjukkan bahwa pendekatan triplet loss pada FaceNet lebih efektif dalam menghasilkan embedding yang optimal untuk dataset yang digunakan dalam penelitian ini.

Pada Gambar 5 terlihat bahwa model InsightFace mampu menghasilkan performa klasifikasi yang stabil dengan tingkat akurasi yang tinggi dalam mendeteksi ASD berdasarkan citra wajah.



Gambar 5. Hasil Pelatihan InsightFace

3.4 Perbandingan Performa Model

Untuk memahami lebih jelas bagaimana masing-masing model berkinerja, dilakukan perbandingan berdasarkan metrik evaluasi yang digunakan.

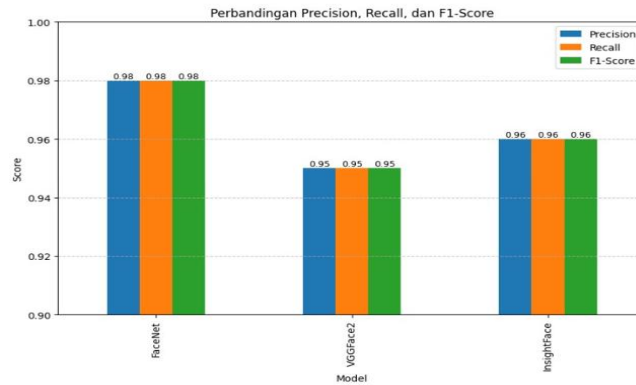
Tabel 1. Perbandingan Performa Model

Model	Accuracy	Precision	Recall	F1-Score
FaceNet	0.98	0.98	0.98	0.98
InsightFace	0.96	0.96	0.96	0.96
VGGFace2	0.95	0.95	0.95	0.95

Berdasarkan Tabel 1, dapat dilihat bahwa FaceNet memiliki performa terbaik di antara ketiga model berdasarkan metrik accuracy, precision, recall, dan F1-score. InsightFace berada di posisi kedua, sedangkan VGGFace2 memiliki performa terendah, meskipun masih dalam kategori baik.

Perbedaan performa ini menunjukkan bahwa metode ekstraksi fitur memiliki pengaruh yang signifikan terhadap hasil klasifikasi. Model yang dirancang khusus untuk face embedding seperti FaceNet dan InsightFace cenderung memiliki performa yang lebih baik dibandingkan model berbasis CNN umum.

Pada Gambar 6 terlihat perbandingan performa ketiga model berdasarkan metrik evaluasi yang digunakan dalam penelitian.



Gambar 6. Perbandingan Model

3.5 Analisis Confusion Matrix

Confusion matrix digunakan untuk menganalisis distribusi kesalahan klasifikasi pada masing-masing model. Pada model FaceNet, sebagian besar data berhasil diklasifikasikan dengan benar, dengan jumlah kesalahan yang sangat kecil. Hal ini menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam membedakan antara kelas ASD dan non-ASD. Pada model InsightFace, terdapat beberapa kesalahan klasifikasi, namun jumlahnya masih relatif kecil. Kesalahan ini umumnya terjadi pada citra yang memiliki karakteristik wajah yang mirip antara kedua kelas. Sementara itu, pada model VGGFace2, jumlah kesalahan klasifikasi sedikit lebih tinggi dibandingkan dua model lainnya. Hal ini menunjukkan bahwa model masih mengalami kesulitan dalam membedakan beberapa pola wajah yang memiliki kemiripan tinggi. Analisis ini menunjukkan bahwa kualitas embedding fitur sangat berpengaruh terhadap performa model dalam melakukan klasifikasi.

3.6 Pembahasan

Berdasarkan hasil eksperimen yang telah dilakukan, dapat disimpulkan bahwa penggunaan model berbasis face recognition memberikan hasil yang sangat baik dalam deteksi ASD berbasis citra wajah. FaceNet menunjukkan performa terbaik karena menggunakan pendekatan triplet loss yang secara langsung mengoptimalkan jarak antar kelas dalam ruang fitur. Hal ini membuat model mampu menghasilkan embedding yang sangat diskriminatif. InsightFace juga menunjukkan performa yang tinggi karena mampu meningkatkan separasi antar kelas melalui margin sudut. Namun, pendekatan ini masih sedikit kurang optimal dibandingkan triplet loss dalam konteks dataset yang digunakan. Sementara itu, VGGFace2 meskipun memiliki performa yang baik, masih kalah dibandingkan dua model lainnya karena pendekatannya lebih umum dan tidak secara spesifik dirancang untuk optimasi embedding wajah.

Hasil penelitian ini sejalan dengan beberapa penelitian sebelumnya yang menunjukkan bahwa model berbasis face embedding memiliki performa yang lebih baik dibandingkan CNN konvensional dalam tugas pengenalan wajah. Selain itu, hasil penelitian ini juga menunjukkan bahwa analisis citra wajah dapat digunakan sebagai alternatif dalam mendukung deteksi dini ASD. Dengan memanfaatkan teknologi deep learning, proses deteksi dapat dilakukan secara otomatis, cepat, dan objektif.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa pengembangan model deteksi Autism Spectrum Disorder (ASD) berbasis analisis citra wajah menggunakan pendekatan deep learning menunjukkan performa yang sangat baik dalam mengklasifikasikan individu ASD dan non-ASD. Penelitian ini berhasil membuktikan bahwa pemanfaatan model berbasis face recognition seperti FaceNet, VGGFace2, dan InsightFace mampu menghasilkan representasi fitur wajah yang efektif untuk mendukung proses klasifikasi. Dari ketiga model yang diuji, FaceNet menunjukkan performa terbaik dengan tingkat akurasi sebesar 98%, diikuti oleh InsightFace sebesar 96% dan VGGFace2 sebesar 95%. Hal ini menunjukkan bahwa pendekatan berbasis face embedding, khususnya dengan metode triplet loss pada FaceNet, memiliki kemampuan yang lebih optimal dalam membedakan karakteristik wajah individu ASD dibandingkan pendekatan CNN konvensional. Selain itu, hasil penelitian ini juga menunjukkan bahwa analisis citra wajah

dapat menjadi alternatif yang potensial dalam mendukung deteksi dini ASD secara otomatis, cepat, dan objektif. Penggunaan teknologi ini dapat membantu mengurangi ketergantungan pada metode diagnosis konvensional yang bersifat subjektif dan memerlukan waktu yang relatif lama. Dengan demikian, penelitian ini memberikan kontribusi dalam pengembangan sistem berbasis kecerdasan buatan di bidang kesehatan, khususnya dalam deteksi gangguan perkembangan saraf. Meskipun demikian, penelitian ini masih memiliki beberapa keterbatasan, di antaranya penggunaan dataset yang terbatas pada jumlah dan variasi tertentu, serta belum mempertimbangkan faktor lain seperti kondisi pencahayaan ekstrem, variasi usia yang lebih luas, dan keberagaman etnis secara menyeluruh. Selain itu, penelitian ini hanya berfokus pada perbandingan model ekstraksi fitur tanpa mengintegrasikan metode klasifikasi lanjutan atau pendekatan ensemble yang berpotensi meningkatkan performa sistem. Oleh karena itu, penelitian selanjutnya diharapkan dapat menggunakan dataset yang lebih besar dan beragam, serta mengembangkan metode yang lebih kompleks seperti integrasi ensemble learning atau pendekatan multimodal untuk meningkatkan akurasi dan kemampuan generalisasi model dalam mendeteksi ASD.

REFERENCES

- [1] A. Ahmad Hania, "Mengenal Artificial Intelligence, Machine Learning, & Deep Learning," *J. Teknol. Indones.*, vol. 1, no. June, pp. 1–6, 2017, [Online]. Available: <https://amt-it.com/mengenal-perbedaan-artificial-intelligence-machine-learning-deep-learning/>
- [2] S. Ahmad, P. Singh, and A. K. Sagar, "A Survey on Big Data Analytics," in *2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, 2018, pp. 256–260. doi: 10.1109/ICACCCN.2018.8748774.
- [3] C. A. Goodfellow I, Bengio Y, "Deep learning," *MIT Press. Cambridge, MA, USA*, vol. 22, no. 4, pp. 351–354, 2016.
- [4] G. H. Yann LeCun, Yoshua Bengio, "Deep learning," vol. 521, pp. 436–444, [Online]. Available: <https://www.nature.com/articles/nature14539>
- [5] F. Chollet, "Deep learning with Python," pp. 113–115, 2022, doi: 10.31211/interacoes.n42.2022.r1.
- [6] J. Heaton, "Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning," *Genet. Program. Evolvable Mach.*, vol. 19, no. 1, pp. 305–307, 2018, doi: 10.1007/s10710-017-9314-z.
- [7] M. S. Alam, M. M. Rashid, A. Jazlan, E. E. Alahi, M. Kchaou, and K. A. B. Alharthi, "Robust Autism Spectrum Disorder Screening Based on Facial Images (For Disability Diagnosis): A Domain-Adaptive Deep Ensemble Approach," pp. 1–32, 2025.
- [8] C. Commons, A. License, and T. H. H. Aldhyani, "Retracted: Facial Features Detection System to Identify Children with Autism Spectrum Disorder: Deep Learning Models," vol. 2022, 2023, doi: 10.1155/2022/3941049.
- [9] L. Darmiyati, "Analisis Geometri Ciri Morfologis Wajah Anak Autistik Indonesia dan Non-Indonesia Berdasarkan Metrik Jarak Geometris," vol. 10, no. 1, pp. 723–730, 2025.
- [10] S. Alam, M. M. Rashid, R. Roy, A. R. Faizabadi, K. D. Gupta, and M. Ahsan, "Empirical Study of Autism Spectrum Disorder Diagnosis Using Facial Images by Improved Transfer Learning Approach," pp. 1–18, 2022.
- [11] F. Rosindo *et al.*, "Computer aided diagnosis of neurodevelopmental disorders and genetic syndromes based on facial images – A systematic literature review," vol. 9, no. October, 2023, doi: 10.1016/j.heliyon.2023.e20517.
- [12] S. Z. J. Deng, J. Guo, N. Xue, I. Kotsia, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pp. 4690–4699, 2019, doi: 10.1109/CVPR.2019.00482.
- [13] M. Faheem, I. Ahmad, J. Rashid, N. A. Khan, and R. Amin, "Autism spectrum disorder detection using facial images : A performance comparison of pretrained convolutional neural networks," no. December 2023, pp. 227–239, 2024, doi: 10.1049/htl2.12073.
- [14] X. Xu, M. Du, H. Guo, J. Chang, and X. Zhao, "Lightweight FaceNet Based on MobileNet," pp. 1–16, 2021, doi: 10.4236/ijis.2021.111001.
- [15] J. P. F. Schroff, D. Kalenichenko, "FaceNet: A Unified Embedding for Face Recognition and Clustering," 2015, doi: 10.1109/CVPR.2015.7298682.
- [16] A. Z. Q. Cao, L. Shen, W. Xie, O. M. Parkhi, "VGGFace2: A Dataset for Recognising Faces across Pose and Age," *2018 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG 2018)*, pp. 67–74, 2018, doi: 10.1109/FG.2018.00020.
- [17] J. G. Deepinsight, "InsightFace: State-of-the-art 2D & 3D Face Analysis Toolbox." [Online]. Available: <https://github.com/deepinsight/insightface>
- [18] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, 2019, doi: 10.1186/s40537-019-0197-0.
- [19] F. Marpaung, F. Aulia, and R. C. Nabila, *Computer Vision Dan Pengolahan Citra Digital*. 2022. [Online]. Available: www.pustakaaksara.co.id
- [20] Y. Amrozi, D. Yuliati, A. Susilo, N. Novianto, and R. Ramadhan, "Klasifikasi Jenis Buah Pisang Berdasarkan Citra Warna dengan Metode SVM," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 11, no. 3, pp. 394–399, 2022, doi: 10.32736/sisfokom.v11i3.1502.