

Implementation of Identity Loss Function on Face Recognition of Low-Resolution Faces With Light CNN Architecture

Tsaqif Mu'tashim Mufid ^{1*}, Riza Ibnu Adam ^{2*}, Jajam Khaeru Jaman ^{3*}, Garno ^{4*}, Iqbal Maulana ^{5*}
* Informatika, Universitas Singaperbangsa Karawang
tsaqif.mutashim19144@student.unsika.ac.id ¹, riza.adam@staff.unsika.ac.id ², jajam.haeruljaman@staff.unsika.ac.id ³,
garno@staff.unsika.ac.id ⁴, iqbal.maulana@staff.unsika.ac.id ⁵

Article Info

Article history:

Received 2023-08-15
Revised 2024-05-21
Accepted 2024-06-06

Keyword:

AdaFace,
Face recognition,
Low-resolution images,
Super-resolution,
Light CNN architecture.

ABSTRACT

Face recognition in low-resolution images has seen significant advancements over the past few decades. Although extensive research has been conducted to improve accuracy in these conditions, one of the main challenges remains the difficulty in identifying unique facial features in low-resolution images, leading to high error rates in identification. The use of Deep Convolutional Neural Networks (DCNN) for low-resolution face recognition is still limited. However, employing super-resolution models like REAL-ESRGAN can enhance recognition accuracy in low-resolution images. This study utilizes the Light CNN architecture and applies the margin-based identity loss function AdaFace on low-resolution datasets. The model is trained using the Casia-WebFace dataset and evaluated using the LFW and TinyFace test datasets. Based on the evaluation results on the LFW test data, the best model is Light CNN9-AdaFace, achieving the highest accuracy of 97.78% at 128x128 resolution. For images with the lowest resolution of 16x16, an accuracy of 83.37% was achieved using super-resolution techniques. On the TinyFace test data, the use of super-resolution resulted in performance metrics with a Rank-1 accuracy of 47.26%, Rank-5 accuracy of 55.25%, Rank-10 accuracy of 58.61%, and Rank-20 accuracy of 61.90% using the Light CNN9-AdaFace architecture.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Dalam beberapa dekade terakhir, pengenalan wajah yang merupakan sub-bidang di dalam Computer Vision telah mengalami kemajuan yang luar biasa [1]. Teknologi ini mengandalkan karakteristik wajah seseorang untuk tujuan verifikasi dan identifikasi [2]. Berbagai sektor industri, mulai dari keamanan dan pengawasan hingga kontrol akses telah menggunakan teknologi pengenalan wajah [3]. Meskipun telah diaplikasikan di banyak sektor, pengenalan wajah beresolusi rendah masih menjadi tantangan utama. Dalam banyak situasi praktis, seperti sistem keamanan CCTV dan pengambilan gambar dari jarak jauh, diperoleh gambar wajah memiliki resolusi rendah, yang mempengaruhi kinerja sistem pengenalan wajah. Meskipun terdapat metode yang dikembangkan untuk meningkatkan akurasi pengenalan

wajah pada resolusi rendah, tetap ada kelemahan dalam penggunaannya. Salah satu kelemahan utama adalah bahwa semakin rendah resolusi gambar, semakin sulit untuk mengidentifikasi fitur wajah yang unik, sehingga dapat menghasilkan kemungkinan kesalahan identifikasi yang tinggi.

Berdasarkan kelemahan tersebut terdapat beberapa penelitian yang membahas teknik pengenalan wajah pada gambar dengan resolusi rendah. Salah satu metode pendekatan *Local Binary Pattern* (LBP) dan *Principal Component Analysis* (PCA) digunakan dalam sebuah penelitian oleh [4] untuk meningkatkan akurasi pengenalan wajah pada foto beresolusi rendah. Hasil penelitian tersebut menunjukkan bahwa menggunakan algoritma LBP dan PCA pada foto beresolusi rendah dapat meningkatkan akurasi pengenalan wajah hingga 94%.

Deep Convolutional Neural Network (DCNN) untuk pengenalan wajah skala besar adalah metode pilihan untuk pengenalan wajah. LightCNN [5] termasuk dalam kategori DCNN. LightCNN sendiri adalah sebuah framework yang dirancang untuk menghasilkan representasi yang mendalam dari wajah khususnya pada "Noisy Label".

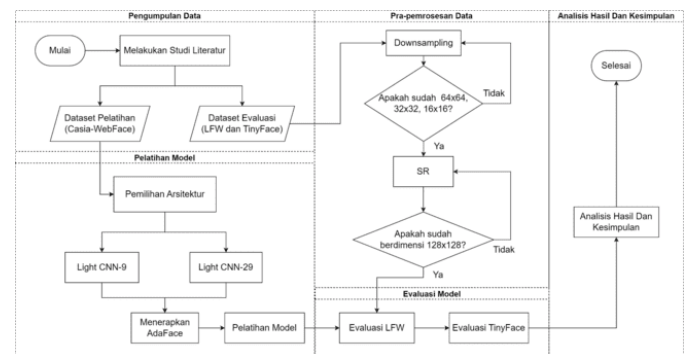
Selain itu, penggunaan super resolusi telah digunakan untuk memperbaiki kualitas gambar resolusi rendah seperti yang pada penelitian sebelumnya [6] Sebagian besar evaluasi pada pengenalan wajah menggunakan gambar resolusi tinggi. Namun, berbeda dengan pengenalan wajah pada gambar dalam skenario resolusi rendah yang dapat mengurangi akurasi prediksi karena gambar resolusi rendah biasanya kurang memiliki detail diskriminatif. Untuk mengatasi masalah ini, salah satu solusi yang umum digunakan adalah dengan menerapkan konsep *super resolution* (SR). Metode SR dapat membantu meningkatkan akurasi pengenalan wajah dengan menghasilkan gambar yang lebih berkualitas dan mendekati resolusi tinggi dari gambar resolusi rendah [7].

DCNN menghasilkan fitur dengan jarak antar-kelas yang sangat besar dan jarak intrakelas yang kecil dari gambar wajah. Ada dua jalur penelitian utama dalam melatih DCNN untuk pengenalan wajah: pendekatan multi-kelas yang menggunakan kelas softmax untuk mengidentifikasi identitas yang berbeda dalam training set dan cara-cara yang secara langsung mempelajari embedding, seperti metode triplet loss [8]. Pada metode jaringan syaraf tiruan penulis akan menggunakan *identity loss* berupa *Quality Adaptive Margin for Face Recognition* (AdaFace) sebuah margin yang dapat beradaptasi berdasarkan kualitas gambarnya. AdaFace yang dikembangkan oleh [9] merupakan salah satu *identity loss* yang dapat meningkatkan performa pengenalan wajah pada skenario beresolusi rendah karena cara kerjanya yang adaptif terhadap kualitas gambar, AdaFace sendiri akan memberikan bobot yang berbeda pada sampel-sampel dengan tingkat kesulitan yang berbeda berdasarkan kualitas gambar mereka. Hal ini memungkinkan jaringan syaraf tiruan dengan *identity loss* yang merupakan fungsi *loss* yang digunakan dalam proses training CNN untuk menghasilkan representasi fitur wajah yang lebih kuat dan mampu mengenali identitas wajah dengan lebih akurat.

Berdasarkan pembahasan tersebut diketahui bahwa banyaknya teknik pengenalan wajah beresolusi rendah, namun pengenalan wajah resolusi rendah menggunakan DCNN masih belum banyak dilakukan oleh para peneliti. Untuk itu penelitian saat ini akan menggunakan LightCNN [5] dengan super resolusi REAL-ESRGAN [10] sebagai fitur ekstraksi gambar pada wajah beresolusi rendah dan penggunaan fungsi *identity loss* berbasis AdaFace [9] pada dataset beresolusi rendah. Dalam proses pelatihan model, digunakan dataset CASIA-WebFace [11] dengan ukuran dimensi 128x128. Sedangkan untuk evaluasi model, digunakan dataset LFW (Labeled Faces in the Wild) [12] dengan variasi ukuran dimensi yaitu 128x128, 64x64, 32x32, 16x16, dan juga dataset TinyFace [13].

II. METODE PENELITIAN

Metode yang dilakukan pada penelitian ini meliputi beberapa langkah penelitian (Gambar 1) dimulai dari tahap pengumpulan data berupa dataset, pra-pemrosesan data untuk pembuatan skenario *low resolution* dimana prosesnya merupakan penyesuaian terhadap perubahan dimensi resolusi tinggi menjadi resolusi rendah ditambah dengan bantuan model super resolusi REAL-ESRGAN [10] pada dataset LFW, pelatihan model yang akan menggunakan arsitektur Light CNN-9, Light CNN-29 [5] dan AdaFace [9], setelah itu masuk pada tahapan evaluasi yang akan menggunakan tugas evaluasi metrik yaitu *face verification* dan *face identification*, lalu terakhir akan adanya analisis hasil dan kesimpulan.



Gambar 1. Diagram Alur Penelitian

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Pengumpulan data pada penelitian ini terdapat 2 jenis dataset yaitu dataset pelatihan dan evaluasi. **Dataset pelatihan** yang akan digunakan pada penelitian ini adalah Casia-WebFace [11]. Dataset ini berjumlah kurang lebih 494,414 gambar dengan 10,575 identitas dengan dimensi 112x112. Dalam dataset Casia-WebFace, wajah – wajah yang berasal dari identitas yang sama disusun secara kolektif dalam suatu folder, yang kemudian diberikan label numerik sebagai pengganti pengidentifikasi perseorangan. Dataset ini diambil dari repositori InsightFace [15] yang dimana merupakan dataset yang sudah *aligned*.



Gambar 2. Sample Wajah Dataset CASIA-WebFace

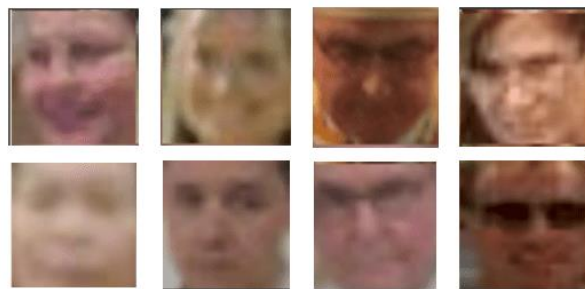
Dataset evaluasi sendiri terbagi atas 2 dataset yang akan digunakan, salah satunya dataset LFW (Labeled Faces in the Wild) [12]. Dataset ini berdimensi 112x112 dengan jumlah

gambar wajah sebanyak 13,233 dari 5,749 identitas berbeda. Nantinya dataset ini akan dibuat skenario dengan metode *bicubic interpolation* ke variasi dimensi 64x64, 32x32, dan 16x16. Dataset ini juga diambil dari repositori InsightFace [15] yang merupakan dataset wajah yang telah diluruskan (*aligned*).



Gambar 2. Sample Wajah Dataset LFW

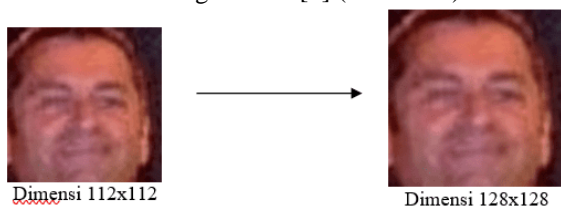
Dataset selanjutnya adalah **TinyFace** [13] yang diambil pada repositori AdaFace [9] merupakan data gambar wajah yang telah diluruskan (*aligned*) menggunakan model *face detection* MTCNN. Dataset ini terdiri dari 5.139 identitas (label) dengan total 169.403 gambar wajah yang memiliki resolusi rendah, dengan rata-rata dimensi sekitar 20x16 piksel dan mencakup berbagai pose, ekspresi, dan kondisi pencahayaan yang berbeda.



Gambar 3. Sample Wajah Dataset TinyFace

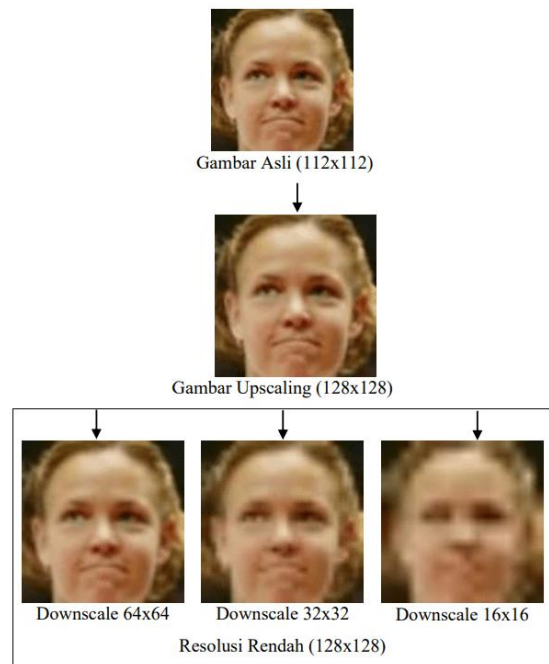
B. Pra-pemrosesan Data

Pra-pemrosesan data dilakukan terhadap dataset pelatihan yaitu Casia-WebFace [11] dan dataset evaluasi yaitu LFW (Labeled Faces in the Wild) [12] serta TinyFace [13] dengan penambahan model *super resolution* REAL-ESRGAN [10] sebagai bentuk eksperimen. Pra-pemrosesan pada dataset pelatihan dilakukan untuk membuat dataset Casia-WebFace [11] dalam skenario untuk bisa dilatih oleh model. Tahapan pertama yang dilakukan adalah melakukan *upscaling* data yang mempunyai dimensi 112x112 diubah menjadi dimensi 128x128. Hal ini dilakukan untuk penyesuaian dalam *input size* dalam arsitektur Light CNN [5] (Gambar 2).



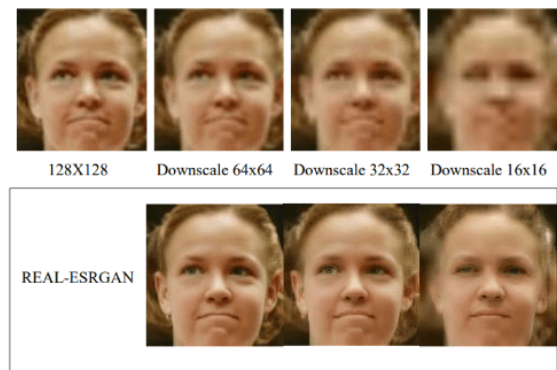
Gambar 4. Hasil Upscaling 112x112 menjadi 128x128

Setelah itu, dataset LFW [12] akan dilakukan pra-pemrosesan data yang bertujuan untuk membuat skenario menjadi *low resolution* (Gambar 3) dengan cara *downscaling* dataset LFW yang berdimensi 128x128. Tahapan pertama yang dilakukan disini adalah *upscaling* pada dataset LFW yang berdimensi 112x112 untuk menyesuaikan input arsitektur Light CNN [5], selanjutnya setelah dilakukan *upscaling* dataset LFW akan melakukan proses *downsampling* pada data yang memiliki dimensi 128x128 kedalam dimensi yang lebih rendah yaitu 64x64, 32x32, serta 16x16 piksel kemudian di *upscaling* kembali pada dimensi 128x128. Hal ini dilakukan untuk membuat skenario *low resolution* dengan cara mengaburkan fitur yang terdapat pada gambar wajah.



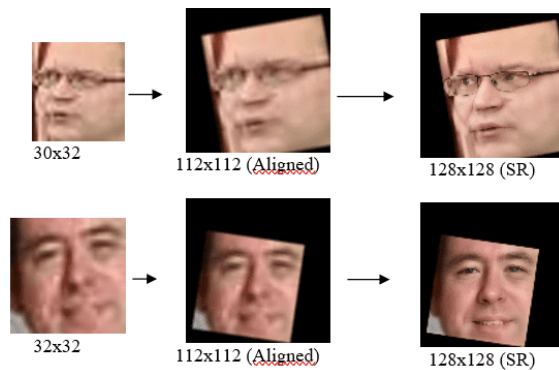
Gambar 5. Hasil Proses Pembuatan Skenario Low Resolution

Setelah itu dilakukan pembuatan skenario low resolution. Selanjutnya, gambar pada wajah akan dibantu dengan model *super resolution* REAL-ESRGAN [10] (Gambar 4). Penggunaan model *super resolution* bertujuan untuk mengembalikan fitur yang hilang pada proses sebelumnya.



Gambar 6. Hasil Proses Model SR Pada Dataset LFW

Terakhir pada dataset TinyFace [13] akan dilakukan juga pra-pemrosesan data menggunakan bantuan model *super resolution* REAL-ESRGAN (Gambar 5) [10].



Gambar 7. Hasil Proses Restorasi Model SR Pada Dataset TinyFace

C. Pelatihan Model

Model akan dilatih menggunakan arsitektur LightCNN [5] sebagai *backbone* dengan *identity loss* berbasis *margin* AdaFace [9] dengan konfigurasi *batch size* 512, *save frequency* 1, *test frequency* 5, dan total *epoch* 70. Dataset CASIA-WebFace [11] berdimensi 128x128 digunakan sebagai data latih sedangkan LFW [12] berdimensi 128x128 digunakan sebagai data evaluasi. Dalam pelatihan digunakan optimizer Adam dengan *weight decay* $5e-4$ serta *learning rate* sebesar 0.0001 dan *learning_rate_scheduler* dengan *gamma* 0.1 dimana *learning rate* akan dibagi 10 pada epochs 36, 52 dan 58. Sedangkan untuk konfigurasi parameter AdaFace, digunakan *margin* 0.4, *hardness* 0.333, *scale* 64.0, dan *t_alpha* 0.01. Adapun pelatihan model menggunakan Nvidia Tesla A100 dengan 40GB RAM. Arsitektur Light CNN-9 [5]

THE ARCHITECTURES OF THE LIGHT CNN-9 MODEL.

Type	Filter Size /Stride, Pad	Output Size	#Params
Conv1	$5 \times 5/1, 2$	$128 \times 128 \times 96$	2.4K
MFM1	-	$128 \times 128 \times 48$	-
Pool1	$2 \times 2/2$	$64 \times 64 \times 48$	-
Conv2a	$1 \times 1/1$	$64 \times 64 \times 96$	4.6K
MFM2a	-	$64 \times 64 \times 48$	-
Conv2	$3 \times 3/1, 1$	$64 \times 64 \times 192$	165K
MFM2	-	$64 \times 64 \times 96$	-
Pool2	$2 \times 2/2$	$32 \times 32 \times 96$	-
Conv3a	$1 \times 1/1$	$32 \times 32 \times 192$	18K
MFM3a	-	$32 \times 32 \times 96$	-
Conv3	$3 \times 3/1, 1$	$32 \times 32 \times 384$	331K
MFM3	-	$32 \times 32 \times 192$	-
Pool3	$2 \times 2/2$	$16 \times 16 \times 192$	-
Conv4a	$1 \times 1/1$	$16 \times 16 \times 384$	73K
MFM4a	-	$16 \times 16 \times 192$	-
Conv4	$3 \times 3/1, 1$	$16 \times 16 \times 256$	442K
MFM4	-	$16 \times 16 \times 128$	-
Conv5a	$1 \times 1/1$	$16 \times 16 \times 256$	32K
MFM5a	-	$16 \times 16 \times 128$	-
Conv5	$3 \times 3/1, 1$	$16 \times 16 \times 256$	294K
MFM5	-	$16 \times 16 \times 128$	-
Pool4	$2 \times 2/2$	$8 \times 8 \times 128$	-
fc1	-	512	4,194K
MFM_fc1	-	256	-
Total	-	-	5,556K

TABEL I
HASIL PELATIHAN LIGHT CNN-9 ADAFACE

Epoch	Total Loss	Mean Accuracy
1	34.9052	-
5	27.4693	91.0667
10	19.2367	95.0833
15	14.7672	96.4
.....
60	7.7095	97.6
65	7.7094	97.6167
70	7.711	97.6

Hasil pelatihan model pada Tabel I menunjukkan bahwa proses training pada model Light CNN-9 [5] dengan *identity loss* berbasis *margin* AdaFace [9] diperoleh nilai *total_loss* yang terus menurun sampai dengan titik terendah sebesar 7,711 pada epoch ke-70 sedangkan akurasi model mampu meningkat sampai mencapai titik konvergen sebesar 97,6. Namun, pada epoch 70 terjadi penurunan *mean accuracy* terhadap epoch 65 sebesar 0,01%. Oleh karena itu akan dilakukan evaluasi model untuk pengujian ulang menggunakan dataset LFW 128x128 pada tugas *face verification*.

D. Evaluasi Model

Hasil dari pelatihan model akan menghasilkan berupa file *checkpoint* yang akan digunakan pada tugas evaluasi *face verification* pada dataset LFW[12] dan *face identification* pada TinyFace[13]. Evaluasi ini dilakukan untuk mengetahui seberapa mampu model pengenalan wajah dengan penggunaan fungsi *loss* berbasis *margin* AdaFace dapat memprediksi dan mengklasifikasikan wajah dengan baik khususnya dalam skenario *low resolution*. Berikut merupakan hasil evaluasi model berdasarkan model hasil pelatihan pada sebelumnya yang bisa dilihat di bawah ini.

I. Face Verification

Terdapat dua skema evaluasi yaitu dataset LFW yang menggunakan *super resolution* dan *non super resolution*. Selain itu juga terdapat penurunan dimensi yang semula dari 128x128 menjadi 64x64, 32x32 dan 16x16 yang menjadi fokus utama dalam pengenalan wajah beresolusi rendah. Semua proses evaluasi akan menggunakan metode 10-fold-cross-validation yaitu proses dimana dataset dibagi menjadi 10 bagian yang seimbang secara acak dan dievaluasi sebanyak 10 iterasi untuk memberikan gambaran yang lebih komprehensif tentang kinerja model dengan *metrics mean accuracy*. Sebelum memulai tahap evaluasi, data gambar akan dipecah menjadi dua kategori: pasangan positif dan pasangan negative berdasarkan informasi yang terdapat dalam file pairs.txt. Contoh pembagian ini untuk data pasangan wajah dapat dilihat pada (Gambar 8) dalam dataset LFW.



Gambar 8. Gambar Wajah Pasangan Positif dan Negatif pada LFW

Selain itu, metode ini juga membantu dalam mengukur seberapa baik model mampu melakukan generalisasi pada data yang belum pernah dilihat sebelumnya. Berikut hasil evaluasi tersebut dalam ke dua skema yaitu menggunakan *identity loss* dan model terbaik tanpa *identity loss* dari [5] yang dibantu menggunakan super resolusi pada masing – masing perubahan resolusi.

TABEL II
EVALUASI MODEL LIGHT CNN-9 PADA LFW

ARSITEKTUR	ITERASI	AKURASI			
		128x128	64x64	32x32	16x16
LIGHT CNN-9	1	70.17	68.67	68.67	67.33
	2	65.83	64.67	63.17	63.50
	3	69.67	66.17	67.33	65.33
	4	71.83	70.00	72.00	69.67
	5	68.00	66.83	66.33	67.67
	6	69.50	69.67	67.67	69.33
	7	67.67	67.67	67.50	63.33
	8	62.17	65.33	66.83	62.33
	9	66.00	65.83	67.17	65.67
	10	68.33	67.00	67.67	65.33
MEAN ACCURACY		67.92	67.18	67.43	65.95

Pada Tabel II menunjukkan bahwa hasil performa model LightCNN-9 tanpa menggunakan *loss* berbasis *margin* AdaFace dan juga tidak menggunakan model *super resolution* memiliki rata – rata akurasi cenderung rendah dengan 67.18% pada dimensi 64x64, 67.43% pada dimensi 32x32, dan 65,95 pada dimensi 16x16. Hal ini menandakan bahwa tingkat keberhasilan dari model terbaik dari [5] belum bisa mengenali data secara general yang belum pernah dilihat.

TABEL III
EVALUASI MODEL LIGHT CNN-9 SUPER RESOLUSI PADA LFW

ARSITEKTUR	ITERASI	AKURASI			
		128x128	64x64	32x32	16x16
LIGHT CNN-9 + REAL-ESRGAN	1	70.17	69	69	67.5
	2	65.83	68.83	65	64.67
	3	69.67	69	67.17	65.33
	4	71.83	72.5	68.33	69.67
	5	68.00	68.17	66.17	65.5
	6	69.50	68	67.17	66
	7	67.67	68.83	65.5	65.33
	8	62.17	65.83	63.83	64.67
	9	66.00	70.17	65.5	65.67
	10	68.33	69.17	65.33	65.5
MEAN ACCURACY		67.92	68.95	66.3	65.98

Sedangkan pada Tabel III menunjukkan bahwa penggunaan super resolusi REAL-ESRGAN [10] cukup meningkat dibandingkan pada tabel sebelumnya yaitu rata – rata akurasi di dapatkan dengan 67.92% pada dimensi 128x128, 68.9% pada dimensi 64x64, 66.3% pada dimensi 32x32, dan 65,98. Hal ini masih sama pada sebelumnya bahwa tingkat keberhasilan dari model terbaik dari [5] masih belum bisa mengenali data secara general yang belum pernah dilihat.

TABEL IV
EVALUASI MODEL LIGHT CNN-9 SUPER RESOLUSI + ADAFACE PADA LFW

ARSITEKTUR	ITERASI	AKURASI			
		128x128	64x64	32x32	16x16
LIGHT CNN-9 + REAL-ESRGAN + ADAFACE	1	97.83	97.67	95.5	83.83
	2	97.5	97	95.5	85
	3	98	97.83	95.83	83
	4	96.5	96.5	94.67	83.17
	5	96.83	97.67	93.83	82.5
	6	98.33	97.67	95	82.33
	7	98.17	97.5	96	85.83
	8	98	96.67	95.83	82.17
	9	98.33	97.83	96.33	81.67
	10	98.33	96.83	95.17	84.17
MEAN ACCURACY		97.78	97.32	95.37	83.37

Namun pada Tabel IV menunjukkan hasil bahwa penggunaan *loss* berbasis *margin* AdaFace menghasilkan peningkatan yang signifikan dibarengi model *super resolution* REAL-ESRGAN dibandingkan dengan model terbaik LightCNN-9 dari [5]. Pada dimensi 128x128 mendapatkan rata – rata akurasi secara signifikan meningkat sebesar 97.78%, 97.32% pada dimensi 64x64, 95.37% pada dimensi 32x32, dan 83.37% pada dimensi 16x16. Performa model LightCNN-9 tanpa menggunakan *loss* berbasis *margin* AdaFace [9] tidak mampu mendapatkan rata – rata akurasi yang baik dalam proses tugas evaluasi menggunakan matriks *face verification*.

II. Face Identification

Sementara itu untuk mengevaluasi performa model dalam tugas *face identification* [14] maka model pengenalan wajah akan diujikan pada dataset TinyFace [13]. Evaluasi akan dilakukan menggunakan protokol *face identification* dengan Rank-1, Rank-5, Rank-10, Rank-20 sebagai matriks pengujian. Selama proses evaluasi, gambar di folder Probe akan diuji satu per satu, mencari gambar dengan skor kemiripan tertinggi dari galeri, yang merupakan kombinasi gambar dari folder *Gallery_Match* dan *Gallery_Distractor*. Skema pembagian data untuk evaluasi TinyFace dapat dilihat pada (Gambar 9).



Gambar 9. Skema Pembagian Data Pada TinyFace

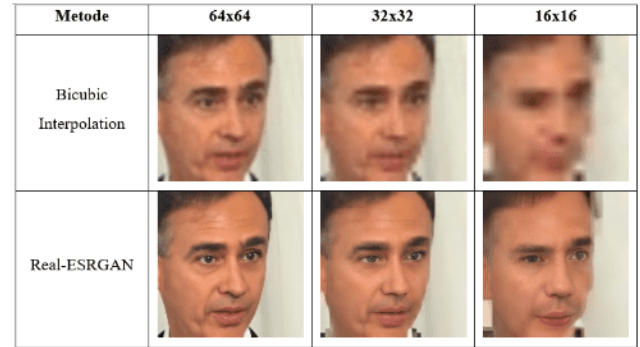
Berikut hasil dari evaluasi menggunakan model dengan *identity loss* berbasis *margin* AdaFace untuk hasil evaluasi pada *face identification*.

TABEL V
EVALUASI MODEL LIGHT CNN-9 SUPER RESOLUSI + ADAFACE PADA TINYFACE

ARSITEKTUR	AKURASI			
	RANK-1	RANK-5	RANK-10	RANK-20
LIGHT CNN-9 + ADAFACE	47.26%	55.25%	58.61%	61.90%

E. Analisis Hasil dan Kesimpulan

Berdasarkan hasil uji sintesis evaluasi model sebelumnya terdapat beberapa indikator untuk bisa ditarik kesimpulan yang pertama tingkat akurasi yang tinggi dan konsistensi kinerja model terhadap perubahan resolusi gambar. Perubahan resolusi yang digunakan juga mempengaruhi dalam menentukan kedekatan jarak antara gambar wajah dengan pasangannya serta penggunaan model super resolusi REAL-ESRGAN dapat meningkatkan kualitas gambar pada perubahan resolusi dibandingkan dengan metode *bicubic interpolation*. Hasil ini bisa dilihat pada (Gambar 10).



Gambar 10. Perbandingan Downsampling Bicubic Interpolation dan Real-ESRGAN

Dari hasil perbandingan tersebut terlihat bahwa khususnya pada dataset LFW berdasarkan hasil evaluasi model pada dimensi 128x128 Light CNN9-AdaFace lebih unggul sebesar 97,78% dibandingkan dengan arsitektur tanpa menggunakan loss berbasis margin AdaFace. Selanjutnya jika menggunakan variasi dimensi dalam skenario low resolution hasil yang didapatkan sangat bervariasi meskipun terdapat perbedaan rata – rata akurasi dalam proses tugas evaluasi metrik menggunakan bantuan super resolusi. Maka dengan model SR sendiri pada dimensi 64x64 berdasarkan data pada sebelumnya Light CNN9-AdaFace masih unggul sebesar 97,32% dibandingkan arsitektur lainnya, namun pada variasi dimensi 32x32 terjadi penurunan cukup jauh pada Light CNN9-AdaFace yaitu 95.37%.

Pada variasi dimensi 16x16 pun sama Light CNN9-AdaFace justru turun jauh rata – rata akurasinya namun masih bisa menahan di angka 83.37%. Sedangkan jika berdasarkan hasil evaluasi model tanpa menggunakan *identity loss* rata – rata akurasi yang didapatkan cenderung lebih kecil dibandingkan dengan yang menggunakan loss berbasis margin AdaFace. Maka pada proses perubahan resolusi gambar wajah pada dataset LFW secara keseluruhan Light CNN9-AdaFace memiliki performa paling baik jika dinilai berdasarkan asumsi indikator diatas dengan sama - sama menggunakan model super resolution. Hasil bisa dilihat pada Tabel VI.

TABEL VI
PERBANDINGAN PERFORMA MODEL BERDASARKAN EVALUASI VARIASI DIMENSI PADA LFW

ARSITEKTUR	128X128	64X64	32X32	16X16
LIGHT CNN-9 [5]	67.92	67.18	67.43	65.95
LIGHT CNN-9 + SR [5]	67.92	68.95	66.3	65.98
LIGHT CNN-9 + SR + ADAFACE	97.78	97.32	95.37	83.37

Namun berbeda pada dataset TinyFace, Light CNN-9 menggunakan AdaFace disini terlihat lebih unggul dibandingkan dengan lainnya. Pada tabel uji sintesis sebelumnya, terdapat empat jenis peringkat yang diukur, yaitu Rank-1, Rank-5, Rank-10, dan Rank-20. Analisis data menunjukkan bahwa Light CNN-9 memiliki performa yang cukup baik dalam tugas pengenalan wajah. Pada Rank-1, Light CNN-9 mencapai akurasi sebesar 47.26%. Hal ini

menunjukkan bahwa Light CNN-9 memiliki kemampuan dalam mengenali individu pada peringkat teratas.

TABEL VII
PERBANDINGAN PERFORMA MODEL BERDASARKAN EVALUASI VARIASI DIMENSI PADA TINYFACE

ARSITEKTUR	AKURASI TINYFACE			
	128x128	64x64	32x32	16x16
LIGHT CNN-9 + SR + ADAFACE [5]	67.92	67.18	67.43	65.95

Berdasarkan hasil yang diperoleh dapat disimpulkan bahwa penerapan fungsi margin loss AdaFace dengan bantuan model super resolution REAL-ESRGAN pada dataset pengujian dapat meningkatkan akurasi pada perubahan resolusi skenario low resolution dibandingkan arsitektur tanpa fungsi margin loss AdaFace sedangkan untuk skenario penggunaan model super resolution REAL-ESRGAN sendiri secara konsisten memberikan peningkatan akurasi pada berbagai perubahan resolusi dibandingkan dengan tanpa model super resolution REAL-ESRGAN.

IV. KESIMPULAN

Dari keseluruhan tahapan penelitian yang telah dilakukan model pengenalan wajah menggunakan arsitektur Light CNN dengan menerapkan fungsi margin loss AdaFace dapat mengekstrak 512 fitur wajah dari gambar berdimensi 128x128. Hasil tersebut didapatkan akurasi sebesar 97.78% pada dataset LFW dimensi 128x128, sedangkan pada data dimensi kecil yaitu 16x16 didapatkan akurasi sebesar 83.37% dengan bantuan model super resolution REAL-ESRGAN. Pada dataset TinyFace pun didapatkan akurasi Rank-1, Rank-5, Rank-10, dan Rank-20 berturut - turut sebesar 47.26%, 55.25%, 58.61%, 61.90%.

DAFTAR PUSTAKA

- [1] S. Yulina, "Penerapan Haar Cascade Classifier dalam Mendeteksi Wajah dan Transformasi Citra Grayscale Menggunakan OpenCV," 2021. [Online]. Available: <https://jurnal.pcr.ac.id/index.php/jkt/>
- [2] C. Fani, "Perbandingan Identifikasi Wajah Dengan Ekstraksi Fitur Haralick Dan CNN," 2020. [Online]. Available: <http://index.unper.ac.id>
- [3] H. M. Claus, "The Importance of Hyperparameter Optimisation for Facial Recognition Applications," 2022. [Online]. Available: www.aaai.org
- [4] H. Sabilal Rasyad, F. Sthevanie, and A. Arifanto, "Pengenalan Wajah Menggunakan Metode Local Binary Pattern Dan Principal Component Analysis," 2021.
- [5] X. Wu, R. He, Z. Sun, and T. Tan, "A Light CNN for Deep Face Representation with Noisy Labels," Nov. 2018, [Online]. Available: <http://arxiv.org/abs/1511.02683>
- [6] W. Astuti, "Implementasi Metode Super Resolusi Untuk Meningkatkan Kualitas Citra Hasil Screenshot," *JURIKOM (Jurnal Riset Komputer)*, vol. 7, no. 3, p. 432, Jun. 2020, doi: 10.30865/jurikom.v7i3.2129.
- [7] K. S. Zebua, I. Herwidiana Kartowisastro, and G. P. Kusuma, "Low Resolution Face Recognition Using Combination Of Open Super Resolution And Facenet," *J Theor Appl Inf Technol*, vol. 30, no. 12, 2023, [Online]. Available: www.jatit.org
- [8] J. Deng, J. Guo, J. Yang, N. Xue, I. Kotsia, and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," Jan. 2018, doi: 10.1109/TPAMI.2021.3087709.
- [9] M. Kim, A. K. Jain, and X. Liu, "AdaFace: Quality Adaptive Margin for Face Recognition," Apr. 2022, [Online]. Available: <http://arxiv.org/abs/2204.00964>
- [10] X. Wang, L. Xie, C. Dong, and Y. Shan, "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data," 2021. [Online]. Available: <https://github.com/xinntao/Real-ESRGAN>
- [11] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning Face Representation from Scratch," Nov. 2014, [Online]. Available: <http://arxiv.org/abs/1411.7923>
- [12] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments." [Online]. Available: <http://vis-www.cs.umass.edu/lfw/>.
- [13] Z. Cheng, X. Zhu, and S. Gong, "Low-Resolution Face Recognition," Nov. 2018, [Online]. Available: <http://arxiv.org/abs/1811.08965>
- [14] D. E. Kurniawan, K. Adi, and A. F. Rohim, "Sistem Identifikasi Biometrika Wajah Menggunakan Metode Gabor KPCA dan Mahalanobis Distance," *J SINBIS J. Sist. Inf. Bisnis*, vol. 2, no. 1, pp. 006–010, Jan. 2014, doi: 10.21456/vol2iss1pp006-010
- [15] "InsightFace: an open source 2D&3D deep face analysis library." Accessed: Jun. 14, 2023. [Online]. Available: <https://insightface.ai/>