

Rekonstruksi Citra Wajah Resolusi Rendah Menggunakan Peningkatan pada Model FSRCNN dengan *Parallel Feature Extraction* dan *Residual Connections*

Reconstruction of Low-Resolution Facial Images Using Improvements to the FSRCNN Model with Parallel Feature Extraction Layers and Residual Connections

Tommy^{1*}, Rosyidah Siregar², Edy Rahman Syahputra³

^{1,2}Program Studi Teknik Informatika, Fakultas Teknik dan Komputer,

³Program Studi Sistem Informasi, Fakultas Teknik dan Komputer,
Universitas Harapan Medan

Jl. Imam Bonjol, J A T I, Kec. Medan Maimun, Kota Medan, Sumatera Utara 20151, Indonesia

*Corresponding author: tomshirakawa@gmail.com

ABSTRAK

DOI:
[10.30595/jrst.v9i2.24237](https://doi.org/10.30595/jrst.v9i2.24237)

Histori Artikel:

Diajukan:
17/10/2024

Diterima:
28/05/2025

Diterbitkan:
08/09/2025

Peningkatan kualitas rekonstruksi citra wajah beresolusi rendah menjadi tantangan penting dalam berbagai aplikasi, terutama dalam pengawasan berbasis CCTV. Penelitian ini mengembangkan model Fast Super-Resolution Convolutional Neural Network (FSRCNN) dengan penambahan parallel feature extraction layer dan residual connection dari lapisan tersebut ke expanding layer untuk meningkatkan kualitas rekonstruksi wajah dari resolusi rendah. Model yang diusulkan diuji menggunakan dua dataset: CelebA dan rekaman CCTV. Pengujian dilakukan untuk dua skala rekonstruksi, yakni dari LR ke HR dan dari LR ke intermediate. Hasil pengujian menunjukkan bahwa model yang diusulkan secara konsisten memberikan peningkatan performa dibandingkan dengan FSRCNN standar. Pada rekonstruksi LR ke HR, model yang diusulkan menghasilkan peningkatan rata-rata PSNR sebesar 1.24 dB dan SSIM sebesar 0.0584 pada dataset CelebA. Sementara itu, untuk dataset CCTV, peningkatan rata-rata PSNR sebesar 0.52 dB dan SSIM sebesar 0.025 tercapai. Meskipun model ini mampu meningkatkan kualitas citra secara signifikan, beberapa batasan ditemukan dalam hal kinerja pada citra CCTV yang lebih bervariasi dan kompleksitas model yang lebih tinggi.

Kata Kunci: FSRCNN; Rekonstruksi Citra; Super-Resolusi; *Parallel Feature Extraction*; *Residual Connection*

ABSTRACT

Enhancing the quality of reconstructing low-resolution facial images poses a significant challenge in various applications, particularly in CCTV-based surveillance. This study develops a Fast Super-Resolution Convolutional Neural Network (FSRCNN) model with the addition of parallel feature extraction layers and residual connections from these layers to the expanding layer to improve the reconstruction quality of faces from low resolution. The proposed model is tested using two datasets: CelebA and CCTV footage. The tests are conducted for two reconstruction scales, namely from LR to HR and from LR to intermediate. The results demonstrate that the proposed model consistently outperforms the standard FSRCNN. In the LR to HR reconstruction, the proposed model achieves an average PSNR improvement of 1.24 dB and an SSIM increase of 0.0584 on the CelebA dataset. For the CCTV dataset, an average PSNR enhancement of 0.52 dB and an SSIM improvement of 0.025 are attained. While this model significantly enhances image quality, several limitations were identified regarding performance on more varied CCTV images and the higher complexity of the model.

Keywords: *FSRCNN; Image Reconstruction; Super-Resolution; Parallel Feature Extraction; Residual Connection*

1. PENDAHULUAN

Dalam beberapa tahun terakhir, penggunaan sistem pengawasan berbasis CCTV semakin meluas, terutama untuk keperluan keamanan di ruang publik. Namun, keterbatasan kualitas rekaman, terutama citra wajah yang beresolusi rendah, sering kali menghambat proses identifikasi dan pelacakan yang akurat (Cheng et al., 2020) (Satyagama & Widyantoro, 2020).

Citra wajah dengan detail yang kurang, baik karena kondisi pencahayaan yang buruk, jarak kamera, atau kualitas sensor, menjadi tantangan dalam menghasilkan rekonstruksi wajah yang berkualitas. Tantangan ini penting untuk diatasi, mengingat kebutuhan untuk mengidentifikasi individu dengan cepat dan akurat dalam berbagai situasi keamanan.

Teknologi super-resolution berbasis deep learning telah menjadi solusi potensial untuk meningkatkan kualitas citra wajah yang direkam dalam resolusi rendah. Salah satu model yang sering digunakan untuk tugas ini adalah Fast Super-Resolution Convolutional Neural Network (FSRCNN), yang menawarkan kecepatan dan efisiensi dalam meningkatkan resolusi gambar.

FSRCNN sendiri merupakan pengembangan dari Super-Resolution Convolutional Neural Network (SRCNN) yang memiliki arsitektur yang lebih sederhana (Dong et al., 2016). FSRCNN banyak diaplikasikan sampai sekarang seperti yang dilakukan oleh Bai dalam rekonstruksi digital rock (Bai et al., 2020), face detection (Ivanov et al., 2020) dan penelitian – penelitian lainnya.

Namun, versi sederhana dari model ini masih memiliki keterbatasan dalam menangkap detail halus pada citra wajah dengan resolusi rendah. Hal ini menjadi penghalang dalam meningkatkan akurasi pengenalan wajah pada citra dengan resolusi rendah. Peningkatan terhadap arsitektur FSRCNN terus dilakukan dalam mengatasi keterbatasan dari FSRCNN.

Zhang dalam penelitiannya mengaplikasikan jaringan residual pada model FSRCNN yang mampu memberikan hasil rekonstruksi yang lebih baik dengan peningkatan PSNR rata-rata sebesar 0.38 dB (J. Zhang et al., 2021). Passarella dalam penelitiannya menggunakan tambahan convolution layer dalam arsitektur FSRCNN

untuk meningkatkan fitur yang lebih baik dari citra hasil rekonstruksi (Passarella et al., 2022).

Peningkatan terhadap FSRCNN juga dilakukan oleh Saehyun yang dalam penelitiannya melakukan modifikasi terhadap arsitektur FSRCNN sehingga dapat dioptimalkan pada FPGA (Ahn & Kang, 2021). Selain peningkatan terhadap arsitektur FSRCNN, penelitian sebelumnya juga melakukan kombinasi dengan metode lain (Khachatryan et al., 2023).

Berdasarkan penelitian-penelitian sebelumnya, pengembangan arsitektur dari FSRCNN memberikan hasil rekonstruksi yang lebih baik. Pengembangan pada extraction layer merupakan salah satu pengembangan yang cukup memberikan dampak terhadap performa dari model FSRCNN. Zhang pada penelitiannya menggunakan multi scale extraction layer dalam model yang dikembangkannya (F. Zhang et al., 2018).

Begitu juga dengan penelitian yang dilakukan oleh Wang dalam rekonstruksi Low Light Level Image (Wang et al., 2021). Pengembangan arsitektur lainnya pada FSRCNN yang banyak diaplikasikan adalah penggunaan residual yang dapat membantu untuk meningkatkan gradient model dalam proses pelatihan (Hou et al., 2020).

Berdasarkan permasalahan yang ada pada citra wajah resolusi rendah dan keterbatasan FSRCNN, penelitian ini melakukan pengembangan terhadap arsitektur FSRCNN sederhana dengan berfokus pada modifikasi extraction layer dan pengaplikasian teknik residual. Model yang dikembangkan pada penelitian ini menggabungkan lapisan paralel extraction feature dan koneksi residual untuk meningkatkan kualitas rekonstruksi citra wajah.

Pendekatan ini diharapkan dapat memperkaya informasi fitur yang dihasilkan oleh model, sehingga mampu memberikan rekonstruksi wajah yang lebih tajam dan detail dibandingkan dengan model FSRCNN yang sederhana. Penelitian ini bertujuan untuk mengembangkan dan menguji model FSRCNN yang dimodifikasi serta membandingkan performanya dengan model sederhana yang ada.

Model yang dikembangkan akan dievaluasi menggunakan beberapa metrik kualitas gambar, seperti PSNR dan SSIM, serta diuji pada dataset standar dan CCTV untuk

melihat kemampuannya dalam menghadapi data dunia nyata. Hasil dari penelitian ini diharapkan mampu menunjukkan peningkatan signifikan dalam kualitas rekonstruksi citra wajah, yang pada akhirnya dapat membantu meningkatkan akurasi pengenalan wajah pada sistem pengawasan.

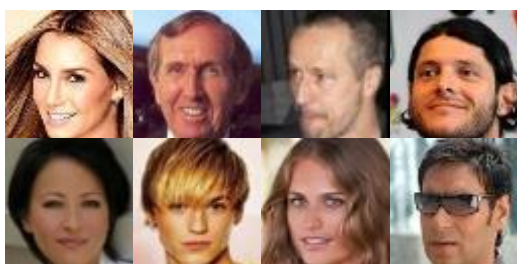
Penelitian ini juga diharapkan berkontribusi pada pengembangan teknologi rekonstruksi citra yang lebih efisien dan efektif, yang bermanfaat dalam aplikasi keamanan dan pengawasan.

2. METODE PENELITIAN

Penelitian ini dilakukan dengan empat tahapan utama yaitu pengumpulan data, perancangan model FSRCNN dengan parallel extraction layer dan koneksi residual, pelatihan model dan terakhir melakukan pengujian dan evaluasi model yang akan dibandingkan dengan FSRCNN sederhana. Pada bagian ini akan dibahas mengenai dataset yang digunakan, perancangan model yang dilakukan, skema pelatihan dan pengujian serta metrik evaluasi yang digunakan dalam pelatihan dan pengujian.

2.1 Dataset

Penelitian ini menggunakan dataset CelebA yang terdiri dari 60.000 citra wajah yang memiliki ukuran 64x64 piksel. Adapun pemilihan CelebA sebagai dataset yang digunakan pada penelitian ini dikarenakan banyaknya penelitian sebelumnya yang memanfaatkan dataset ini (Y. Zhang et al., 2020) (Wu et al., 2023).



Gambar 1. Dataset CelebA

Adapun versi CelebA (Gambar 1) yang digunakan di penelitian ini mengambil dari dataset yang digunakan pada penelitian yang digunakan oleh Bulat (Bulat et al., 2018). Selain menggunakan CelebA, penelitian ini juga menggunakan citra wajah yang diperoleh dari rekaman CCTV (Gambar 2).



Gambar 2. Dataset CCTV

2.2 Model FSRCNN Sederhana

Model FSRCNN sederhana yang digunakan dalam penelitian ini dirancang untuk meningkatkan resolusi citra wajah dari kualitas rendah menjadi kualitas yang lebih tinggi dengan memanfaatkan arsitektur konvolusional. Model ini terdiri dari beberapa lapisan yang berfungsi secara terkoordinasi untuk menghasilkan citra rekonstruksi yang lebih baik.

Arsitektur dasar FSRCNN sederhana terdiri dari lima jenis lapisan utama: extraction layer, shrinking layer, mapping layers, expanding layer, dan deconvolution layer. Adapun penjabaran dari masing-masing layer adalah sebagai berikut (Dong et al., 2016):

1. Extraction Layer: Lapisan ini bertugas untuk mengambil fitur awal dari citra masukan yang beresolusi rendah.
2. Shrinking Layer: Setelah ekstraksi fitur, shrinking layer berfungsi untuk mengurangi dimensi fitur yang dihasilkan, sehingga mengurangi kompleksitas data.
3. Mapping Layers: Terdapat empat mapping layers yang digunakan untuk memproses fitur yang telah dikurangi dimensinya. Masing-masing layer mapping ini berfungsi untuk menangkap pola yang lebih kompleks dari data dengan menerapkan operasi konvolusi, yang memperkaya representasi fitur yang dihasilkan dari citra masukan.
4. Expanding Layer: Setelah pemetaan fitur, expanding layer digunakan untuk memperbesar kembali dimensi fitur ke ukuran yang diinginkan.
5. Deconvolution Layer (Conv2DTranspose): Akhirnya, deconvolution layer berfungsi untuk menghasilkan citra akhir dengan resolusi tinggi dari fitur yang telah diproses.

Model yang digunakan pada penelitian ini dibangun menggunakan Tensorflow pada Bahasa pemrograman Python, Adapun deskripsi arsitektur model FSRCNN sederhana dapat dilihat pada tabel 1.

Tabel 1. Arsitektur FSRCNN Sederhana

Layer	Deskripsi	Parameter
Input Layer	Lapisan Input yang menerima input citra masukan.	input_shape: (None, None, 3)
Feature Extraction	Lapisan konvolusi dengan 56 filter.	filters: 56 kernel_size: 5 activation: ReLU
Shrinking	Lapisan konvolusi dengan 12 filter.	filters: 12 kernel_size: 1 activation: ReLU
Mapping	Terdiri dari 4 lapisan konvolusi, dengan masing – masing lapisan terdiri dari layer dengan 12 filter.	filters: 12 kernel_size: 3 activation: ReLU
Expanding	Lapisan konvolusi dengan 56 filter.	filters: 56 kernel_size: 1 activation: ReLU
Deconvolution	Lapisan konvolusi transpose.	filters: 3 kernel_size: 9 strides: (scale_factor, scale_factor) Padding : same

2.3 Model Pengembangan FSRCNN

Model pengembangan FSRCNN yang digunakan pada penelitian ini menggunakan model FSRCNN sederhana sebagai model dasar yang kemudian ditambahkan dengan parallel extraction layer dan koneksi residu. Pendekatan penambahan parallel extraction layer yang digunakan pada penelitian ini didasari oleh beberapa penelitian sebelumnya yang mana menunjukkan peningkatan dalam ekstraksi fitur dari citra input (Das et al., 2022) (Zhao et al., 2021).

Parallel extraction layer (**Gambar 3**) yang digunakan pada penelitian ini terdiri dari tiga lapisan paralel yang terpisah yang berhubungan langsung dengan lapisan input. Lapisan pertama dari parallel extraction layer merupakan lapisan konvolusi dengan 32 filter dan ukuran kernel 5x5, lapisan kedua merupakan lapisan konvolusi dengan 48 filter dan ukuran kernel 3x3, lapisan ketiga merupakan lapisan konvolusi dengan 64 filter dan kernel 1x1.

Setelah hasil dari parallel extraction layer digabungkan melalui lapisan concatenate, fitur akan diteruskan ke lapisan yang sama dengan lapisan FSRCNN sederhana yaitu dilanjutkan ke lapisan shrinking, mapping dan

seterusnya. Koneksi residual yang digunakan pada penelitian ini bukan lah koneksi residual seperti residual block yang banyak digunakan pada penelitian sebelumnya (Qiu et al., 2021).

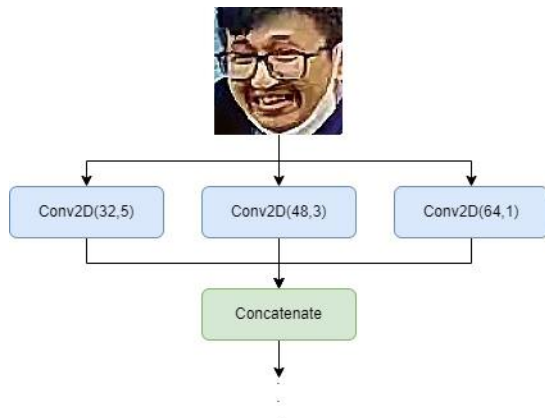
Pada penelitian ini koneksi residual menghubungkan fitur hasil dari lapisan concatenate dengan luaran dari lapisan expanding. Dengan menggunakan pendekatan koneksi residual ini, fitur-fitur penting yang diperoleh dari parallel extraction layer tetap terjaga pada saat proses upscaling dan meningkatkan aliran gradient pada saat pelatihan.

2.4 Skema Pelatihan

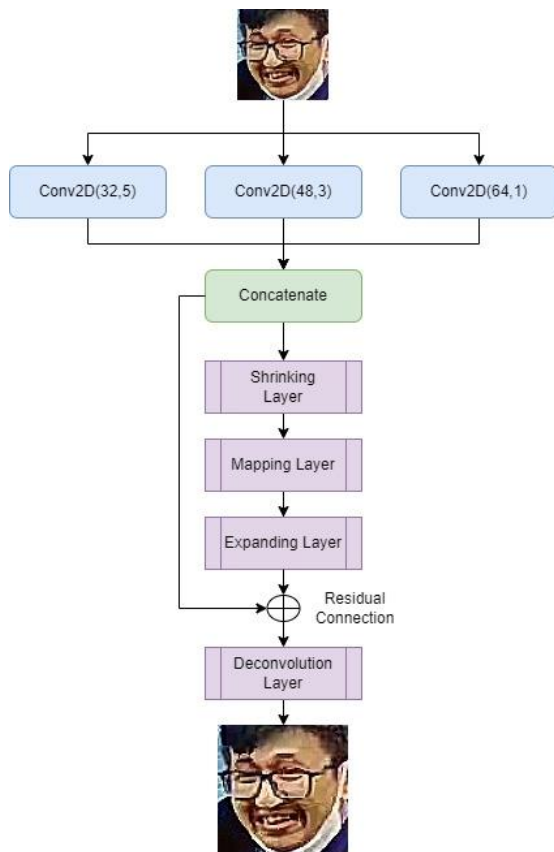
Model FSRCNN sederhana dan model FSRCNN yang dilengkapi dengan parallel extraction layer dan residual connection (**Gambar 4**) akan dilatih menggunakan dataset CelebA yang mana dataset akan dibagi menjadi tiga jenis ukuran yaitu x1 (64x64), x2 (32x32) dan x4 (16x16). Adapun proses downscaling menggunakan bicubic interpolation yang dapat mempertahankan banyak detail dari gambar seiring proses downscaling (Niu et al., 2020).

Adapun parameter pelatihan untuk kedua model tersebut menggunakan jumlah epoch sebesar 200 epoch, batch size 32 dan

learning rate sebesar 0.00001. Kedua model akan dilatih sebanyak dua kali yaitu pelatihan rekonstruksi dari resolusi x4 ke x2 dan pelatihan rekonstruksi dari x4 ke x1. Hal ini bertujuan untuk mengamati performa dari kedua model dalam melakukan rekonstruksi dari Low Resolution ke Intermediate Resolution dan dari Low Resolution ke High resolution.



Gambar 3. Parallel Extraction Layer



Gambar 4. Residual Connection

2.5 Skema Pengujian

Model yang telah dilatih kemudian akan di uji untuk melakukan rekonstruksi terhadap

dataset citra uji menggunakan dataset CelebA dan dataset CCTV. Pengujian rekonstruksi akan dilakukan pada rekonstruksi dari x4 ke x2 dan dari x4 ke x1. Pengujian terhadap dataset CelebA dan dataset CCTV akan dilakukan secara terpisah agar dapat diamati kemampuan dari kedua model terhadap citra umum seperti CelebA dan citra dari keadaan nyata yaitu dari citra CCTV. Adapun jumlah citra uji akan menggunakan 1.000 citra uji dari CelebA dan 1.000 citra uji dari dataset CCTV.

2.6 Metrik Evaluasi

Metrik evaluasi yang digunakan pada penelitian ini adalah metrik PSNR dan SSIM. Berdasarkan penelitian – penelitian sebelumnya, PSNR dan SSIM merupakan metrik evaluasi yang dapat menggambarkan ukuran kualitas dari citra hasil rekonstruksi terhadap citra aslinya (Ground Truth Image) (Li et al., 2022). Adapun PSNR dapat di-formulasikan sebagai berikut:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX^2}{MSE} \right) \quad (1)$$

Dimana MAX = 255 untuk citra 8-bit, dan MSE merupakan mean squared error antara citra asli dengan citra hasil rekonstruksi. SSIM dapat diformulasikan sebagai berikut:

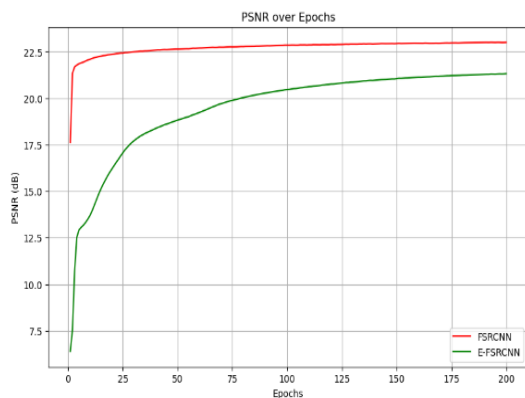
$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (2)$$

Dimana x dan y adalah citra input, μ_x dan μ_y adalah rata-rata local, σ_x^2 dan σ_y^2 adalah nilai variance, σ_{xy} adalah nilai covariance, dan C_1 dan C_2 adalah konstanta.

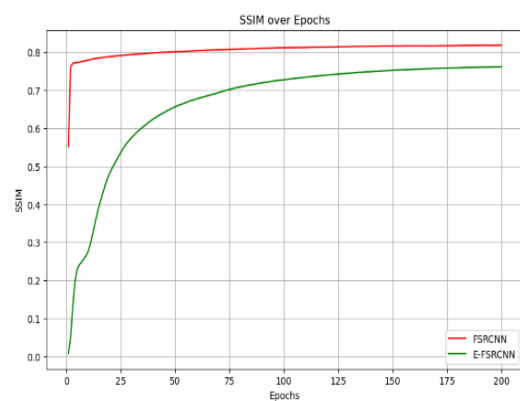
3. HASIL DAN PEMBAHASAN

Tahapan awal dalam melihat bagaimana performa pengembangan model yang dilakukan pada penelitian ini adalah dengan melakukan pelatihan terhadap model FSRCNN sederhana dan model pengembangan FSRCNN yang dapat disingkat dengan E-FSRCNN. Pelatihan dilakukan menggunakan dataset CelebA dengan dua jenis rekonstruksi yaitu x4 ke x2 dan x4 ke x1.

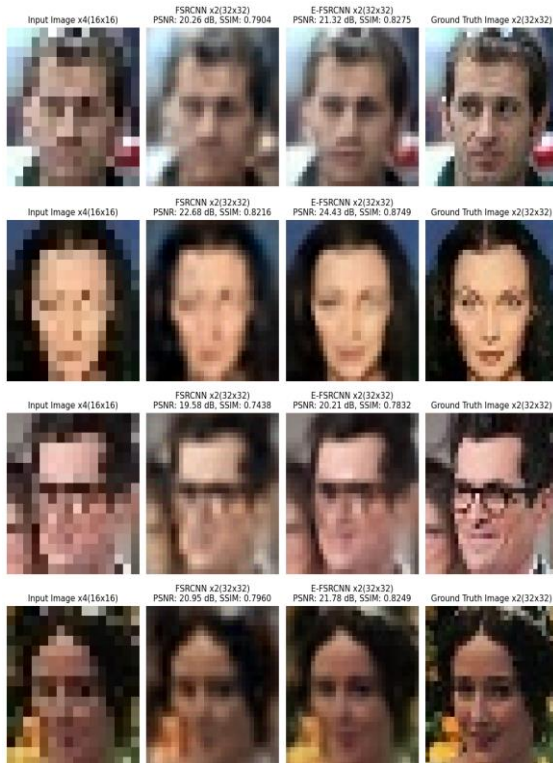
Performa pelatihan model seperti yang terlihat pada Gambar 5 dan Gambar 6 menunjukkan bahwa skor PSNR dan SSIM dari model yang dikembangkan (E-FSRCNN) lebih baik dibandingkan dengan model FSRCNN sederhana. Di awal epoch, FSRCNN sederhana memiliki skor PSNR sebesar 6.39 dB dan terus meningkat sampai 21.32 dB di akhir epoch. Untuk skor SSIM, model FSRCNN sederhana memiliki skor sebesar 0.0078 dan terus meningkat hingga 0.7618 di akhir epoch.



Gambar 5. Kurva PSNR Pelatihan x4 ke x2



Gambar 6. Kurva SSIM Pelatihan x4 ke x2



Gambar 7. Hasil rekonstruksi CelebA FSRCNN dan E-FSRCNN (x4 ke x2)



Gambar 8. Hasil rekonstruksi Dataset CCTV FSRCNN dan E-FSRCNN (x4 ke x2)

Skor FSRCNN sederhana lebih rendah dibandingkan pada model FSRCNN yang dikembangkan dimana pada awal *epoch* model FSRCNN yang dikembangkan mampu memiliki skor PSNR sebesar 17.62 dB dan terus meningkat hingga mencapai 23.0021 dB di akhir *epoch*. Sedangkan untuk skor SSIM, model yang dikembangkan memiliki skor sebesar 0.5522 di awal *epoch* dan meningkat hingga 0.8183 di akhir *epoch*.

Pengujian rekonstruksi pada dataset CCTV seperti yang terlihat pada Gambar 8 juga menunjukkan bahwa model FSRCNN yang dikembangkan (E-FSRCNN) memberikan hasil rekonstruksi yang lebih baik dibandingkan dengan model FSRCNN sederhana.

Hasil rekonstruksi dari kedua model seperti yang terlihat pada Gambar 7 menunjukkan hasil rekonstruksi citra wajah dari model FSRCNN pada berbagai tahap. Pada kolom paling kiri, terlihat citra input yang memiliki resolusi rendah dengan ukuran 16×16 piksel. Citra ini menunjukkan detail yang sangat kabur dan kehilangan banyak informasi penting, sehingga menyulitkan untuk mengidentifikasi fitur wajah dengan jelas.

Pada kolom kedua, dapat dilihat hasil rekonstruksi dari model FSRCNN yang dilatih untuk meningkatkan resolusi dari 16×16 menjadi 32×32 piksel. Meskipun ada peningkatan, hasil rekonstruksi ini masih menunjukkan beberapa keaburan. Hal ini menunjukkan bahwa meskipun model dapat

menambah detail pada citra, masih ada ruang untuk perbaikan dalam menghasilkan gambar yang lebih tajam dan informatif.

Di kolom ketiga, ditampilkan hasil dari model FSRCNN yang lebih ditingkatkan, yang juga merekonstruksi citra dari 16×16 ke 32×32 piksel. Dimana model ini menunjukkan kinerja yang lebih baik dalam hal kualitas gambar yang dihasilkan. Peningkatan dalam PSNR dan SSIM ini mencerminkan kemampuan model untuk menangkap dan memulihkan fitur wajah yang lebih jelas dan tajam dibandingkan dengan model sebelumnya.

Adapun rekapitulasi pengujian pada rekonstruksi x4 ke x2 pada model FSRCNN dan E-FSRCNN pada dataset CelebA dan CCTV seperti yang terlihat pada **Tabel 2** dan **Tabel 3** menunjukkan bahwa secara umum model E-FSRCNN memiliki skor yang lebih baik daripada FSRCNN sederhana. Pelatihan dan pengujian kemudian dilanjutkan pada rekonstruksi x4 ke x1 untuk melihat performa kedua model dalam proses rekonstruksi dari citra LR ke citra HR. Pada pelatihan rekonstruksi x4 ke x1 baik pada PSNR dan SSIM seperti yang terlihat pada **Gambar 9** dan **Gambar 10** memiliki pergerakan yang mirip dengan pelatihan rekonstruksi x4 ke x2.

Adapun dari hasil pelatihan dapat dilihat model pengembangan FSRCNN (E-FSRCNN) memiliki nilai akhir yang lebih baik dimana pengembangan FSRCNN memiliki nilai PSNR akhir sebesar 22.5055 dan SSIM sebesar 0.6937 dibandingkan dengan FSRCNN sederhana yang memperoleh PSNR akhir sebesar 20.9774 dan SSIM sebesar 0.6245.

Pengujian kemudian dilakukan terhadap kedua model yang telah dilatih menggunakan dataset pengujian dari CelebA dan CCTV pada resolusi x4 ke x1. Hasil rekonstruksi dari CelebA seperti yang dapat dilihat pada **Gambar 11** dan **Gambar 12** menunjukkan bahwa hasil rekonstruksi menggunakan pengembangan FSRCNN (E-FSRCNN) memberikan hasil yang lebih baik dimana selain dari ukuran PSNR dan SSIM yang diperoleh, secara visual fitur wajah hasil rekonstruksi dari E-FSRCNN lebih terlihat dibandingkan dengan hasil rekonstruksi dari FSRCNN sederhana.

Secara umum, model E-FSRCNN menunjukkan perbaikan yang lebih signifikan dalam hal kualitas rekonstruksi dibandingkan dengan FSRCNN sederhana. Hal ini terlihat dari nilai PSNR dan SSIM yang lebih tinggi di sebagian besar gambar, menandakan bahwa model E-

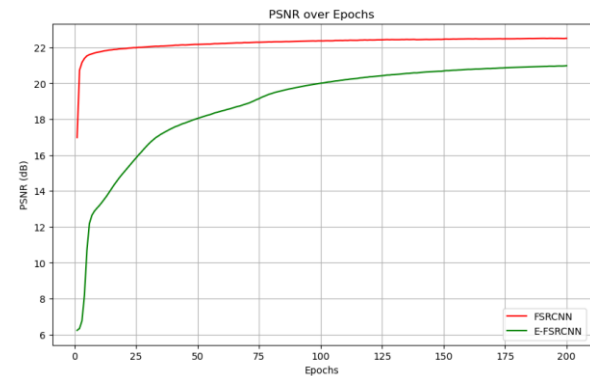
FSRCNN mampu mempertahankan struktur dan detail wajah lebih baik.

Tabel 2. PSNR Hasil Pengujian x4 ke x2

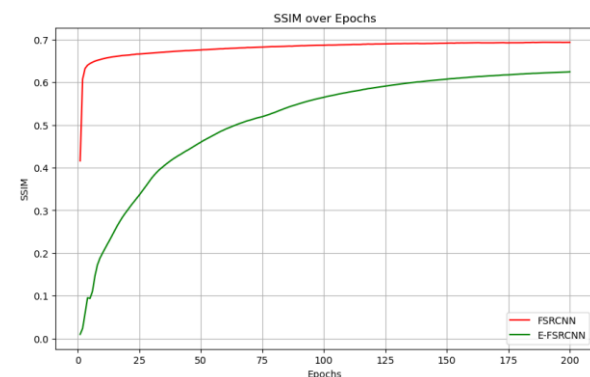
Dataset	Model	PSNR		
		Min	Max	Avg
CelebA	FSRCNN	15.79	29.49	21.39
	E-FSRCNN	16.7	31.2	22.74
CCTV	FSRCNN	12.31	34.18	22.6
	E-FSRCNN	12.35	35.58	23.14

Tabel 3. SSIM Hasil Pengujian x4 ke x2

Dataset	Model	SSIM		
		Min	Max	Avg
CelebA	FSRCNN	0.5825	0.9047	0.7953
	E-FSRCNN	0.6701	0.9329	0.8371
CCTV	FSRCNN	0.2754	0.9379	0.7688
	E-FSRCNN	0.2799	0.9519	0.7837



Gambar 9. Kurva PSNR Pelatihan x4 ke x1



Gambar 10. Kurva SSIM Pelatihan x4 ke x1

Sebagai contoh, pada citra kedua di baris pertama, FSRCNN menghasilkan PSNR sebesar 19.06 dB dan SSIM sebesar 0.6850, sementara model E-FSRCNN mencapai PSNR 19.95 dB dan SSIM 0.7099, menunjukkan peningkatan yang

nyata dalam kualitas rekonstruksi. Hasil rekonstruksi pada dataset yang digunakan mengindikasikan bahwa modifikasi yang dilakukan pada arsitektur E-FSRCNN, seperti penambahan lapisan paralel feature extraction dan koneksi residual, membantu meningkatkan performa model dalam merekonstruksi citra wajah dari resolusi rendah ke resolusi yang lebih tinggi.



Gambar 11. Hasil rekonstruksi Dataset CelebA FSRCNN dan E-FSRCNN (X4 ke x1)



Gambar 12. Hasil rekonstruksi Dataset CCTV FSRCNN dan E-FSRCNN (X4 ke x1)

Tabel 4. PSNR Hasil Pengujian x4 ke x1

Dataset	Model	PSNR		
		Min	Max	Avg
CelebA	FSRCNN	15.03	29.57	20.99
	E-FSRCNN	15.75	30.99	22.23
CCTV	FSRCNN	12.81	32.35	20.63
	E-FSRCNN	12.43	33.41	21.15

Tabel 5. SSM Hasil Pengujian x4 ke x1

Dataset	Model	SSIM		
		Min	Max	Avg
CelebA	FSRCNN	0.4775	0.8138	0.6566
	E-FSRCNN	0.54	0.8512	0.715
CCTV	FSRCNN	0.2683	0.8771	0.6343
	E-FSRCNN	0.2803	0.8721	0.6593

Model E-FSRCNN pada sebagian besar hasil rekonstruksi menghasilkan PSNR yang lebih tinggi dibandingkan FSRCNN sederhana, seperti pada gambar kedua dan ketiga yang mencapai PSNR 23.42 dB (E-FSRCNN) dibandingkan dengan 23.20 dB (FSRCNN). Model E-FSRCNN umumnya memberikan nilai SSIM yang lebih tinggi, menunjukkan bahwa model ini lebih baik dalam mempertahankan struktur dan detail wajah. Pada gambar kedua, E-FSRCNN mencapai SSIM 0.6805 dibandingkan dengan FSRCNN yang hanya mencapai 0.6716.

Pada dataset CelebA, E-FSRCNN secara konsisten memberikan hasil rekonstruksi yang lebih baik dibandingkan FSRCNN. Peningkatan rata-rata PSNR sebesar 1.24 dB menunjukkan kemampuan E-FSRCNN dalam memperbaiki detail citra wajah beresolusi rendah dengan lebih baik. Pada dataset CCTV, model E-FSRCNN juga memberikan peningkatan dibandingkan FSRCNN.

Meskipun PSNR minimum sedikit lebih rendah di E-FSRCNN, rata-rata PSNR lebih tinggi sebesar 0.52 dB, menunjukkan bahwa model ini mampu menangani citra CCTV dengan lebih baik, terutama dalam memulihkan detail citra wajah di lingkungan yang bervariasi.

Pada dataset CelebA, E-FSRCNN memberikan rekonstruksi yang lebih akurat dibandingkan FSRCNN, dengan peningkatan rata-rata SSIM sebesar 0.0584. Ini menunjukkan bahwa E-FSRCNN mampu menjaga struktur citra wajah dengan lebih baik selama proses super-resolusi. Pada dataset CCTV, meskipun nilai SSIM maksimum sedikit lebih rendah untuk E-FSRCNN Tabel 4, rata-rata SSIM Tabel 5 lebih tinggi sebesar 0.025, menunjukkan peningkatan kemampuan E-FSRCNN dalam menjaga kesamaan struktur pada citra wajah yang direkonstruksi, terutama di lingkungan yang bervariasi seperti yang tercermin dalam data CCTV.

4. SIMPULAN

Model FSRCNN yang dikembangkan berhasil meningkatkan kualitas rekonstruksi

citra dibandingkan dengan FSRCNN sederhana. Hal ini terlihat dari peningkatan yang signifikan pada nilai PSNR dan SSIM baik untuk dataset CelebA maupun CCTV. Hal ini dapat dilihat pada nilai PSNR meningkat rata-rata sekitar 1.24 dB pada rekonstruksi dari x4 ke x1 dan 1.35 dB pada rekonstruksi dari x4 ke x2 pada dataset CelebA.

Sedangkan pada nilai SSIM meningkat rata-rata sekitar 0.0584 pada rekonstruksi dari x4 ke x1 dan 0.0418 pada rekonstruksi dari x4 ke x2 pada dataset CelebA, menunjukkan kualitas struktur yang lebih baik dalam citra hasil rekonstruksi.

Dengan input gambar yang sangat rendah (16x16 piksel), baik untuk dataset CelebA maupun CCTV, model FSRCNN yang dikembangkan mampu menghasilkan citra yang lebih jelas dan dengan detail yang lebih baik, yang mendekati citra ground truth dengan resolusi yang lebih tinggi (64x64 atau 32x32).

Hal ini membuktikan bahwa pendekatan pengembangan arsitektur FSRCNN menggunakan *parallel extraction layer* dan koneksi residual memberikan dampak yang baik. Meskipun model FSRCNN yang dikembangkan menunjukkan peningkatan kinerja pada dataset CCTV, hasil PSNR dan SSIM masih lebih bervariasi dibandingkan dengan dataset CelebA.

Hal ini disebabkan oleh perbedaan karakteristik antara kedua dataset, seperti kondisi pencahayaan yang tidak merata, variasi wajah yang lebih besar, dan gangguan lain yang umum dalam rekaman CCTV. Oleh karena itu, model masih memiliki batasan dalam menghadapi citra dengan kondisi yang lebih menantang seperti ini.

Penelitian selanjutnya dapat memperluas cakupan skala resolusi untuk mengevaluasi batas performa dari arsitektur FSRCNN yang dikembangkan. Hal ini akan memberikan gambaran yang lebih lengkap tentang kemampuan generalisasi model.

Pengembangan arsitektur pada model yang dikembangkan juga masih dapat dikembangkan lebih lanjut untuk meningkatkan kualitas dari hasil rekonstruksi tapi tetap menjaga performa kecepatan model.

UCAPAN TERIMA KASIH

Kami ingin menyampaikan terima kasih yang sebesar-besarnya kepada Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia (Direktorat Jenderal Pendidikan Tinggi Kementerian Pendidikan dan Kebudayaan Republik Indonesia) atas pendanaan dan sumber daya yang diperlukan

untuk melaksanakan penelitian ini. Kami juga ingin mengucapkan terima kasih kepada LLDIKTI Wilayah I atas bimbingan dan bantuan mereka sepanjang proses penelitian. Akhirnya, kami sampaikan rasa terima kasih yang tulus kepada Universitas Harapan Medan yang telah menyediakan fasilitas penelitian dan lingkungan akademik yang kolaboratif.

REFERENSI

- Ahn, S., & Kang, S.-J. (2021). Deep Learning-based Real-Time Super-Resolution Architecture Design. *Journal of Broadcast Engineering*, 26(2), 167–174. <https://doi.org/10.5909/JBE.2021.26.2.167>
- Bai, Y., Berezovsky, V., & Popov, V. (2020). Super Resolution for Digital Rock Core Images via FSRCNN. *Proceedings of the 2020 4th High Performance Computing and Cluster Technologies Conference & 2020 3rd International Conference on Big Data and Artificial Intelligence*, 78–81. <https://doi.org/10.1145/3409501.3409528>
- Bulat, A., Yang, J., & Tzimiropoulos, G. (2018). To learn image super-resolution, use a gan to learn how to do image degradation first. *Proceedings of the European Conference on Computer Vision (ECCV)*, 185–200. https://doi.org/10.1007/978-3-030-01231-1_12
- Cheng, Z., Zhu, X., & Gong, S. (2020). Face re-identification challenge: Are face recognition models good enough? *Pattern Recognition*, 107, 107422. <https://doi.org/https://doi.org/10.1016/j.patcog.2020.107422>
- Das, S. K., Roy, P., & Mishra, A. K. (2022). DFU_SPNet: A stacked parallel convolution layers based CNN to improve Diabetic Foot Ulcer classification. *ICT Express*, 8(2), 271–275. <https://doi.org/https://doi.org/10.1016/j.ict.2021.08.022>
- Dong, C., Loy, C. C., & Tang, X. (2016). Accelerating the super-resolution convolutional neural network. *Computer Vision--ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*, 391–407. https://doi.org/10.1007/978-3-319-46475-6_25

- Hou, J., Si, Y., & Yu, X. (2020). A novel and effective image super-resolution reconstruction technique via fast global and local residual learning model. *Applied Sciences*, 10(5), 1856.
<https://doi.org/10.3390/app10051856>
- Ivanov, N. S., Arzhskov, A. V., & Ivanenko, V. G. (2020). Combining Deep Learning and Super-Resolution Algorithms for Deep Fake Detection. *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 326–328.
<https://doi.org/10.1109/EIConRus49466.2020.9039498>
- Khachatryan, T., Galstyan, D., & Harutyunyan, E. (2023). A Comprehensive Approach for Enhancing Deep Learning Datasets Quality Using Combined SSIM Algorithm and FSRCNN. *2023 IEEE East-West Design & Test Symposium (EWDTS)*, 1–4.
<https://doi.org/10.1109/EWDTS59469.2023.10297040>
- Li, Z., Liu, Y., Chen, X., Cai, H., Gu, J., Qiao, Y., & Dong, C. (2022). Blueprint Separable Residual Network for Efficient Image Super-Resolution. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 832–842.
<https://doi.org/10.1109/CVPRW56347.2022.00099>
- Niu, Y., Weng, H., Lin, J., & Liu, G. (2020). Single Image Super-Resolution: From Discrete to Continuous Scale Without Retraining. *IEEE Access*, 8, 32121–32136.
<https://doi.org/10.1109/ACCESS.2020.2973283>
- Passarella, L. S., Mahajan, S., Pal, A., & Norman, M. R. (2022). Reconstructing high resolution ESM data through a novel fast super resolution convolutional neural network (FSRCNN). *Geophysical Research Letters*, 49(4), e2021GL097571.
<https://doi.org/10.1029/2021GL097571>
- Qiu, D., Zheng, L., Zhu, J., & Huang, D. (2021). Multiple improved residual networks for medical image super-resolution. *Future Generation Computer Systems*, 116, 200–208.
<https://doi.org/https://doi.org/10.1016/j.future.2020.11.001>
- Satyagama, P., & Widyantoro, D. H. (2020). Low-Resolution Face Recognition System Using Siamese Network. *2020 7th International Conference on Advance Informatics: Concepts, Theory and Applications (ICAICTA)*, 1–6.
<https://doi.org/10.1109/ICAICTA49861.2020.9428885>
- Wang, B., Zou, Y., Zhang, L., Hu, Y., Yan, H., Zuo, C., & Chen, Q. (2021). Low-Light-Level Image Super-Resolution Reconstruction Based on a Multi-Scale Features Extraction Network. *Photonics*, 8(8).
<https://doi.org/10.3390/photonics8080321>
- Wu, H., Bezold, G., Günther, M., Boulton, T., King, M. C., & Bowyer, K. W. (2023). Consistency and Accuracy of CelebA Attribute Values. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 3258–3266.
<https://doi.org/10.1109/CVPRW59228.2023.00328>
- Zhang, F., Cai, N., Cen, G., Li, F., Wang, H., & Chen, X. (2018). Image super-resolution via a novel cascaded convolutional neural network framework. *Signal Processing: Image Communication*, 63, 9–18.
<https://doi.org/https://doi.org/10.1016/j.image.2018.01.009>
- Zhang, J., Liu, M., Wang, X., & Cao, C. (2021). Residual Net Use on FSRCNN for Image Super-Resolution. *2021 40th Chinese Control Conference (CCC)*, 8077.
<https://doi.org/10.23919/CCC52363.2021.9550205>
- Zhang, Y., Yin, Z., Li, Y., Yin, G., Yan, J., Shao, J., & Liu, Z. (2020). CelebA-Spoof: Large-Scale Face Anti-spoofing Dataset with Rich Annotations. In A. Vedaldi, H. Bischof, T. Brox, & J.-M. Frahm (Eds.), *Computer Vision -- ECCV 2020* (pp. 70–85). Springer International Publishing.
https://doi.org/10.1007/978-3-030-58610-2_5
- Zhao, Z., Li, Q., Zhang, Z., Cummins, N., Wang, H., Tao, J., & Schuller, B. W. (2021). Combining a parallel 2D CNN with a self-attention Dilated Residual Network for CTC-based discrete speech emotion recognition. *Neural Networks*, 141, 52–60.
<https://doi.org/https://doi.org/10.1016/j.neunet.2021.03.013>