

BAB I PENDAHULUAN

1.1 Latar Belakang

Autism Spectrum Disorder (ASD) adalah gangguan perkembangan saraf yang secara signifikan memengaruhi perilaku adaptif, interaksi sosial, dan komunikasi. Laporan dari World Health Organization (WHO) [1] menunjukkan bahwa prevalensi ASD meningkat secara signifikan setiap tahun, dan keterlambatan diagnosis sering menyebabkan intervensi tidak optimal. Berdasarkan data Badan Pusat Statistik (BPS) pada tahun 2023, diperkirakan terdapat 3,2 juta anak di Indonesia yang menderita ASD [2]. Di sisi lain, diagnosis ASD konvensional masih bergantung pada observasi klinis, wawancara, dan instrumen seperti ADOS-2 atau DSM-5 yang membutuhkan tenaga ahli, biaya tinggi, dan tidak tersedia merata. Dalam situasi seperti ini, metode pendukung diagnosis yang lebih cepat, objektif, dan mudah diakses diperlukan, terutama di daerah dengan keterbatasan layanan kesehatan.

Sebaliknya, kemajuan *computer vision* berbasis *deep learning* membuka peluang pengembangan sistem skrining ASD melalui citra wajah. Sejumlah studi menunjukkan bahwa pola morfologi wajah, ekspresi mikro, arah pandang, dan struktur *craniofacial* dapat menjadi *biomarker* awal untuk indikasi ASD [3]. Akibatnya, ini mendorong penggunaan teknik otomatis berbasis *deep learning*. Studi yang menggunakan arsitektur seperti VGG19, ResNet50, Xception, dan EfficientNetB0 menunjukkan tingkat akurasi yang tinggi antara 85–95% [4], [5]. Namun, sebagaimana dikritik tentang penggunaan *heavyweight models* [6] kinerja tinggi yang dihasilkan belum menunjukkan kemampuan generalisasi yang baik, karena model cenderung terlalu menyesuaikan dengan dataset yang lebih kecil dan dapat menyebabkan hilangnya data karena pemisahan data yang tidak benar-benar independen.

Fakta ini menunjukkan adanya celah metodologis. Misalnya, meskipun penelitian sebelumnya telah berhasil menggunakan berbagai arsitektur *deep*

learning untuk mencapai akurasi dan AUC yang tinggi, sebagian besar penelitian masih menitikberatkan pada model berparameter besar, yang kurang efektif untuk digunakan di perangkat yang memiliki keterbatasan komputasi. Selain itu, belum banyak studi yang secara eksplisit membahas pengaruh variasi *preprocessing* wajah maupun evaluasi efisiensi komputasi model. Oleh karena itu, masih ada ruang untuk penelitian tentang bagaimana arsitektur *lightweight* dapat menjadi lebih kompetitif sekaligus lebih sesuai untuk digunakan pada sistem skrining ASD yang realistis.

Studi ini berfokus pada evaluasi arsitektur ringan untuk klasifikasi ASD berbasis gambar wajah, dengan MobileViTv2 sebagai model utama setelah melalui evaluasi komparatif pada fase awal eksperimen. Sebagai metrik utama untuk evaluasi *Area Under the Curve* (AUC) digunakan. Ini dipilih karena metrik ini lebih kuat, tidak bergantung pada batas, dan memiliki kemampuan untuk menunjukkan perbedaan antara *sensitivity* dan *specificity* yang relevan dalam konteks skrining klinis. Diharapkan bahwa metode ini dapat memberikan gambaran yang tidak bias tentang kinerja dan efisiensi komputasi model *lightweight* serta tentang bagaimana dua skema *preprocessing* gambar—*dataset default* dan *cropped*—berpengaruh terhadap kemampuan klasifikasi.

1.2 Rumusan Masalah

Berdasarkan informasi sebelumnya, penelitian ini menghasilkan dua pertanyaan utama:

1. Algoritma apakah yang cocok digunakan untuk mendeteksi autisme berdasarkan citra wajah yang memiliki performa bagus dan efisien?
2. Apa pengaruh *preprocessing* citra wajah (*default dataset* vs *cropped dataset*) terhadap kemampuan klasifikasi model *lightweight* sistem deteksi autisme?

1.3 Batasan Masalah

Batas-batas berikut ditetapkan untuk menjaga fokus penelitian:

1. *Autism Spectrum Disorder* (ASD) yang dibahas dalam penelitian ini dibatasi sebagai gangguan perkembangan saraf dalam spektrum autisme secara umum, yang ditinjau melalui indikasi visual pada citra

wajah anak. Penelitian ini tidak membedakan jenis atau subtype ASD tertentu, serta tidak mempertimbangkan secara langsung aspek perilaku, komunikasi, sensorik, maupun faktor genetik.

2. Studi ini menggunakan eksperimen komparatif bertahap untuk mengevaluasi tiga arsitektur ringan: MobileViTv2, EfficientNetV2-S, dan MobileNetV3-Large. Evaluasi dilakukan dalam tiga tahap. Pertama, instruksi dasar dengan backbone yang didinginkan; kedua, optimasi hyperparameter menggunakan rangka kerja Optuna dan penyempurnaan dengan pemanasan partial backbone; dan terakhir, evaluasi akhir dengan *Test Time Augmentation* (TTA). Pada tahap akhir penelitian, model yang paling efektif dipilih untuk dianalisis.
3. Data yang digunakan terbatas pada dataset gambar wajah gangguan spektrum autisme (ASD), yang terdiri dari 2940 gambar yang termasuk dalam dua kelas ketersediaan, yaitu *autistik* dan *non-autistik* [7].
4. Metrik valuasi utama yang dipakai adalah AUC karena mengukur kemampuan model untuk membedakan kelas positif dan negatif di seluruh rentang ambang batas (threshold). Ini relevan untuk skrining klinis.
5. Dalam analisis efisiensi, faktor-faktor seperti jumlah parameter, ukuran model, FLOPs, dan trade-off dibatasi waktu dan sumber daya. Tidak ada faktor eksternal seperti kebijakan organisasi.
6. Teknik *preprocessing* yang dikaji hanya mencakup *full face* (default dataset) dan *cropped dataset* yang diekstraksi menggunakan algoritma *Multi-task Cascaded Convolutional Networks* (MTCNN).
7. Model dilatih menggunakan strategi transfer learning dua tahap untuk optimasi performa dan mencegah overfitting. Detail konfigurasi pelatihan dijelaskan pada Bab III.
8. Untuk memastikan distribusi kelas yang seimbang pada setiap subset, evaluasi dilakukan dengan skema pembagian data train/validation/test

yang konsisten (70%: 15%: 15%) dengan sampel stratified.

9. Maksimal lima belas tes dilakukan untuk melakukan optimasi hyperparameter menggunakan framework optimasi Bayesian Optuna. Setiap tes dijalankan selama dua puluh epoch, dan untuk efisiensi komputasi digunakan mekanisme pruning otomatis.
10. Pada fase evaluasi akhir, *Test-Time Augmentation* (TTA) diterapkan untuk meningkatkan kekuatan prediksi dengan menggabungkan hasil inferensi dari empat transformasi gambar: *original*, *horizontal flip*, *scaling*, dan *rotation* ringan ($\pm 10^\circ$).

1.4 Tujuan Penelitian

Tujuan yang akan dicapai oleh peneliti dalam penelitiannya adalah:

1. Mengevaluasi performa dan efisiensi komputasi arsitektur *lightweight* MobileViTv2 dalam klasifikasi ASD berbasis citra wajah, berdasarkan metrik AUC sebagai metrik utama, akurasi, jumlah parameter.
2. Menganalisis pengaruh *preprocessing* citra wajah (*default dataset* vs *cropped dataset*) terhadap performa klasifikasi model MobileViTv2 pada dataset ASD.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat bagi beberapa pihak:

1. Manfaat Teoritis:

Memberi kontribusi empiris tentang manfaat antara kompleksitas model dan kinerja klasifikasi dalam domain *screening* medis, terutama dalam hal deteksi ASD berbasis citra wajah menggunakan arsitektur *deep learning lightweight*.

2. Manfaat Praktis:

- a. Untuk pengembangan teknologi kesehatan: Menyediakan evaluasi komparatif arsitektur ringan yang dapat digunakan

untuk membangun sistem skrining ASD pada perangkat edge computing atau mobile.

- b. Untuk masyarakat: Mendukung upaya untuk meningkatkan aksesibilitas skrining dini autisme spektral (ASD), terutama di wilayah dengan keterbatasan tenaga ahli dan fasilitas diagnostik formal.

3. Manfaat bagi Peneliti:

Sebagai alat untuk meningkatkan kemampuan dalam penglihatan komputer dan pembelajaran mendalam, terutama dalam hal penerapan dan evaluasi model *lightweight* untuk aplikasi kesehatan.

1.6 Sistematika Penulisan

Sistematika penulisan pada laporan ini dibagi menjadi lima bab, untuk memberikan gambaran secara menyeluruh terkait laporan penelitian ini, bagian-bagian tersebut adalah sebagai berikut.

BAB I PENDAHULUAN, Bab ini membahas latar belakang masalah, rumusan masalah, keterbatasan masalah, tujuan penelitian, keuntungan dari penelitian, dan prosedur untuk menyusun tugas akhir.

BAB II TINJAUAN PUSTAKA, Bab ini membahas penelitian terdahulu tentang deteksi gangguan spektrum autisme berbasis citra wajah. Ini juga mencakup dasar teori tentang konsep seperti Autism Spectrum Disorder, Deep Learning, Convolutional Neural Network, Vision Transformer, Light Weight Architecture (MobileViTv2, EfficientNetV2, MobileNetV3), Transfer Learning, dan metrik evaluasi (AUC).

BAB III METODE PENELITIAN, Bab ini membahas alur penelitian. Ini mencakup pengumpulan dan *preprocessing* data menggunakan MTCNN, augmentasi data, proses seleksi model melalui evaluasi komparatif beberapa arsitektur ringan, konfigurasi pelatihan dua tahap, metrik untuk mengevaluasi performa dan efisiensi komputasi, dan metode untuk mengevaluasi pengaruh *preprocessing*.

BAB IV HASIL DAN PEMBAHASAN, Bab ini membahas hasil evaluasi komparatif arsitektur ringan untuk memilih model terbaik. Ini membahas performa model yang dipilih (MobileViTv2), dan mengevaluasi pengaruh preprocessing (*dataset default* versus *dataset cropped*) terhadap kemampuan klasifikasi. Bab ini juga membahas temuan penelitian dalam konteks aplikasi praktis dan keterbatasan penelitian.

BAB V PENUTUP, Bab ini berisi kesimpulan dari hasil penelitian dan saran yang dapat menjadi pertimbangan untuk penelitian berikutnya.

