

COMPARISON OF ARCFACE AND DLIB PERFORMANCE IN FACE RECOGNITION WITH DETECTION USING YOLOV8

PERBANDINGAN KINERJA ARCFACE DAN DLIB DALAM PENGENALAN WAJAH DENGAN DETEKSI MENGUNAKAN YOLOV8

Berliana Wahyu Nurlita¹, Sri Winarno², Adhitya Nugraha³, Almas Najiib Imam Muttaqin⁴, Yasmin Zarifa⁵, Pramesya Mutia Salsabila⁶, Ghina Fairuz Mumtaz⁷,
^{1, 2, 3, 4, 5, 6, 7} Universitas Dian Nuswantoro,
Jl. Imam Bonjol No.207, Semarang, Indonesia
*Email: litaber2706@gmail.com¹, sri.winarno@dsn.dinus.ac.id², adhitya@dsn.dinus.ac.id³,
almasnajiib27@gmail.com⁴, yasminezarifaa@gmail.com⁵, pramesyamutiasalsabila@gmail.com⁶,
ghinafairuz2321@gmail.com⁷*

Abstract - This study compares the performance of ArcFace and Dlib models in face recognition with YOLOv8 used for face detection on a limited dataset. The evaluation used metrics such as accuracy, F1-score, recall, and precision. ArcFace, which employs the Additive Angular Margin Loss method, demonstrated superior performance with the highest accuracy of 0.90, precision of 0.90, recall of 1.00, and an F1-score of 0.95. Meanwhile, Dlib achieved an accuracy of 0.57, precision of 0.57, recall of 1.00, and an F1-score of 0.73. The aim of the study was to find the best model in terms of accuracy. ArcFace proved to be more accurate and suitable for applications requiring high reliability, such as advanced security systems, identity verification, and research that demands high precision in face recognition. Dlib, although less accurate, offers speed and simplicity, making it suitable for rapid prototyping and lightweight applications with limited resources. The results indicate that ArcFace outperforms in face recognition on limited datasets, while Dlib is more appropriate for simple applications requiring lightweight computation. This study provides guidance for developers in selecting the appropriate face recognition model to meet specific needs in both industry and research.

Keywords - Face Recognition, YOLOv8, ArcFace, Dlib, Image Processing

Abstrak - Penelitian ini membandingkan kinerja model ArcFace dan Dlib dalam pengenalan wajah dengan bantuan deteksi wajah menggunakan YOLOv8 pada dataset terbatas. Evaluasi dilakukan dengan metrik akurasi, F1-score, recall, dan precision. ArcFace, yang menggunakan metode Additive Angular Margin Loss, menunjukkan performa yang lebih unggul dengan akurasi tertinggi 0,90, precision 0,90, recall 1,00, dan F1-score 0,95. Sementara itu, Dlib mencapai akurasi 0,57, precision 0,57, recall 1,00, dan F1-score 0,73. Tujuan penelitian ini adalah menemukan model terbaik dari segi akurasi. ArcFace terbukti lebih akurat dan cocok untuk aplikasi yang memerlukan keandalan tinggi, seperti sistem keamanan, verifikasi identitas, dan penelitian yang membutuhkan akurasi tinggi. Dlib, meskipun akurasinya lebih rendah, menawarkan kecepatan dan kesederhanaan yang dapat digunakan untuk prototipe cepat dan aplikasi ringan dengan keterbatasan sumber daya. Kesimpulan menunjukkan bahwa ArcFace unggul dalam pengenalan wajah pada dataset terbatas, sedangkan Dlib lebih cocok untuk aplikasi sederhana dengan kebutuhan komputasi ringan. Hasil penelitian ini memberikan panduan bagi pengembang dalam memilih model pengenalan wajah yang sesuai dengan kebutuhan spesifik di industri maupun penelitian.

Kata Kunci - Pengenalan Wajah, YOLOv8, ArcFace, Dlib, Pemrosesan Gambar

I. PENDAHULUAN

Pengenalan wajah merupakan aspek krusial dalam teknologi pengolahan citra dan kecerdasan buatan, dengan aplikasi luas dalam keamanan, verifikasi identitas, dan interaksi manusia dengan mesin[1]–[3]. Teknologi ini memanfaatkan algoritma untuk mengenali dan mengidentifikasi individu berdasarkan ciri-ciri wajah mereka. Dengan kemajuan dalam teknik pembelajaran mendalam, akurasi dan efisiensi pengenalan wajah terus meningkat, memungkinkan implementasi yang lebih efektif di berbagai sektor[4], [5]. Meskipun teknologi pengenalan wajah telah berkembang pesat, tantangan muncul saat menghadapi dataset terbatas, terutama yang hanya menyediakan satu foto per subjek dan dikumpulkan secara manual. Keterbatasan ini dapat memengaruhi kemampuan model dalam mengenali wajah dengan akurat, karena banyak algoritma pembelajaran mendalam bergantung pada jumlah dan keragaman data untuk pelatihan. Tantangan ini semakin kompleks dengan adanya faktor variabilitas kondisi pencahayaan, ekspresi wajah, dan sudut pengambilan gambar yang memengaruhi kinerja model[6]. Untuk mengatasi tantangan ini, penelitian ini menerapkan teknik augmentasi data dan strategi evaluasi berbasis metrik komprehensif guna mengurangi dampak keterbatasan dataset[7].

Salah satu solusi untuk mengatasi masalah dataset terbatas adalah menggunakan teknik deteksi wajah yang efektif, seperti YOLOv8 (You Only Look Once versi 8). Penelitian sebelumnya menunjukkan bahwa YOLOv8 memiliki keunggulan dalam kecepatan, akurasi, dan efisiensi komputasi, memungkinkan deteksi wajah secara real-time dalam berbagai kondisi[8]. YOLOv8 dipilih karena algoritma ini dapat mendeteksi objek dengan cepat dan akurat, membuatnya ideal untuk aplikasi yang memerlukan proses deteksi secara real-time[9], [10]. Keunggulan ini memberikan landasan yang solid untuk mendeteksi wajah dengan efisiensi tinggi, yang kemudian dapat diproses lebih lanjut oleh model pengenalan wajah.

ArcFace adalah algoritma pengenalan wajah yang menggunakan Additive Angular Margin Loss untuk menghasilkan embedding wajah yang lebih terpisah dan robust[11]. Mekanisme ini meningkatkan kemampuan model dalam membedakan individu dengan lebih akurat, bahkan dalam kondisi pencahayaan dan ekspresi yang bervariasi. Hal ini membuat ArcFace sangat cocok untuk aplikasi yang memerlukan tingkat presisi tinggi dalam pengenalan wajah. Di sisi lain, Dlib menggunakan pendekatan histogram of oriented gradients (HOG) dan support vector machine (SVM) untuk mendeteksi dan mengenali wajah[12]. Meskipun lebih sederhana dan cepat dalam implementasi, Dlib memiliki keterbatasan dalam hal akurasi ketika berhadapan dengan dataset yang kurang bervariasi.

Penelitian ini juga memperhitungkan batasan-batasan yang dapat memengaruhi hasil, seperti keterbatasan dataset, variabilitas dalam kondisi perekaman data (cahaya, sudut pandang, ekspresi wajah), dan kemungkinan bias dalam model. Batasan dataset diatasi dengan augmentasi data untuk meningkatkan variasi, sementara pengukuran kinerja dilakukan dengan metrik komprehensif seperti akurasi, precision, recall, dan F1-score guna meminimalkan bias dalam evaluasi. ArcFace dan Dlib memiliki potensi penggunaan yang signifikan dalam industri dan penelitian. ArcFace, dengan representasi wajah yang robust melalui pendekatan fungsi loss berbasis margin, ideal untuk aplikasi yang memerlukan keandalan tinggi, seperti sistem keamanan canggih, verifikasi identitas, dan penelitian yang membutuhkan presisi tinggi. Di sisi lain, Dlib, dengan kelebihan kesederhanaan dan kecepatan, cocok untuk prototipe cepat dan aplikasi ringan dengan keterbatasan sumber daya. Dengan membandingkan kedua model ini, penelitian ini bertujuan untuk menentukan mana yang lebih

efektif dalam menghadapi tantangan dataset terbatas dan berkontribusi terhadap pengembangan teknologi pengenalan wajah yang lebih robust dan akurat.

Pertanyaan penelitian yang ingin dijawab dalam studi ini adalah Apakah ArcFace lebih unggul dibandingkan Dlib dalam konteks akurasi dan efisiensi pada pengenalan wajah dengan deteksi menggunakan YOLOv8 ketika diaplikasikan pada dataset terbatas yang hanya terdiri dari satu foto per subjek, dan bagaimana perbandingan keandalan serta kecepatan masing-masing model dalam mendukung aplikasi pengenalan wajah di industri maupun penelitian. Penelitian ini bertujuan untuk membandingkan kinerja ArcFace dan Dlib dalam pengenalan wajah pada dataset satu foto per subjek dengan deteksi wajah menggunakan YOLOv8. Melalui penelitian ini, diharapkan dapat memberikan pemahaman yang lebih baik serta dapat menjadi acuan untuk penelitian selanjutnya mengenai efektivitas kedua model dalam konteks dataset terbatas dan memberikan kontribusi pada pengembangan teknologi.

II. SIGNIFIKANSI STUDI

A. Studi Literatur

Berikut ini merupakan beberapa penelitian terdahulu yang berkaitan dengan pengenalan wajah menggunakan ArcFace dan Dlib, serta pendekatan lain yang relevan.

TABEL I. STUDI LITERATUR

| No. | Judul Penelitian | Tahun | Objektif | Model | Hasil | Kekurangan |
|-----|---|-------|---|--|---|--|
| 1 | AdvHat: Serangan Dunia Nyata pada Sistem ArcFace Face ID | 2019 | Mengembangkan serangan adversarial praktis pada sistem ArcFace Face ID menggunakan stiker di topi. | ArcFace and ms1m-refine-v2, LResNet100E-IR | Menunjukkan efektivitas serangan dalam berbagai kondisi, dengan efek yang dapat ditransfer ke model lain. | Bergantung pada kondisi pengambilan gambar dan konfigurasi model tertentu. |
| 2 | Bergantung pada Kondisi Pengambilan Gambar dan Konfigurasi Model Tertentu | 2019 | Memperkenalkan fungsi kehilangan ArcFace untuk meningkatkan pengenalan wajah dengan menambahkan margin sudut. | ArcFace dengan arsitektur DCNN | Mengungguli model canggih pada berbagai tolok ukur, dengan implementasi yang sederhana. | Membutuhkan data yang signifikan untuk pelatihan. |
| 3 | Pengenalan Wajah Tertutup Menggunakan ArcFace dan Pembelajaran Ensemble | 2022 | Mengatasi tantangan pengenalan wajah tertutup menggunakan pembelajaran ensemble dengan ArcFace. | MobileNetV2, DenseNet201, ResNet50V2, VGG16 dengan ArcFace | Mencapai akurasi 93.65% pada dataset wajah tertutup dan tidak tertutup. | Performa mungkin menurun dalam skenario dunia nyata dengan pencahayaan buruk atau masker yang kompleks.. |
| 4 | Desain Sistem Deteksi Kantuk Real-time Menggunakan | 2020 | Mengembangkan sistem deteksi kantuk real-time menggunakan | Dlib dengan OpenCV untuk deteksi landmark | Mencapai akurasi pengenalan 96.71% pada | Terpengaruh oleh halangan seperti kaca dan |

| | Dlib | | Dlib. | wajah. | dataset uji. | membutuhkan penyesuaian untuk kondisi beragam.. |
|---|---|------|---|--|---|---|
| 5 | Pendekatan Pengenalan Wajah Menggunakan Dlib dan K-NN | 2023 | Menerapkan pengenalan wajah dengan Dlib dan K-NN untuk sistem kontrol akses. | Dlib, classifier k-NN, classifier Haar cascade | Mencapai akurasi 90% pada dataset ORL_Face. | Performa terpengaruh oleh variasi pencahayaan dan resolusi gambar. |
| 6 | Implementasi Teknologi Pengakuan Wajah Pembelajaran Mendalam Dlib | 2020 | Meningkatkan deteksi dan pengenalan wajah dengan Dlib untuk skenario real-time. | Dlib dengan Python, algoritma ERT | Meningkatkan ketepatan pengenalan dibandingkan dengan OpenCV. | Membutuhkan pengaturan yang kompleks dan ketergantungan lingkungan. |

Berdasarkan Tabel 1, penelitian-penelitian sebelumnya menunjukkan keunggulan ArcFace dalam memaksimalkan pemisahan antar-kelas dan konsistensi dalam satu kelas melalui penggunaan margin sudut tambahan (Additive Angular Margin)[11], [13], [14]. ArcFace terbukti mampu mencapai tingkat akurasi yang tinggi pada berbagai dataset besar yang mengandung keragaman wajah, meskipun efektivitasnya pada dataset yang lebih kecil dan terbatas masih belum banyak dibahas. Dalam situasi dengan dataset terbatas, model sering menghadapi tantangan overfitting yang dapat mengurangi kinerja ArcFace. Oleh karena itu, penelitian lebih lanjut dibutuhkan untuk mengevaluasi apakah keunggulan margin sudut ini masih relevan saat diterapkan pada dataset yang lebih kecil.

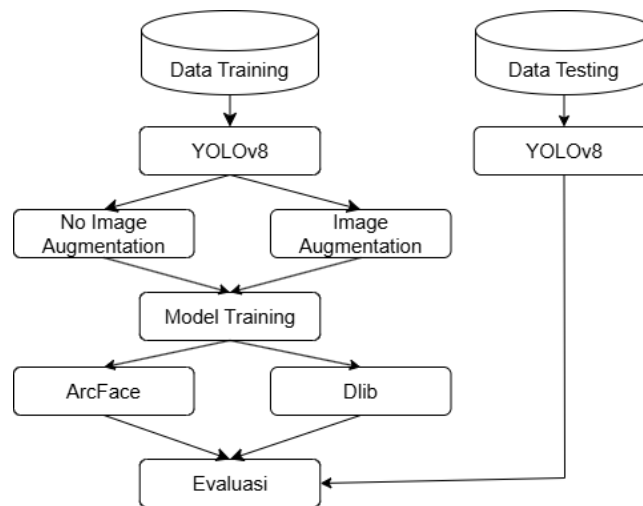
Di sisi lain, metode Dlib yang berbasis ekstraksi fitur sederhana, seperti Histogram of Oriented Gradients (HOG) dan algoritma k-Nearest Neighbors (k-NN), sering kali lebih tahan terhadap overfitting pada dataset kecil karena kompleksitas model yang lebih rendah[15]–[17]. Dalam beberapa penelitian, Dlib menunjukkan performa yang stabil bahkan ketika ukuran dataset terbatas. Namun, ketahanannya dalam menghasilkan akurasi tinggi pada dataset terbatas belum dibandingkan secara menyeluruh dengan ArcFace, yang dirancang untuk menghasilkan fitur yang lebih diskriminatif. Dengan membandingkan kedua metode ini dalam konteks dataset terbatas, penelitian dapat memberikan gambaran lebih jelas tentang model yang paling efisien untuk kondisi data yang tidak ideal atau terbatas. Selain itu, pada dataset terbatas, kualitas gambar yang sedikit juga menjadi tantangan. ArcFace dirancang untuk mengoptimalkan margin antar kelas sehingga fitur wajah lebih mudah dibedakan, tetapi metode ini mungkin tidak ideal pada dataset yang memiliki jumlah sampel terbatas per kelas. Sebaliknya, pendekatan Dlib yang lebih sederhana berpotensi memberikan hasil yang lebih stabil meski dalam kondisi keterbatasan sampel. Penelitian yang mengeksplorasi kinerja ArcFace dan Dlib pada dataset terbatas akan membantu memahami bagaimana kedua metode ini menangani tantangan yang umum dihadapi dalam aplikasi pengenalan wajah praktis.

Terakhir, perbandingan kinerja ArcFace dan Dlib pada dataset terbatas penting dalam konteks aplikasi pengenalan wajah yang sering kali tidak memiliki akses ke dataset besar, seperti dalam pengaturan institusi kecil atau aplikasi keamanan berbasis komunitas. Sementara ArcFace menawarkan kompleksitas dan presisi yang lebih tinggi, Dlib menyediakan solusi yang lebih efisien secara komputasi. Evaluasi yang mendalam mengenai bagaimana kedua

metode ini mengadaptasi performanya pada dataset terbatas akan sangat bermanfaat dalam memilih model pengenalan wajah yang paling sesuai, terutama untuk aplikasi dengan sumber daya dan akses data yang terbatas.

B. Metode Penelitian

Pada metode penelitian ini, setiap aspek proses penelitian akan dibahas secara rinci, mulai dari persiapan data hingga evaluasi akhir, sehingga dapat memberikan hasil yang dapat diandalkan untuk menentukan keunggulan antara ArcFace dan Dlib dalam pengenalan wajah. Gambar 1 menunjukkan alur proses penelitian :



Gambar 1. Alur Penelitian

Gambar 1 Merupakan proses dalam menjalankan penelitian yang terdiri dari 5 proses yakni:

1. Data Training dan Data Testing: Penelitian dimulai dengan pemisahan dataset menjadi dua bagian utama: data training dan data testing. Data training digunakan untuk melatih model, sementara data testing digunakan untuk mengevaluasi kinerja model setelah pelatihan selesai.
2. Proses Deteksi Wajah dengan YOLOv8: Data training dan data testing diproses menggunakan model YOLOv8 untuk mendeteksi wajah. YOLOv8 dipilih karena kemampuannya mendeteksi wajah dengan cepat dan akurat dalam berbagai kondisi, memastikan bahwa data yang digunakan untuk pelatihan dan pengujian berfokus pada area wajah yang relevan.
3. Augmentasi Gambar: Data training kemudian dibagi ke dalam dua alur—satu alur tanpa augmentasi gambar dan alur lainnya dengan augmentasi gambar. Augmentasi gambar dilakukan untuk meningkatkan variasi data, seperti rotasi, perubahan pencahayaan, dan penyesuaian lainnya, guna mengatasi keterbatasan data dan meningkatkan kemampuan generalisasi model.
4. Model Training: Setelah data disiapkan melalui proses augmentasi dan deteksi wajah, tahap berikutnya adalah pelatihan model. Pada tahap ini, dua model yang berbeda, yaitu ArcFace dan Dlib, dilatih menggunakan data training yang telah diproses. ArcFace dikenal dengan pendekatan loss berbasis margin yang menghasilkan embedding wajah yang lebih robust, sedangkan Dlib menggunakan metode yang lebih sederhana seperti HOG dan SVM.

5. Evaluasi Model: Setelah model ArcFace dan Dlib selesai dilatih, keduanya diuji dengan data testing yang telah melalui deteksi wajah dengan YOLOv8. Evaluasi ini melibatkan metrik akurasi, precision, recall, dan F1-score untuk membandingkan performa kedua model. Tujuan evaluasi ini adalah menentukan model mana yang lebih unggul dalam hal akurasi dan efisiensi, terutama dalam konteks dataset terbatas.

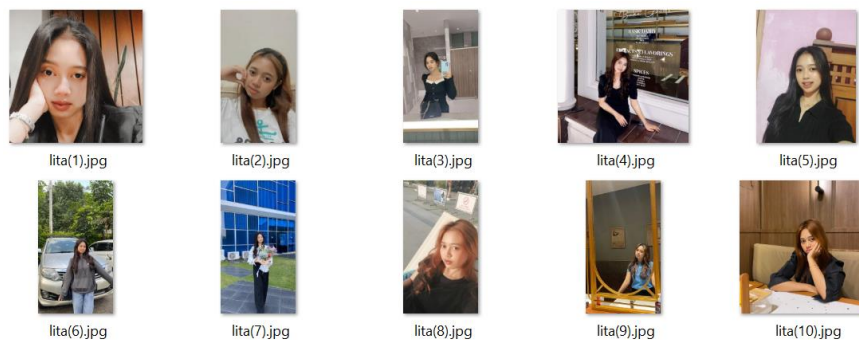
C. Dataset

Dalam penelitian ini, dataset berfungsi sebagai kumpulan data yang akan dimanfaatkan dalam proses pelatihan dan pengujian model. Data tersebut dipisahkan menjadi dua bagian yaitu Dataset training, yang digunakan untuk melatih model, dan Dataset Testing, yang berperan dalam mengevaluasi performa model dengan menggunakan data baru yang belum pernah terlibat dalam proses pelatihan.



Gambar 2. Data Training

Data training digunakan untuk melatih model, di mana gambar wajah diproses terlebih dahulu menggunakan YOLOv8 untuk mendeteksi area wajah yang relevan. Data training mencakup 10 foto formal mahasiswa Universitas Dian Nuswantoro yang terbagi menjadi 10 kelas. Data ini kemudian dibagi menjadi dua alur satu alur tanpa augmentasi gambar, yang digunakan apa adanya, dan alur lain dengan augmentasi gambar, di mana data dimodifikasi untuk meningkatkan variasi melalui rotasi, flipping, dan penyesuaian pencahayaan. Tujuan dari data training ini adalah membantu model membangun kemampuan mengenali fitur-fitur wajah secara akurat, baik dalam kondisi data yang standar maupun yang bervariasi, guna memastikan model lebih robust dan mampu mengatasi kondisi nyata yang berbeda.



Gambar 3. Data Testing

Data Testing digunakan untuk mengevaluasi kinerja model setelah proses pelatihan. Seperti data training, data testing juga diproses dengan YOLOv8 untuk memastikan fokus pada area wajah. Data testing terdiri dari 100 foto, dengan 10 gambar non-formal dari masing-masing

kelas. Sebagai contoh, data testing pada gambar 3 menunjukkan sepuluh foto seseorang dengan label "lita" dengan nomor dari 1 hingga 10. Data ini tidak melalui augmentasi agar evaluasi dapat mengukur performa model dalam kondisi yang lebih realistis. Penggunaan data testing bertujuan untuk memverifikasi kemampuan model dalam mengenali wajah yang belum pernah dilihat sebelumnya, sehingga menguji sejauh mana model dapat menggeneralisasi pengetahuannya. Evaluasi ini membantu memastikan bahwa model tidak hanya menghafal data training tetapi juga dapat mengaplikasikan pengetahuan tersebut dengan akurasi tinggi dalam situasi nyata.

D. YOLOv8

Proses Deteksi Wajah dengan YOLOv8 merupakan langkah penting dalam alur penelitian ini untuk memastikan bahwa data training dan data testing berfokus pada area wajah yang relevan[18]. Proses dimulai dengan memasukkan gambar-gambar dari dataset ke dalam pipeline deteksi YOLOv8. YOLOv8 dipilih karena kemampuannya yang andal dalam mendeteksi objek, termasuk wajah, dengan cepat dan akurat dalam berbagai kondisi. Model ini memindai setiap gambar dan menandai area wajah dengan bounding box, memastikan bahwa hanya area wajah yang diinputkan ke dalam model pelatihan ArcFace dan Dlib. Gambar yang dipotong menjadi 64x64 piksel kemudian dikonversi ke format grayscale, yang hanya memerlukan satu saluran warna dibandingkan gambar berwarna (RGB) dengan tiga saluran. Format grayscale dipilih karena elemen visual seperti tepian, kontur, dan bentuk wajah lebih berperan dalam deteksi wajah dibandingkan informasi warna[19]. Tahap ini membantu membuang elemen latar belakang yang tidak dibutuhkan, sehingga kualitas data yang masuk tetap terjaga. Setelah proses deteksi, gambar dari data training dipisahkan menjadi alur tanpa augmentasi dan alur dengan augmentasi untuk meningkatkan variasi data. Di sisi lain, data testing juga diproses dengan YOLOv8 untuk menjaga konsistensi hasil antara pelatihan dan pengujian, tanpa melalui proses augmentasi. Proses ini memastikan bahwa model dilatih dan diuji dengan gambar wajah yang seragam dan fokus, meningkatkan keakuratan evaluasi dan performa model dalam kondisi nyata.

E. Augmentasi Gambar

Pada proses augmentasi gambar mencakup berbagai teknik untuk meningkatkan variasi data training agar model pengenalan wajah lebih robust dan mampu mengenali wajah dalam kondisi nyata[20]. Proses augmentasi membantu mengatasi keterbatasan jumlah data training yang hanya terdiri dari 10 foto formal per kelas. Dalam penelitian ini parameter data augmentasi yang digunakan diberikan seperti pada Tabe 2.

TABEL II. AUGMENTASI GAMBAR

| Tenik Augmentasi | | |
|---------------------|---|--|
| Jenis | Nilai | Efek |
| Rotasi | Sudut: -15 hingga 15 | Mengubah orientasi gambar |
| Skala | Faktor Skala : 0.9-1.2 | Mengubah ukuran objek dalam gambar, mensimulasikan variasi jarak subjek terhadap kamera. |
| Kemiringan | Kemiringan (shear) : -15 hingga 15 derajat. | Gambar akan dimiringkan, mensimulasikan distorsi perspektif (sudut pandang). |
| Flip Horizontal | Diterapkan | Gambar akan dimiringkan, mensimulasikan distorsi perspektif. |
| Blur Gaussian | Sigma: 0 hingga 0.5. | Memberikan efek blur untuk menyimulasikan kondisi gambar yang tidak fokus. |
| Perubahan Kecerahan | Kecerahan: 50% hingga 150%. | Gambar dapat dibuat lebih gelap atau lebih terang. |

| | | |
|-----------|--|--|
| Resize | Lebar digandakan (2.0), tinggi tetap (1.0). | Membuat gambar lebih lebar, mensimulasikan perubahan aspek rasio. |
| Penajaman | Alpha: 0.0 hingga 1.0 dan lightness: 0.75 hingga 1.5. | Gambar dipertajam untuk menonjolkan detail, meningkatkan tepi dan tekstur. |

Pada tabel 2 menunjukkan parameter yang digunakan untuk proses data augmentasi. Teknik augmentasi pertama adalah rotasi, di mana gambar dapat diputar dengan sudut antara -15 hingga 15 derajat untuk mengubah orientasi gambar, membantu model mengenali wajah yang sedikit miring. Skala juga diterapkan dengan faktor antara 0.9 hingga 1.2 untuk mengubah ukuran objek dalam gambar, mensimulasikan variasi jarak antara subjek dan kamera. Selanjutnya, teknik kemiringan (shear) dilakukan dengan derajat kemiringan antara -15 hingga 15 derajat untuk menciptakan distorsi perspektif, yang penting agar model dapat mengenali wajah dalam sudut pandang yang lebih ekstrem. Flip horizontal diterapkan untuk membalik gambar, menambah variasi posisi wajah yang umum dalam kondisi nyata. Teknik blur Gaussian digunakan dengan parameter sigma antara 0 hingga 0.5 untuk memberikan efek buram, mensimulasikan kondisi di mana gambar mungkin tidak sepenuhnya fokus.

Perubahan kecerahan diatur dalam rentang 50% hingga 150% untuk menciptakan gambar yang lebih gelap atau lebih terang, membantu model mengenali wajah di berbagai kondisi pencahayaan. Resize dilakukan dengan memperbesar lebar hingga dua kali lipat sementara tinggi tetap sama, mensimulasikan perubahan aspek rasio yang mungkin terjadi. Terakhir, penajaman dengan parameter alpha antara 0.0 hingga 1.0 dan lightness antara 0.75 hingga 1.5 digunakan untuk menonjolkan detail dalam gambar, meningkatkan tepi dan tekstur sehingga fitur wajah lebih jelas. Semua teknik ini bertujuan untuk memperkaya variasi data training, memastikan model mampu mengatasi tantangan nyata dengan lebih baik.



Gambar 4. Hasil Augmentasi Gambar

Pada gambar 4 menampilkan serangkaian variasi foto wajah yang dihasilkan dari parameter proses augmentasi untuk meningkatkan keragaman data training dan membuat model pengenalan wajah lebih adaptif. Beberapa gambar tampak diputar ke kiri atau kanan, yang bertujuan untuk mensimulasikan posisi wajah yang miring, sehingga model dapat mengenali wajah dari sudut yang berbeda. Ada pula gambar dengan variasi kecerahan, baik yang lebih gelap maupun lebih terang, guna melatih model agar dapat mengenali wajah dalam berbagai

kondisi pencahayaan, seperti di bawah sinar matahari atau di dalam ruangan redup. Teknik flipping horizontal yang tampak pada gambar menambah variasi orientasi wajah, membantu model mengenali wajah yang menghadap ke arah berlawanan. Selain itu, terlihat adanya perubahan ukuran dan distorsi perspektif yang mensimulasikan perbedaan jarak antara wajah dan kamera. Beberapa gambar juga memiliki efek blur, yang membantu model mempelajari pola wajah dalam kondisi gambar yang kurang fokus. Sebaliknya, gambar yang lebih tajam dengan detail yang jelas digunakan untuk menonjolkan elemen penting seperti kontur mata dan bibir. Hasil augmentasi ini bertujuan agar model dapat mengenali wajah dalam berbagai situasi nyata, meningkatkan ketahanannya terhadap variasi data, dan mengurangi risiko overfitting.

F. Model Training

Pada tahap Model Training, penelitian ini melibatkan pelatihan model ArcFace dan Dlib menggunakan data yang sudah diproses melalui deteksi wajah dengan YOLOv8 dan augmentasi gambar. Data training mencakup foto formal mahasiswa Universitas Dian Nuswantoro yang dikonversi ke format grayscale agar elemen penting seperti kontur wajah lebih terlihat. ArcFace dilatih dengan pendekatan Additive Angular Margin Loss yang membantu model membedakan wajah satu dengan yang lain lebih jelas, sehingga lebih tangguh dalam mengenali wajah meski ada perbedaan kecil dalam ekspresi atau pencahayaan[21]. Pelatihan ini dilakukan dalam beberapa tahap (epoch), di mana model terus menyesuaikan parameter untuk meningkatkan keakuratannya. Sementara itu, Dlib dilatih dengan metode yang lebih sederhana, menggunakan HOG (Histogram of Oriented Gradients) untuk mengekstraksi fitur wajah dan SVM (Support Vector Machine) untuk klasifikasi[15]. Sama seperti ArcFace, Dlib dilatih dengan data yang sama, baik dengan maupun tanpa augmentasi, untuk membandingkan performanya. Meskipun Dlib terkenal karena kecepatan dan kemudahan penggunaannya, model ini mungkin kurang handal ketika menghadapi variasi data yang lebih rumit. Selama pelatihan, parameter-parameter seperti kecepatan belajar (learning rate), ukuran batch, dan jumlah epoch diatur sedemikian rupa untuk menjaga keseimbangan performa dan menghindari model overfitting. Model diuji secara berkala untuk memantau kinerjanya dan memastikan bahwa model bisa bekerja dengan baik pada data baru yang belum pernah dilihat sebelumnya.

G. Evaluasi

Dalam penelitian ini, evaluasi model pengenalan wajah dilakukan untuk menilai performa dua pendekatan yang digunakan, yaitu ArcFace dan Dlib[22]. Evaluasi dilakukan dengan menggunakan Data Testing, yaitu data uji yang tidak pernah digunakan dalam proses pelatihan model. Untuk mengukur performa model secara menyeluruh, digunakan beberapa metrik utama, yaitu Akurasi, Precision, Recall, dan F1-score. Metrik-metrik ini dipilih karena mampu memberikan gambaran yang lengkap tentang kemampuan model dalam mengenali dan mengklasifikasikan wajah secara akurat. Akurasi merupakan metrik yang menunjukkan persentase prediksi benar dari keseluruhan prediksi yang dihasilkan oleh model, baik untuk kategori positif maupun negatif. Precision mengukur seberapa tepat model dalam mengenali wajah yang benar dari seluruh prediksi positif yang dibuat. Dalam hal ini, precision penting untuk memahami seberapa jarang model membuat kesalahan dalam mengidentifikasi wajah yang salah. Sebaliknya, Recall atau sensitivitas mengukur seberapa banyak wajah yang sebenarnya dikenali dengan benar oleh model dari seluruh wajah yang ada. Nilai recall yang tinggi menunjukkan bahwa model mampu menangkap sebagian besar wajah yang seharusnya dikenali, meskipun precision mungkin tidak setinggi recall.

F1-score, yang merupakan rata-rata harmonis antara precision dan recall, digunakan untuk menyeimbangkan kedua metrik tersebut, terutama dalam kondisi di mana precision dan recall tidak seimbang. F1-score memberikan gambaran yang lebih akurat tentang kinerja model secara keseluruhan. Dengan menggunakan keempat metrik ini, evaluasi kinerja model akan memberikan pandangan yang lebih komprehensif mengenai kelebihan dan kekurangan dari masing-masing model serta dampak dari penerapan augmentasi gambar terhadap performa pengenalan wajah. Hasil evaluasi akan dianalisis secara statistik untuk memastikan kesimpulan yang valid dan mendukung tujuan penelitian.

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Gambar}} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{2}$$

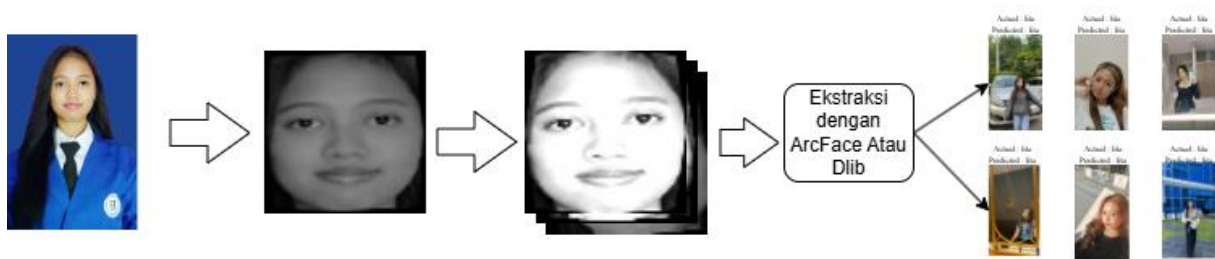
$$\text{Recall} = \frac{TP}{TP+FN} \tag{3}$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

True Positive (TP) merujuk pada prediksi yang tepat di mana model berhasil mengenali wajah yang benar, sedangkan True Negative (TN) mengacu pada kasus di mana model dengan benar tidak mengenali objek yang bukan wajah atau wajah yang tidak relevan. False Positive (FP) terjadi ketika model salah mengidentifikasi objek yang bukan wajah atau wajah yang tidak sesuai sebagai wajah yang benar. False Negative (FN) muncul ketika model gagal mendeteksi wajah yang sebenarnya ada dalam gambar.

III. HASIL DAN PEMBAHASAN

Pada penelitian ini membandingkan performa ArcFace dan Dlib untuk pengenalan wajah yang menggunakan foto dari 10 wajah. Setiap wajah akan diambil masing - masing 10 foto, oleh karena itu dalam penelitian ini akan menggunakan 100 foto. Foto-foto tersebut dipilih untuk mewakili kondisi lingkungan yang cukup beragam, termasuk variasi pencahayaan dan sudut pengambilan gambar. Sebelum pengenalan wajah dilakukan, setiap foto diproses menggunakan YOLOv8 untuk mendeteksi dan memotong area wajah secara otomatis.



Gambar 5. Hasil Pengenalan Wajah

Gambar 5 menampilkan proses pengenalan wajah dengan model ArcFace atau Dlib, yang dimulai dari tahap awal menggunakan foto formal sebagai input referensi. Foto formal ini diambil dalam kondisi pencahayaan standar dan dengan latar belakang polos untuk meminimalkan gangguan, sehingga fitur wajah dapat terekam dengan jelas. Langkah selanjutnya adalah deteksi dan pemotongan wajah menggunakan YOLOv8, yang memusatkan

perhatian hanya pada area wajah sebagai region of interest (ROI). YOLOv8 digunakan karena kemampuannya yang cepat dan akurat dalam mendeteksi wajah, memastikan bahwa area wajah yang relevan terfokus dengan baik. Setelah terdeteksi, wajah kemudian dikonversi menjadi format grayscale. Konversi ini bertujuan untuk meningkatkan efisiensi komputasi dengan hanya menggunakan satu saluran warna dan membantu model fokus pada fitur visual utama, seperti bentuk dan kontur wajah, yang lebih penting dalam pengenalan dibandingkan dengan warna.

Selanjutnya, wajah yang telah dipotong dan dikonversi ke grayscale mungkin mengalami normalisasi atau penyesuaian kontras untuk memperjelas fitur dan mengurangi pengaruh variasi pencahayaan. Hal ini membuat model lebih mudah mengenali fitur wajah yang konsisten, terlepas dari variasi pencahayaan. Pada tahap ekstraksi fitur, baik ArcFace maupun Dlib digunakan untuk menghasilkan representasi wajah yang unik. ArcFace menggunakan pendekatan Additive Angular Margin Loss yang menghasilkan embedding wajah yang lebih robust dan terpisah, sehingga lebih akurat dalam membedakan wajah yang mirip. Sementara itu, Dlib menggunakan pendekatan yang lebih sederhana dengan Histogram of Oriented Gradients (HOG) dan Support Vector Machine (SVM), yang mengandalkan pola kontur dan tepi wajah, tetapi cenderung kurang akurat dalam situasi pencahayaan kompleks.

Pada tahap akhir, hasil pengenalan wajah ditampilkan pada berbagai foto non-formal yang diambil dalam kondisi berbeda-beda, seperti di luar ruangan atau dalam ruangan, dengan pencahayaan dan ekspresi yang beragam. Setiap foto memiliki label "Actual" dan "Predicted" untuk menunjukkan kecocokan prediksi model dengan identitas wajah pada gambar referensi. Model menggunakan fitur wajah yang diekstraksi dari foto referensi untuk mengidentifikasi apakah individu pada gambar non-formal adalah orang yang sama. Visualisasi ini menggambarkan bagaimana model pengenalan wajah bekerja untuk mengenali individu dalam kondisi nyata. ArcFace terbukti lebih efektif dalam mengenali wajah dengan presisi tinggi dalam berbagai situasi dibandingkan dengan Dlib, yang lebih sederhana tetapi kurang kuat untuk menangani variasi data yang kompleks.

TABEL III. TABEL HASIL PENGUJIAN

| Model Jumlah Augmentasi | ArcFace | | | | | Dlib | | | | |
|----------------------------|---------|---------|------------|--------|----------|---------|----------|------------|--------|----------|
| | Time | Akurasi | Precisi on | Recall | F1 Score | Time | Aku rasi | Precisi on | Recall | F1 Score |
| - | 6m 52s | 0.80 | 0.80 | 1.00 | 0.89 | 10m 2s | 0.57 | 0.57 | 1.00 | 0.73 |
| 1 | 4m 36s | 0.80 | 0.80 | 1.00 | 0.89 | 9m 52s | 0.57 | 0.57 | 1.00 | 0.73 |
| 5 | 4m 35s | 0.82 | 0.82 | 1.00 | 0.90 | 9m 50s | 0.57 | 0.57 | 1.00 | 0.73 |
| 10 | 4m 38s | 0.82 | 0.82 | 1.00 | 0.90 | 10m | 0.55 | 0.55 | 1.00 | 0.71 |
| 20 | 4m 44s | 0.83 | 0.83 | 1.00 | 0.91 | 10m 7s | 0.55 | 0.55 | 1.00 | 0.71 |
| 30 | 4m 40s | 0.83 | 0.83 | 1.00 | 0.91 | 9m 50s | 0.57 | 0.57 | 1.00 | 0.73 |
| 40 | 6m 57s | 0.84 | 0.84 | 1.00 | 0.91 | 19m 14s | 0.52 | 0.52 | 1.00 | 0.68 |
| 50 | 6m 54s | 0.90 | 0.90 | 1.00 | 0.95 | 19m 10s | 0.54 | 0.54 | 1.00 | 0.70 |

Tabel 3 membandingkan performa model ArcFace dan Dlib dalam pengenalan wajah pada berbagai tingkat augmentasi data menggunakan metrik waktu eksekusi, akurasi, presisi, recall, dan F1 score. Dari analisis tabel, terlihat bahwa model ArcFace menunjukkan peningkatan performa seiring bertambahnya jumlah augmentasi. Akurasi awal sebesar 0,80 tanpa augmentasi meningkat hingga 0,90 saat jumlah augmentasi mencapai 50. Skor F1 pada

ArcFace juga mengalami peningkatan yang serupa, dari 0,89 menjadi 0,95, mengindikasikan peningkatan kemampuan model dalam mengenali wajah secara konsisten. Selain itu, nilai recall yang stabil pada 1,00 menunjukkan bahwa ArcFace mampu mendeteksi semua kelas yang relevan secara optimal tanpa kehilangan data penting. Waktu pemrosesan ArcFace tetap efisien, hanya mengalami sedikit peningkatan dari 6 menit 52 detik tanpa augmentasi menjadi 6 menit 54 detik pada 50 augmentasi, yang menunjukkan bahwa ArcFace mampu menangani tambahan augmentasi dengan efisiensi waktu yang baik. Penelitian ini menunjukkan bahwa model ArcFace memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan penelitian sebelumnya[22]. Oleh karena itu, model ini sangat direkomendasikan untuk diterapkan dalam sistem absensi kelas.

Sebaliknya, model Dlib tidak menunjukkan peningkatan signifikan dalam akurasi, presisi, maupun skor F1 meskipun augmentasi bertambah. Akurasi pada Dlib tetap di kisaran 0,57 hingga 0,54, dengan skor F1 berkisar antara 0,73 hingga 0,70. Meskipun demikian, nilai recall Dlib juga stabil pada 1,00 untuk setiap jumlah augmentasi, yang berarti model ini mampu mendeteksi semua kelas yang relevan meskipun presisi dan akurasinya relatif rendah. Waktu pemrosesan pada Dlib meningkat tajam seiring bertambahnya augmentasi, dari 10 menit 2 detik tanpa augmentasi menjadi 19 menit 10 detik pada 50 augmentasi, menunjukkan bahwa Dlib kurang efisien dibandingkan ArcFace dalam hal waktu pemrosesan. Penelitian ini mengungkapkan bahwa model Dlib memiliki tingkat akurasi yang lebih rendah dibandingkan dengan penelitian sebelumnya[17]. Oleh karena itu, model ini dianggap kurang cocok untuk direkomendasikan dalam penggunaannya. Secara keseluruhan, model ArcFace menunjukkan performa yang lebih baik dalam penelitian ini dengan keunggulan dalam akurasi, efisiensi waktu, dan responsivitas terhadap augmentasi, menjadikannya pilihan yang lebih optimal untuk tugas pengenalan wajah.

IV. KESIMPULAN

Kesimpulan penelitian ini menunjukkan bahwa model ArcFace memiliki performa yang lebih unggul dibandingkan Dlib dalam pengenalan wajah menggunakan dataset terbatas, terutama dalam hal akurasi dan efisiensi waktu pemrosesan pada berbagai variasi augmentasi data. Hasil pengujian memperlihatkan bahwa ArcFace secara konsisten mencapai akurasi tertinggi sebesar 90% dan F1 score 95 pada level 50 augmentasi, sementara Dlib menunjukkan performa yang lebih rendah dan tidak stabil dengan akurasi maksimal 57% dengan waktu pemrosesan hingga 19 menit pada level 1, 5, dan 30 augmentasi. Dengan kinerja stabil dan efisien, ArcFace menjadi pilihan yang sangat disarankan untuk aplikasi pengenalan wajah yang memerlukan keseimbangan antara ketepatan dan kecepatan, bahkan pada dataset terbatas. Selain itu, penerapan sistem ini dalam lingkungan pendidikan, seperti sistem absensi, perlu diteliti lebih lanjut untuk validasi tambahan dan peningkatan kualitas pengenalan wajah, termasuk melalui teknik upscaling gambar untuk memperbaiki ketajaman dan detail dalam ekstraksi fitur.

REFERENSI

- [1] M. Wang and W. Deng, "Deep face recognition: A survey," *Neurocomputing*, vol. 429, pp. 215–244, 2021, doi: 10.1016/j.neucom.2020.10.081.
- [2] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri, "Face recognition systems: A survey," *Sensors (Switzerland)*, vol. 20, no. 2, 2020, doi: 10.3390/s20020342.
- [3] N. Abudarham, L. Shkiller, and G. Yovel, "Critical features for face recognition," *Cognition*,

- vol. 182, no. April 2018, pp. 73–83, 2019, doi: 10.1016/j.cognition.2018.09.002.
- [4] K. H. Teoh, R. C. Ismail, S. Z. M. Naziri, R. Hussin, M. N. M. Isa, and M. S. S. M. Basir, “Face Recognition and Identification using Deep Learning Approach,” *J. Phys. Conf. Ser.*, vol. 1755, no. 1, 2021, doi: 10.1088/1742-6596/1755/1/012006.
 - [5] L. Li, X. Mu, S. Li, and H. Peng, “A Review of Face Recognition Technology,” *IEEE Access*, vol. 8, pp. 139110–139120, 2020, doi: 10.1109/ACCESS.2020.3011028.
 - [6] J. Chang, Z. Lan, C. Cheng, and Y. Wei, “Data Uncertainty Learning in Face Recognition,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 5709–5718, 2020, doi: 10.1109/CVPR42600.2020.00575.
 - [7] M. O. Oloyede, G. P. Hancke, and H. C. Myburgh, “A review on face recognition systems: recent approaches and challenges,” *Multimed. Tools Appl.*, vol. 79, no. 37–38, pp. 27891–27922, 2020, doi: 10.1007/s11042-020-09261-2.
 - [8] A. Tessa Ningrum, R. Wijay, M. Rizal Abdul Aziz, M. Yudha Mauluda, and P. Rosyani, “Face Deteksi Objek pada Gambar dan Video dengan YOLOv8 (Counting Objects),” *J. Artif. Intel. dan Sist. Penunjang Keputusan*, vol. 2, no. 2, pp. 119–126, 2024, [Online]. Available: <https://jurnalmahasiswa.com/index.php/aidanspk/article/view/1485>.
 - [9] D. Reis, J. Kupec, J. Hong, and A. Daoudi, “Real-Time Flying Object Detection with YOLOv8,” 2023, [Online]. Available: <http://arxiv.org/abs/2305.09972>.
 - [10] M. Sohan, T. Sai Ram, and C. V. Rami Reddy, “A Review on YOLOv8 and Its Advancements,” no. May, pp. 529–545, 2024, doi: 10.1007/978-981-99-7962-2_39.
 - [11] J. Deng, J. Guo, J. Yang, N. Xue, I. Kotsia, and S. Zafeiriou, “ArcFace: Additive Angular Margin Loss for Deep Face Recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 10, pp. 5962–5979, 2022, doi: 10.1109/TPAMI.2021.3087709.
 - [12] S. Suwarno and K. Kevin, “Analysis of Face Recognition Algorithm: Dlib and OpenCV,” *J. Informatics Telecommun. Eng.*, vol. 4, no. 1, pp. 173–184, 2020, doi: 10.31289/jite.v4i1.3865.
 - [13] S. Komkov and A. Petiushko, “AdvHat: Real-world adversarial attack on arcFace face ID system,” *Proc. - Int. Conf. Pattern Recognit.*, pp. 819–826, 2020, doi: 10.1109/ICPR48806.2021.9412236.
 - [14] D. D. Nguyen, M. H. Le, X. H. Nguyen, H. T. Ngo, and M. S. Nguyen, “Smart Desk in Hybrid Classroom: Automatic Attendance System based on Face Recognition using MTCNN and ARCFACE,” *2022 Int. Conf. Multimed. Anal. Pattern Recognition, MAPR 2022 - Proc.*, pp. 1–6, 2022, doi: 10.1109/MAPR56351.2022.9924775.
 - [15] D. Zhang, J. Li, and Z. Shan, “Implementation of Dlib deep learning face recognition technology,” *Proc. - 2020 Int. Conf. Robot. Intell. Syst. ICRIS 2020*, pp. 88–91, 2020, doi: 10.1109/ICRIS52159.2020.00030.
 - [16] M. T. Aydın, O. Menemencioglu, and İ. M. Orak, “Face Recognition Approach Using Dlib and K-Nn,” vol. 1, no. 2, pp. 93–103, 2023, [Online]. Available: <http://ctc.karabuk.edu.tr>.
 - [17] S. Mohanty, S. V. Hegde, S. Prasad, and J. Manikandan, “Design of real-time drowsiness detection system using dlib,” *2019 5th IEEE Int. WIE Conf. Electr. Comput. Eng. WIECON-ECE 2019 - Proc.*, pp. 1–4, 2019, doi: 10.1109/WIECON-ECE48653.2019.9019910.
 - [18] N. P. Motwani and S. S., “Human Activities Detection using DeepLearning Technique-YOLOv8,” *ITM Web Conf.*, vol. 56, p. 03003, 2023, doi: 10.1051/itmconf/20235603003.
 - [19] I. Zeger, S. Grgic, J. Vukovic, and G. Sisul, “Grayscale Image Colorization Methods: Overview and Evaluation,” *IEEE Access*, vol. 9, pp. 113326–113346, 2021, doi: 10.1109/ACCESS.2021.3104515.
 - [20] H. T. Duong and T. A. Nguyen-Thi, “A review: preprocessing techniques and data augmentation for sentiment analysis,” *Comput. Soc. Networks*, vol. 8, no. 1, pp. 1–16, 2021, doi: 10.1186/s40649-020-00080-x.
 - [21] C. Oinar, B. M. Le, and S. S. Woo, “KappaFace: Adaptive Additive Angular Margin Loss for Deep Face Recognition,” *IEEE Access*, vol. 11, no. 1, pp. 137138–137150, 2023, doi: 10.1109/ACCESS.2023.3338648.
 - [22] R. Arun Kumar, V. A. Solayappan, T. Sree Sharmila, and K. Ram Prasad, “Masked Deep Face

Recognition using ArcFace and Ensemble Learning,” *2021 IEEE 2nd Int. Conf. Technol. Eng. Manag. Soc. Impact using Mark. Entrep. Talent. TEMSMET 2021*, 2021, doi: 10.1109/TEMSMET53515.2021.9768777.