

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Penelitian ini mengusulkan modifikasi arsitektur EfficientNetV2-S dengan integrasi *Seat Position Embedding* (SPE), yaitu penambahan *input* informasi data posisi duduk siswa dalam bentuk *embedding* kategorial (kiri, tengah, kanan), untuk mengklasifikasikan arah pose kepala siswa di ruang kelas dengan tujuan utama meningkatkan akurasi interpretasi atensi visual melalui konteks spasial. Pada tahap pengembangan model klasifikasi, ditemukan bahwa penerapan berbagai teknik optimasi secara bertahap memberikan peningkatan performa yang signifikan. Penggunaan augmentasi data terbukti krusial untuk mengatasi *overfitting* dan meningkatkan kemampuan generalisasi model dasar. Selanjutnya, pendekatan *transfer learning* dengan arsitektur EfficientNetV2-S secara drastis meningkatkan akurasi sekaligus efisiensi waktu pelatihan. Optimasi lebih lanjut menggunakan teknik *class weight* berhasil menangani masalah ketidakseimbangan kelas dan memberikan peningkatan performa tambahan. Pada tahap terakhir, penambahan SPE dilakukan dengan cara menghilangkan bagian *classifier* dari EfficientNetV2. Dengan demikian, jaringan dapat menerima *input* tambahan berupa *embedding* posisi duduk siswa yang kemudian diintegrasikan ke dalam proses klasifikasi.

Hasil eksperimen dengan konfigurasi akhir menunjukkan bahwa penambahan informasi posisi duduk dalam bentuk *embedding* kategorial mampu meningkatkan performa model, dengan peningkatan akurasi dari 82,53% menjadi 83,25% dan *f1-score* dari 82,84% menjadi 83,41% pada model EfficientNetV2-S. Pendekatan ini juga berhasil meningkatkan performa pada arsitektur berbasis CNN lainnya seperti MobileNetV3-L dan ConvNeXt-Tiny. Sementara itu, pada model ResNet50, penambahan informasi SPE tidak memberikan peningkatan performa yang signifikan. Hal ini kemungkinan disebabkan oleh karakteristik arsitekturnya yang dalam dengan mekanisme residual, serta kapasitas representasi spasial yang tinggi, sehingga model tersebut mungkin sudah mampu menangkap konteks posisi duduk siswa secara implisit hanya dari informasi visual pada citra. Kelebihan pendekatan ini adalah kemampuannya dalam mempertimbangkan konteks spasial

tanpa perlu melakukan deteksi *landmark* wajah dan sudut Euler yang rentan terhadap kualitas citra rendah. Namun, kelemahannya terletak pada potensi ambiguitas antar kelas yang berdekatan secara visual, seperti antara pose kepala depan–bawah dan depan–atas, serta rotasi kecil ke kiri atau kanan yang terkadang sedikit sulit dibedakan dari kelas depan.

Berdasarkan serangkaian tahap eksperimen dan analisis yang telah dilakukan, penelitian ini berhasil mengembangkan dan mengevaluasi sebuah sistem untuk deteksi arah pose kepala siswa di ruang kelas. Pada tahap integrasi untuk membentuk sistem deteksi pose kepala secara menyeluruh, model yang diusulkan, yaitu EfficientNetV2-S dengan SPE diintegrasikan dengan model deteksi kepala YOLOv8 yang menunjukkan performa deteksi yang tinggi. Hasil evaluasi *end-to-end* memvalidasi bahwa sistem yang menggunakan model EfficientNetV2-S dengan SPE mencapai performa yang cukup baik, dengan nilai mAP sebesar 0,508 dibandingkan dengan mAP *baseline* sebesar 0,490. Peningkatan yang terukur ini membuktikan bahwa SPE memberikan dampak positif pada kinerja sistem secara keseluruhan. Lebih lanjut, analisis efisiensi menunjukkan bahwa peningkatan akurasi dicapai tanpa menambah beban komputasi yang signifikan, baik dari segi ukuran model maupun waktu inferensi.

5.2 Saran

Penelitian selanjutnya disarankan agar dilakukan pengembangan set data yang lebih baik, berikut merupakan serangkaian saran yang dapat dilakukan untuk penelitian selanjutnya dari sisi set data:

1. Jumlah partisipan perlu ditambah dan dibuat lebih bervariasi untuk meningkatkan keberagaman subjek agar model tidak hanya belajar dari ciri visual siswa yang sama, yang dapat menyebabkan *overfitting* serta membuat model mengandalkan ingatan terhadap wajah individu alih-alih mempelajari pola arah pose kepala secara umum.
2. Penambahan jumlah gambar dalam set data juga diperlukan, namun harus tetap menjaga tingkat redundansi visual yang rendah sehingga variasi data tetap tinggi dan setiap sampel tetap mengandung informasi yang bermakna.

3. Keseimbangan jumlah sampel untuk setiap kelas arah pose kepala harus menjadi perhatian penting, karena set data yang seimbang akan membantu model belajar secara adil dari semua kelas dan meningkatkan performa pada kelas-kelas minoritas.

Dari sisi model, disarankan agar pengembangan model dilakukan lebih lanjut melalui eksplorasi secara mendalam. Berikut merupakan serangkaian saran yang dapat dilakukan untuk penelitian selanjutnya dari sisi model:

1. Dalam penelitian ini, posisi duduk siswa dikategorikan secara horizontal (kiri, tengah, kanan), penambahan dimensi vertikal seperti posisi depan dan belakang kelas mungkin dapat memberikan konteks tambahan yang berguna dalam memahami arah pose kepala siswa secara lebih akurat. Dengan demikian, SPE dapat dirancang lebih informatif dan representatif.
2. Pendekatan integrasi fitur yang saat ini masih cukup sederhana (yaitu *concatenation*) dapat ditingkatkan dengan mengeksplorasi mekanisme fusi yang lebih adaptif, seperti *attention mechanism* (misalnya *Cross-Attention*), yang memungkinkan model untuk secara dinamis menyesuaikan bobot antara informasi visual dan spasial dalam proses klasifikasi.
3. Eksplorasi terhadap arsitektur *hybrid* yang menggabungkan keunggulan CNN dan Transformer (seperti Vision Transformer atau Swin Transformer) juga dapat menjadi arah potensial untuk meningkatkan kapasitas model dalam memahami hubungan spasial global dan lokal.