

Deteksi Antusiasme Siswa dengan Algoritma Yolov8 pada Proses Pembelajaran *Daring*

Kartika Salma ^{1*}, Syarif Hidayat ²

^{1,2}Program Studi Informatika, Fakultas Teknologi Industri, Universitas Islam Indonesia, Kabupaten Sleman, Daerah Istimewa Yogyakarta, Indonesia.

Email: 20523172@students.uui.ac.id ^{1*}, syarif@uui.ac.id ²

Histori Artikel:

Dikirim 13 Maret 2024; *Diterima dalam bentuk revisi* 9 April 2024; *Diterima* 27 April 2024; *Diterbitkan* 10 Mei 2024. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

Abstrak

Penerapan teknologi Face Emotion Recognition (FER) dalam kelas daring membuka peluang baru untuk memantau respon emosional siswa secara efektif dan menyesuaikan pendekatan pembelajaran. Melalui FER, pengajar dapat memantau respon emosional siswa secara real-time terhadap materi pembelajaran dan memungkinkan penyesuaian yang cepat sesuai dengan kebutuhan individu. Selain itu, teknologi ini juga dapat digunakan untuk mendeteksi tingkat antusiasme atau ketidakantusiasan siswa terhadap pembelajaran yang memungkinkan untuk pengoptimalan strategi pengajaran. Penelitian ini memfokuskan perhatian pada implementasi algoritma YOLOv8 dalam deteksi antusiasme siswa, dengan membandingkan performa model YOLOv8n, YOLOv8s, YOLOv8m, dan YOLOv8l. Hasil pengujian menunjukkan bahwa YOLOv8n memberikan performa terbaik dengan tingkat akurasi mencapai 95,3% dan waktu inferensi yang cepat, yaitu 62ms, memungkinkan deteksi objek secara real-time. Dengan demikian, penerapan YOLOv8 dalam konteks ini bertujuan untuk mendeteksi antusiasme siswa secara real-time dan memungkinkan pengajar dapat dengan cepat menyesuaikan pendekatan mereka dengan kebutuhan siswa. Selain itu, penelitian ini juga berkontribusi pada peningkatan kualitas pembelajaran daring dengan memberikan wawasan tentang keterlibatan emosional siswa serta sebagai alat yang dapat membantu pengajar untuk lebih memahami dan merespon secara tepat terhadap emosi siswa selama proses pembelajaran daring berlangsung.

Kata Kunci: Face Emotion Recognition (FER); Respon Emosional; Real-Time; Deteksi Antusiasme; Akurasi; Yolov8.

Abstract

The implementation of Face Emotion Recognition (FER) technology in online classes opens up new opportunities to effectively monitor students' emotional responses and adjust the teaching approach. Through FER, instructors can monitor students' emotional responses to learning materials in real-time and enable quick adjustments based on individual needs. Additionally, this technology can also be used to detect the level of enthusiasm or lack thereof among students towards the learning process, allowing for the optimization of teaching strategies. This study focuses on the implementation of the YOLOv8 algorithm in detecting students' enthusiasm, comparing the performance of YOLOv8n, YOLOv8s, YOLOv8m, and YOLOv8l models. Test results show that YOLOv8n performs the best with an accuracy rate of 95.3% and a fast inference time of 62ms, enabling real-time object detection. Thus, the application of YOLOv8 in this context aims to detect students' enthusiasm in real-time and allows instructors to quickly adjust their approach to meet students' needs. Furthermore, this research contributes to improving the quality of online learning by providing insights into students' emotional engagement and serving as a tool to help instructors better understand and respond appropriately to students' emotions during the online learning process.

Keyword: Face Emotion Recognition (FER); Emotional Response; Real-Time; Enthusiasm Detection; Accuracy; Yolov8.

1. Pendahuluan

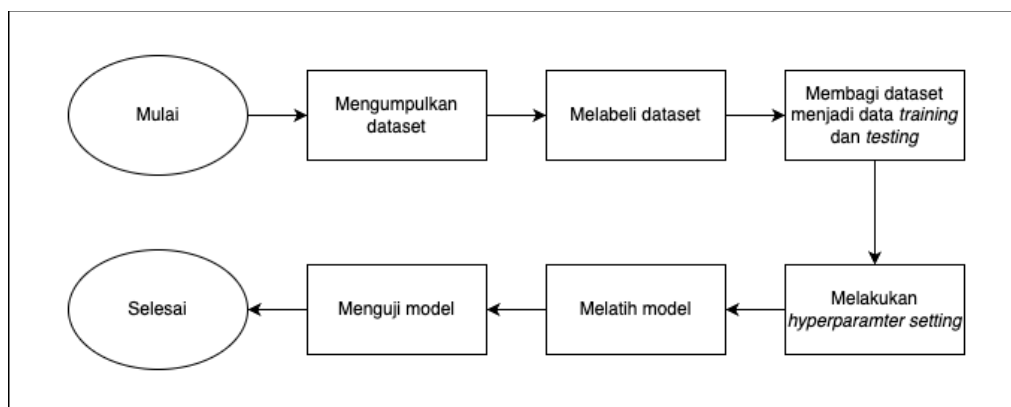
Dalam era kelas daring yang semakin meluas di universitas dan lembaga kursus, muncul kebutuhan akan pemantauan respon emosional siswa sebagai bagian integral dari pengalaman pembelajaran. Pandemi COVID-19 telah mempercepat adopsi kelas daring, tetapi dalam prosesnya, dosen dan pengajar menghadapi tantangan dalam mengevaluasi respon siswa secara efektif. Dalam konteks ini, teknologi *Face Emotion Recognition* (FER) muncul sebagai solusi potensial untuk memahami dinamika emosional siswa. Ekspresi wajah, yang mencerminkan keadaan emosional seseorang, menjadi elemen utama dalam memahami respon emosional manusia. Ekspresi ini melibatkan respon psikologis terhadap stimulus, baik itu berasal dari faktor internal maupun eksternal (Anggraini *et al.*, 2019). Emosi wajah dapat diidentifikasi melalui ekspresi mikro yang umumnya bersifat universal, menjadikannya sarana utama bagi manusia untuk menyampaikan keadaan emosional dan niat mereka (C. Darwin and P. Prodger, 1998). Dalam klasifikasi ekspresi wajah, Ekman dan Friesen mengembangkan *Facial Action Coding System* (FACS) yang mencakup tujuh kategori, seperti marah, jijik, takut, bahagia, netral, sedih, dan terkejut (P. Ekman and W. V. Friesen, 1971). Ekspresi wajah memainkan peran penting dalam komunikasi antarpribadi ketika kita berurusan dengan emosi dan niat kita, serta menginterpretasikan yang orang lain rasakan (Kovalchuk *et al.*, 2022). Dalam konteks pembelajaran, pemahaman terhadap emosi siswa dapat memberikan wawasan berharga bagi pengajar untuk menyesuaikan metode pengajaran dan menciptakan lingkungan pembelajaran yang lebih mendukung. Penting untuk memastikan keterlibatan siswa selama sesi pembelajaran online agar lingkungan pembelajaran tersebut menjadi lebih interaktif, sebagaimana halnya dalam kelas luring. Salah satu pendekatan yang diusulkan adalah menggunakan teknologi berbasis deep learning yang menggunakan ekspresi wajah siswa untuk mendeteksi keterlibatan mereka secara real-time selama sesi pembelajaran daring (Gupta *et al.*, 2023).

Pemanfaatan FER dalam pembelajaran daring memungkinkan pengajar memantau respon emosional siswa dan menyesuaikan pendekatan mereka. Penggunaan FER dalam konteks pembelajaran daring bukan hanya sekadar sebuah terobosan teknologi, tetapi juga sebuah langkah penting menuju transformasi pendidikan yang lebih interaktif dan adaptif. Melalui FER, pengajar dapat secara *real-time* memonitor respon emosional siswa terhadap materi pembelajaran dan dengan cepat menyesuaikan pendekatan mereka sesuai dengan kebutuhan individu. Selain itu, penggunaan FER juga dapat membantu dalam mendeteksi tingkat antusiasme atau ketidakantusiasme siswa terhadap materi pembelajaran. Antusiasme merujuk pada ketertarikan intens terhadap suatu hal yang tercermin dalam tingkat kebahagiaan yang tinggi dan semangat yang menyala-nyala (Ohoirat, 2020). Penggunaan teknologi ini tidak hanya tentang mengikuti perkembangan zaman, tetapi juga menjawab tuntutan akan pembelajaran yang lebih adaptif dan responsif terhadap kebutuhan individual siswa. Penelitian ini dilakukan dengan merujuk pada beberapa karya ilmiah sebelumnya yang telah dilakukan sebagai bahan perbandingan. Berbagai pendekatan telah diusulkan dalam penelitian sebelumnya untuk mengenali ekspresi wajah, mulai dari penggunaan *Convolutional Neural Network* (CNN) hingga algoritma *You Only Look Once* (YOLO). Seperti penelitian yang pernah dilakukan oleh Prasetyawan, beliau menerapkan CNN untuk identifikasi emosi melalui ekspresi wajah manusia (Prasetyawan *et al.*, 2020). Azizi (2021) menguji tiga metode optimasi dan variasi jumlah epoch untuk mendeteksi emosi berdasarkan ekspresi (Azizi, 2021). Kemudian penelitian oleh Adhitama (2022) menerapkan pendekatan yang melibatkan algoritma YOLOv5 (*You Only Look Once*) dan kerangka kerja PyTorch untuk mengembangkan kemampuan *deep learning* dalam mendeteksi ekspresi wajah (*Facial Expression Recognition*) dan mengelaskannya menjadi “antusias” dan “tidak antusias” dengan akurasi mencapai 87% (Adhitama, 2022). Meskipun penelitian terdahulu telah mengeksplorasi berbagai metode termasuk YOLO, tetapi terdapat kesenjangan dalam literatur yang belum sepenuhnya mengeksplorasi potensi YOLOv8, khususnya dalam konteks deteksi antusiasme siswa dalam pembelajaran daring. Oleh karena itu, penelitian ini bertujuan untuk mengisi kesenjangan tersebut dan memfokuskan perhatian pada pengenalan ekspresi wajah, khususnya dalam mengukur respon antusiasme siswa terhadap pembelajaran daring dan meningkatkan tingkat akurasi dalam mendeteksi antusiasme siswa.

2. Metode Penelitian

Metode yang digunakan pada penelitian ini adalah algoritma YOLOv8 yang merupakan algoritma deteksi objek yang populer dan dikenal memiliki tingkat akurasi yang tinggi serta waktu inferensi yang singkat. YOLO dipilih karena keunggulannya dibanding model deteksi objek yang lain, yaitu ukurannya lebih kecil dan kemampuan komputasinya lebih cepat (Jiang *et al.*, 2022). YOLOv8, dikeluarkan oleh Ultralytics pada bulan Januari 2023, yang merupakan pengembangan lanjutan dari karya mereka, yaitu YOLOv5 pada tahun 2020. Meskipun belum terdapat paper resmi yang diterbitkan untuk YOLOv8, model ini sudah dapat digunakan untuk pelatihan dan pengujian oleh masyarakat umum (Terven *et al.*, 2023).

Penelitian ini menggunakan data dari FER 2013 dengan hanya menggunakan 2000 data. Data-data ini dipilih secara proporsional, mewakili 10% dari setiap kelas yang ada dalam dataset FER 2013 dengan tujuan menghindari terjadinya *class imbalance*. Ketidakseimbangan kelas (*class imbalance*) merujuk pada situasi di mana jumlah sampel atau data dalam setiap kelas atau kategori dalam sebuah dataset tidak seimbang. Artinya, ada kelas atau kategori yang memiliki jumlah data yang jauh lebih banyak daripada kelas atau kategori lainnya. Ketidakseimbangan kelas dapat menyebabkan model *machine learning* cenderung memprioritaskan kelas mayoritas dan mengabaikan kelas minoritas, yang berpotensi menghasilkan kinerja klasifikasi yang tidak seimbang atau bias (Saini *et al.*, 2023). Kemudian, untuk alur penelitian dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

2.1 Mengumpulkan Dataset

Dalam penelitian ini, digunakan dataset yang berasal dari *Facial Expression Recognition 2013 (FER-2013)* (Goodfellow *et al.*, 2015). Dataset tersebut memiliki tujuh kategori dan terdiri dari 28.709 gambar yang telah disiapkan untuk proses pelatihan. Dari jumlah total gambar yang tersedia, penelitian ini memilih menggunakan 2000 gambar saja, dengan membagi jumlah tersebut secara proporsional sebanyak 10% untuk setiap kategori ekspresi.

2.2 Melabeli Dataset

Pelabelan atau anotasi merupakan tahap penting dalam pengembangan model deteksi objek. Anotasi sangat penting dalam deteksi objek karena merupakan langkah kunci dalam proses *machine learning* untuk memahami dan mengenali objek dalam gambar. Dalam deteksi objek, anotasi merujuk pada proses menandai lokasi dan kelas objek dalam gambar dengan menggunakan *bounding box* atau metode lainnya (Ma *et al.*, 2022). Dalam konteks ini, penekanan utama pada pelabelan adalah untuk menunjukkan lokasi wajah. Tujuannya adalah memastikan bahwa saat sistem deteksi dijalankan, fokusnya terarah pada identifikasi wajah, tanpa teralihkan atau membuang waktu untuk mendeteksi objek atau area lain yang tidak relevan dalam citra.

2.3 Membagi Dataset

Dataset awal diproses dengan membaginya secara proporsional menjadi data *training* dan data *testing* dengan komposisi 90% dan 10%. Sebanyak 90% atau 1790 gambar dialokasikan sebagai data *training* untuk melatih model, sementara 10% sisanya, yaitu 210 gambar, dijadikan data *testing* untuk menguji kinerja model dalam mengenali ekspresi wajah.

2.4 Hyperparameter Setting

Penelitian ini menggunakan tiga *hyperparameter* utama, yakni *batch size* (ukuran kelompok data) 32, jumlah *epoch* 300, dan resolusi gambar 320x320. *Batch size* adalah parameter penting dalam *deep learning*. Berbagai *batch size* yang dipilih dapat menghasilkan akurasi pengujian dan pelatihan yang berbeda serta waktu proses yang berbeda pula. Oleh karena itu, saat melatih model, orang dapat memilih *batch size* 32 terlebih dahulu dan mengurangnya untuk meningkatkan akurasi atau meningkatkannya untuk efisiensi (Lin, 2022). Kemudian, semakin tinggi resolusi gambar, semakin baik deteksi objeknya, tetapi juga semakin besar biaya komputasinya dan persyaratan penyimpanan (Hao et al., 2022). Oleh karena itu, resolusi gambar yang dipilih adalah 320x320.

2.5 Melatih Model

Setelah mendapatkan anotasi kelas objek dan membagi data, langkah selanjutnya adalah melatih komputer menggunakan model YOLOv8n, YOLOv8s, YOLOv8m, dan YOLOv8l. Proses pelatihan ini melibatkan beberapa tahap penting. Pertama, *batch size* ditetapkan pada 32 untuk memastikan efisiensi dalam mengelola data pada setiap iterasi pelatihan. Selanjutnya, jumlah *epoch* yang digunakan adalah 300, yang menandakan berapa kali seluruh dataset digunakan dalam proses pelatihan. Resolusi gambar yang dipilih untuk pelatihan adalah 320x320, yang memungkinkan model untuk memahami detail-detail kecil dalam gambar. Melatih model dilakukan dengan tujuan agar komputer dapat mengenali pola atau karakteristik setiap kelas dalam gambar, memungkinkan pembuatan keputusan atau prediksi.

2.6 Menguji Model

Model diuji dengan 10% dari total dataset yang telah dipisahkan sebelumnya atau sebanyak 210 gambar. Model yang diuji melibatkan YOLOv8n, YOLOv8s, YOLOv8m, dan YOLOv8l. Metrik yang dievaluasi meliputi presisi, *recall*, mAP, dan waktu inferensi. Presisi yang tinggi menunjukkan bahwa hasil deteksi positif sesuai dengan prediksi positif secara akurat (Hasan et al., 2023). *Recall* merupakan tingkat keberhasilan sistem dalam mendapatkan kembali informasi yang telah diberikan (Hendriyana et al., 2020). *Recall* yang tinggi mengindikasikan bahwa kelas diidentifikasi dengan akurasi yang baik. mAP adalah metrik yang mengukur kemampuan sebuah sistem dalam mengenali objek atau entitas tertentu dengan tingkat akurasi yang tinggi. Pengujian bertujuan untuk menilai kemampuan model dalam mendeteksi ekspresi dan menerjemahkannya ke dalam kategori antusias atau tidak antusias.

3. Hasil dan Pembahasan

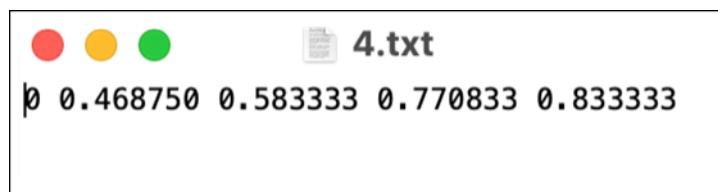
3.1 Hasil Pengolahan Dataset

Proses pelabelan dilaksanakan secara rinci dengan memisahkan data ke dalam dua kelas yang berbeda, yaitu "antusias" dan "tidak antusias," di mana angka 0 mengindikasikan "tidak antusias," sementara angka 1 mengindikasikan "antusias." Pelabelan ini melibatkan partisipasi lima individu dengan latar belakang dan jenis kelamin yang beragam. Hal ini bertujuan untuk memastikan bahwa definisi kelas antusias dan tidak antusias tidak hanya bergantung pada sudut pandang penulis, sehingga diharapkan dapat lebih objektif dan akurat. Individu yang terlibat dalam proses pelabelan termasuk dua dosen dan tiga mahasiswa. Hasil pelabelan oleh lima partisipan disajikan dalam bentuk tabel dan dapat dilihat pada Tabel 1.

Tabel 1. Hasil Pelabelan oleh Lima Partisipan

Ekspresi	Label	Jumlah
Tidak antusias	0	987
Antusias	1	1013

Kemudian hasil pelabelan diolah menggunakan *tool/LabelImg*, mencakup koordinat dan ukuran objek. Gambar 2 menunjukkan hasil pelabelan dengan *LabelImg*. Data hasil pelabelan mencakup id kelas objek, pusat titik (x, y), lebar, dan tinggi kotak pembatas label.



Gambar 2. Hasil Pelabelan dengan LabelImg

Setelah dilabeli, dataset kemudian dibagi menjadi dua bagian dengan komposisi 90% data *training* atau sejumlah 1790 gambar dan 10% data *testing* atau sejumlah 210 gambar.

3.2 Hasil Pelatihan

Hasil pelatihan memuat hasil dari pelatihan YOLOv8n, YOLOv8s, YOLOv8m, dan YOLOv8l dengan menggunakan *epoch*, *batch size*, dan *image size* yang sama, yaitu 300 *epoch*, 32 *batch size*, dan *image size* 320x320. Hasil pelatihan YOLOv8n menunjukkan peningkatan signifikan dalam kecocokannya secara keseluruhan. Nilai loss pada grafik *train/box_loss* dan *train/cls_loss* terus menurun seiring dengan bertambahnya *epoch*. *Train/box_loss* turun dari 1,4 ke 0,9, sementara *train/cls_loss* dari 2,0 ke 0,75. Hal ini mencerminkan peningkatan akurasi model dalam memprediksi posisi dan kelas objek dalam gambar. Selanjutnya, penurunan nilai loss pada data validasi juga terlihat seiring bertambahnya *epoch*. *Val/box_loss* turun dari 1,7 ke 1,1, dan *val/cls_loss* dari 2,0 ke 0,5. Ini menunjukkan bahwa model YOLOv8n efektif tidak hanya pada data pelatihan, tetapi juga pada data yang belum pernah dilihat sebelumnya.

Proses pelatihan YOLOv8s menunjukkan penurunan nilai loss yang signifikan. *Train/box_loss* turun dari 1,4 ke 1,0, dan *train/cls_loss* dari 1,6 ke 0,8. Model YOLOv8s mencapai akurasi optimal pada *epoch* ke-50. Model YOLOv8m memerlukan waktu pelatihan lebih lama, tetapi menghasilkan akurasi yang lebih tinggi. *Train/box_loss* dan *train/cls_loss* menurun dari 1,4 ke 0,8 dan 1,8 ke 0,6. Meskipun pelatihan berhenti setelah lebih dari 100 *epoch*, YOLOv8m memberikan hasil yang akurat pada data pelatihan dan validasi. Hasil pelatihan YOLOv8l menunjukkan peningkatan yang signifikan dalam akurasi. *Train/box_loss* turun dari 1,4 ke 0,8, dan *train/cls_loss* dari 1,75 ke 0,75. Grafik *val/box_loss* menunjukkan fluktuasi, tetapi model YOLOv8l tetap memberikan hasil yang dapat diterima pada data validasi.

3.3 Hasil Pengujian

Setelah melalui tahap pelatihan, langkah selanjutnya adalah pengujian. Pengujian dilakukan dengan menggunakan 10% data yang sudah dibagi di langkah sebelumnya. Pembahasan hasil pengujian masing-masing ukuran akan berfokus pada metrik presisi, *recall*, dan mAP (*mean Average Precision*). Presisi relevan untuk mengukur sejauh mana model memberikan respon yang tepat terhadap deteksi objek, khususnya dalam hal deteksi ekspresi wajah. *Recall* menjadi krusial untuk mengevaluasi kemampuan model mendeteksi objek sebanyak mungkin yang seharusnya terdeteksi. mAP digunakan sebagai metrik komprehensif yang menggabungkan presisi dan *recall*, memberikan ukuran akurasi deteksi objek, terutama dalam tugas *multiple object detection* seperti identifikasi ekspresi wajah. Terakhir, evaluasi *inference time* menjadi parameter penting untuk mengukur efisiensi model dalam mengenali

ekspresi wajah dengan cepat dan efisien. Dari hasil penelitian yang telah dilakukan, diperoleh data mengenai performa YOLOv8 untuk setiap ukuran yang dapat diamati pada Tabel .

Tabel 2. Hasil Pengujian Model YOLOv8

Model YOLOv8	Presisi	Recall	mAP	Inference Time
YOLOv8n	0,888	0,914	0,953	69,2 ms
YOLOv8s	0,882	0,925	0,952	224,7 ms
YOLOv8m	0,911	0,878	0,939	602,5 ms
YOLOv8l	0,855	0,896	0,941	1302,8 ms

Dari hasil penelitian, diperoleh data mengenai performa YOLOv8 untuk setiap ukuran yang tercantum dalam Tabel . Model YOLOv8 ukuran n (nano) mencapai nilai mAP tertinggi sebesar 0,953, menandakan kecocokannya dalam konteks pembelajaran daring. YOLOv8 memberikan waktu inferensi yang sangat singkat, memungkinkan deteksi objek secara real-time, dengan nilai mAP yang tinggi, sehingga memungkinkan deteksi antusiasme siswa dengan lebih akurat. Hal ini dapat membantu pengajar untuk lebih memahami dan merespon secara tepat terhadap emosi siswa selama proses pembelajaran daring berlangsung. Selain itu, hasil pengujian menunjukkan bahwa semakin besar ukuran model, nilai mAP cenderung semakin kecil. Kemungkinan hal ini disebabkan oleh penggunaan 80 kelas pada pengujian YOLOv8, sementara penelitian ini hanya menggunakan 2 kelas, yang mungkin menyebabkan *overfitting*.

4. Kesimpulan

Hasil penelitian ini menghasilkan beberapa kesimpulan signifikan. Pertama, implementasi YOLOv8 untuk deteksi antusiasme siswa dalam pembelajaran daring berhasil dilakukan, dengan pembagian kelas “antusias” dan “tidak antusias” pada dua kelas. Kedua, model yang diuji, termasuk YOLOv8n, YOLOv8s, YOLOv8m, dan YOLOv8l, menunjukkan hasil terbaik oleh YOLOv8n dengan mAP@ sebesar 95,3%, presisi 0,888, dan *recall* 0,914. Ketiga, YOLOv8n menonjol dengan *inference time* tercepat, hanya 69,2ms, memungkinkan deteksi objek secara *real-time* dalam kurun waktu kurang dari 1 detik. Terakhir, meski YOLOv8l diklaim memiliki akurasi lebih tinggi, hasil penelitian menunjukkan mAP@ sebesar 94,1% dengan waktu inferensi yang lebih lama, yaitu 1302,8ms. Meskipun YOLOv8l diklaim memiliki tingkat akurasi yang lebih tinggi daripada YOLOv8n, tetapi dalam penelitian ini, YOLOv8n malah menghasilkan nilai mAP yang lebih tinggi. Selain memberikan gambaran tentang hasil penelitian, temuan penelitian ini juga dapat dapat membawa dampak positif, terutama pada pendidikan daring secara keseluruhan. Dengan adanya temuan ini, pengajar dapat memiliki alat yang efektif untuk memantau tingkat antusiasme siswa dalam pembelajaran daring. Hal ini tidak hanya dapat meningkatkan efisiensi pembelajaran, tetapi juga membuka peluang untuk penyesuaian strategi pembelajaran yang lebih personal dan responsif terhadap kebutuhan individual siswa.

5. Ucapan Terima Kasih

Terima kasih kepada pihak-pihak yang sudah terlibat dalam penelitian ini, terutama partisipan yang turut serta membantu penulis dalam mengelaskan 2000 gambar ke dalam kelas antusias dan tidak antusias. Terima kasih juga penulis sampaikan kepada teman-teman penulis yang selalu memberikan dorongan semangat sehingga penulis dapat menyelesaikan artikel ini.

6. Daftar Pustaka

- Angraini, N. A., & Fadillah, N. (2019). Analisis Deteksi Emosi Manusia dari Suara Percakapan Menggunakan Matlab dengan Metode KNN. *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, 3(2), 176-179.
- AZIZI, F. N. (2021). Deteksi Emosi Menggunakan Citra Ekspresi Wajah Secara Otomatis.
- Darwin, C., Ekman, P., & Prodger, P. (2002). The expression of the emotions in man and animals: Oxford University Press. USA, (1872 reprinted 2002).
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2), 124.
- Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., ... & Bengio, Y. (2013). Challenges in representation learning: A report on three machine learning contests. In *Neural Information Processing: 20th International Conference, ICONIP 2013, Daegu, Korea, November 3-7, 2013. Proceedings, Part III 20* (pp. 117-124). Springer berlin heidelberg.
- Gupta, S., Kumar, P., & Tekchandani, R. K. (2023). Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. *Multimedia Tools and Applications*, 82(8), 11365-11394. DOI: <https://doi.org/10.1109/IROS55552.2023.10342139>.
- Hasan, M. A., & Lazem, A. H. (2023). Facial Human Emotion Recognition by Using YOLO Faces Detection Algorithm. *CENTRAL ASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES*, 4(10), 1-11.
- Hendriyana, H., & Maulana, Y. H. (2020). Identification of types of wood using convolutional neural network with MobileNet architecture. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 4(1), 70-76. DOI: <https://doi.org/10.29207/resti.v4i1.1445>.
- Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). A Review of Yolo algorithm developments. *Procedia computer science*, 199, 1066-1073. DOI: <https://doi.org/10.1016/j.procs.2022.01.135>.
- Kovalchuk, Y., Budini, E., Cook, R. M., & Walsh, A. (2022). Investigating the relationship between facial mimicry and empathy. *Behavioral Sciences*, 12(8), 250.
- Lin, R. (2022, February). Analysis on the selection of the appropriate batch size in CNN neural network. In *2022 International Conference on Machine Learning and Knowledge Engineering (MLKE)* (pp. 106-109). IEEE. DOI: <https://doi.org/10.1109/MLKE55170.2022.00026>.
- Ma, J., Ushiku, Y., & Sagara, M. (2022). The effect of improving annotation quality on object detection datasets: A preliminary study. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4850-4859).
- Ohoirat, A. L. E. PENGEMBANGAN INSTRUMEN ANALISIS VIDEO UNTUK MENENTUKAN PROFIL KARAKTER CALON GURU DALAM MEMBUKA PELAJARAN DITINJAU DARI EKSPRESI WAJAH.

Prasetyawan, D. (2020). Penentuan Emosi pada Video dengan Convolutional Neural Network. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 5(1), 23-35. DOI: <https://doi.org/10.14421/jiska.2020.51-04>.

Saini, M., & Susan, S. (2023). Tackling class imbalance in computer vision: a contemporary review. *Artificial Intelligence Review*, 56(Suppl 1), 1279-1335.

Terven, J., & Cordova-Esparza, D. (2023). A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *arXiv preprint arXiv:2304.00501*.