Journal of Applied Informatics and Computing (JAIC)
Vol.7, No.1, Juli 2023, pp. 57~62
e-ISSN: 2548-6861

Peningkatan Deteksi Kecelakaan di Jalan Raya Menggunakan Real-ESRGAN pada Citra CCTV Persimpangan Jalan

Muhammad Fachry Ikhsal [1]*, Budi Arif Dermawan [2]*, Riza Ibnu Adam [3]*
*Informatika, Universitas Singaperbangsa Karawang
1910631170203@student.unsika.ac.id [1], budi.arif@staff.unsika.ac.id [2], riza.adam@staff.unsika.ac.id [3]

Article Info

Article history:

Received 2023-06-10
Revised 2023-07-29
Accepted 2023-07-30

Keyword:

Transfer Learning,
MobileNetV2,
Real-ESRGAN,
Accident Detection.

ABSTRACT

The failure of the accident detection system on CCTV cameras can affect the increase in the death rate on the highway. The use of the CNN method in the construction of CCTV accident detection systems has been widely used before. However, common problems that are often encountered are dirty lenses and varifocal zooms that don't automatically focus, causing the quality of the resulting CCTV images to decrease, thus affecting system performance. In this research, a model was developed to detect accidents on CCTV images using the MobileNetV2 pre-trained model which was optimized by upscaling the dataset using the Real-ESRGAN model to produce more optimal performance. This study uses a CCTV image dataset totaling 989 and consisting of 2 types of prediction classes including accident and non-accident. The results showed that the MobileNetV2 model succeeded in producing 94% testing accuracy and an average inference time of 3.33 seconds in the GT test scenario. During the testing process, it was found that the model was not optimal if it identified new data with clustered objects. In addition, based on the test scenarios X2, X4, X8 it was found that the image quality calculated based on PSNR and SSIM values greatly influences classification performance such as accuracy, precision, recall, and AUC score.

This is an open access article under the CC–BY-SA license.

I. PENDAHULUAN

Pada tahun 2022 dikutip dari detik.com Korlantas Polri mencatat di Indonesia kecelakaan lalu lintas masih cukup tinggi. Tercatat kasus kecelakaan lalu lintas sepanjang 2022 yaitu 6.707 kasus dengan koban meninggal dunia sebanyak 452 orang [1]. Di negara padat penduduk seperti India 80% dari korban jiwa bukan disebabkan oleh kecelakaan itu sendiri tapi akibat lambatnya penanganan dalam menangani korban kecelakaan [2]. Karena itulah diperlukan sebuah sistem yang dapat mendeteksi kecelakaan melalui kamera pengawas seperti CCTV. Salah satu metode yang biasa digunakan untuk mendeteksi kecelakaan adalah menggunakan metode *deep learning* dengan cara mengambil tiap *frame* dari video CCTV dan mengklasifikasi *frame* dari video ke kategori *frame* kecelakaan dan *frame* bukan kecelakan.

Namun masalah umum yang sering didapatkan adalah kualitas citra CCTV dapat mengalami penurunan akibat beberapa faktor seperti lensa yang kotor, kabel yang terlalu panjang, dan *zoom varifocal* yang tidak fokus secara otomatis. Penurunan kualitas tersebut dapat berupa *noise*, *blur*, dan *distortion*. Hal ini dapat berpengaruh pada performa sistem dalam mendeteksi kecelakaan dalam sebuah *frame*. Seperti yang dibahas pada penelitian yang dilakukan oleh Ng dan Lo pada tahun 2019, pada penelitian tersebut kualitas gambar yang buruk menurunkan performa model sebesar 7% pada *task* klasifikasi *gender* dan 15% pada klasifikasi umur [3]. Di sisi lain, waktu komputasi yang dibutuhkan model untuk mendeteksi objek tidak kalah penting dan harus dijadikan bahan pertimbangan dalam pembangunan model.

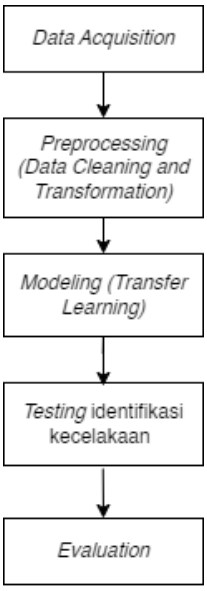
Penelitian tentang *Object Detection* telah banyak dilakukan sebelumnya. Pada tahun 2022 Triana dan Enri melakukan penelitian untuk mendeteksi kecelakaan di jalan raya dengan metode CNN dan berhasil mendapatkan akurasi validasi sebesar 90% [4]. Namun, penelitian tersebut belum melakukan testing dan belum mempertimbangkan waktu *inference*. Di samping itu penelitian masih menggunakan *layer* CNN sederhana tanpa menggunakan arsitektur yang spesifik sehingga kurang efisien. Selanjutnya, pada tahun

2022 Pardede dan Hardiansyah melakukan penelitian pada objek kereta api dengan metode VGG 16 berhasil menghasilkan akurasi 97% pada objek gerbong dan 79% pada objek lokomotif. Dibandingkan arsitektur CNN sederhana arsitektur VGG dinilai lebih efektif dalam hal performa. Namun, pada penelitian tidak dijelaskan waktu komputasi yang dibutuhkan model untuk melakukan *inference* [5]. Selanjutnya, pada penelitian yang dilakukan oleh Hartanto dan Wibowo berhasil menghasilkan akurasi sebesar 86% pada model MobileNetV2. Namun, tidak jelaskan waktu yang dibutuhkan model untuk melakukan *Inference* [6]. Meskipun begitu, MobileNetV2 dikenal sebagai arsitektur yang *lightweight* atau ringan dan didesain untuk dapat bekerja secara efisien pada *embedded system* seperti kamera CCTV. MobileNetV2 didasarkan pada penggunaan Inverted Residuals Blocks dengan Linear Bottlenecks yang menggunakan 3 lapisan konvolusi. Lapisan tersebut berfungsi untuk mengekstraksi fitur-fitur penting dari data dengan mengurangi dan meningkatkan dimensi fitur. MobileNetV2 juga menggunakan fungsi aktivasi linear pada lapisan bottleneck, menggantikan ReLU yang digunakan dalam MobileNetV1. Fungsi aktivasi linear membantu dalam mengatasi masalah *vanishing gradient* dan memperkuat propagasi sinyal melalui jaringan. Secara keseluruhan MobileNetV2 dapat mengurangi jumlah parameter sambil mempertahankan tingkat akurasi [7]. Pada penelitian yang dilakukan oleh Indraswari et al pada tahun 2021 MobileNetV2 berhasil menghasilkan akurasi tertinggi yaitu di atas 85% dan performa komputasi tertinggi dengan *running time* 2,394 s jika dibandingkan model lain seperti VGG 16, ResNet, dan Inception [8].

Berdasarkan pembahasan tersebut, penelitian yang telah dilakukan berupa pembangunan model klasifikasi untuk mendeteksi kecelakaan di jalan raya menggunakan arsitektur MobileNetV2 yang dilakukan optimasi dengan melakukan *upsampling* pada dataset untuk menghasilkan performa yang lebih optimal. Pada penelitian ini, akan dilakukan beberapa skenario uji model klasifikasi menggunakan model MobileNetV2 pada resolusi citra yang bervariasi untuk mengetahui pengaruh kualitas citra terhadap performa model, selanjutnya juga dihitung waktu yang dibutuhkan model untuk melakukan *inference*.

II. METODE PENELITIAN

Pada penelitian ini akan dilakukan beberapa langkah penelitian (Gambar 1) dimulai dari tahap *data acquisition* atau pengumpulan data, *preprocessing* untuk memastikan data yang digunakan bersih dan sesuai dengan tujuan penelitian, *modeling* dengan menggunakan teknik *transfer learning* pada model MobileNetV2, lalu akan dilakukan *testing* dan evaluasi model.



Gambar 1. Diagram Langkah Penelitian. Penelitian dilakukan dari tahap *data acquisition* sampai tahap *evaluation*.

III. HASIL DAN PEMBAHASAN

A. Data Acquisition

Pada tahap ini dilakukan pengumpulan data yang relevan dengan permasalahan. Pada penelitian ini data yang digunakan adalah data sekunder yang berasal dari platform Kaggle. Data tersebut berisi kumpulan citra CCTV jalan raya yang berasal dari dua kategori yaitu *accident* (Gambar 2) dan *non-accident* (Gambar 3).

Gambar 2. Citra *Accident*

Gambar 3. Citra *Non-Accident.*

Lalu pada (Tabel 1) didapatkan hasil akuisisi citra dengan 2 kategori atau kelas, jumlah data *accident* yaitu 462 dengan rasio 53% dan jumlah data *non-accident* yaitu 527 dengan rasio 46%.

TABEL 1
JUMLAH DATA PADA TIAP KELAS

Nama Kelas	Rasio	Jumlah
Accident	46%	462
Non accident	53%	527

Lalu data mentah yang digunakan juga sudah dipisah dengan jumlah dan rasio seperti pada (Tabel 2). Data *train* memilki rasio 80 % dengan jumlah data 791, data *validation* memilki rasio 10% dengan jumlah data 98 dan jumlah data *test* memilki rasio 10% dengan jumlah data 100.

TABEL 2
JUMLAH DATA PADA TIAP *SPLIT*

Tipe split	Rasio	Jumlah
Train	80%	791
Validation	10%	98
Test	10%	100

B. Preprocessing

Pada tahap ini dilakukan dua proses utama yaitu *data cleaning* dan *data transformation*. Pada tahap *data cleaning* dilakukan penghapusan data yang duplikat, tidak relevan, label yang salah, dan data yang tidak memilki resolusi 1280x720. Karena penghapusan data mempengaruhi jumlah data yang telah di *split*, maka dilakukan *split kembali* (Tabel 3) untuk mempertahankan rasio pemisahan 80%, 10%, 10% pada data *train*, *validation*, dan *test*.

TABEL 3
JUMLAH DATA PADA TIAP SPLIT SETELAH PROSES *CLEANING*

Tipe split	Rasio	Jumlah
Train	80%	663
Validation	10%	83
Test	10%	82

Selain itu, pada tahap *data transformation* dilakukan proses *downscaling* (Tabel 4) untuk memperoleh citra *low resolution* (LR) dan *upscaling* untuk memperoleh citra super resolution (SR). Proses ini bertujuan untuk menghasilkan kualitas citra yang bervariasi yang diukur dengan skor metrik PSNR dan SSIM untuk tujuan penelitian [9]. Proses *upscaling* dilakukan dengan menggunakan model Real-ESRGAN yang berbeda tergantung dengan faktor *upscaling* yang sesuai.

Pada *upscaling* dan *downscaling* dengan faktor 2 citra akan di *downscaling* untuk memperoleh citra LR dengan resolusi 640x360 yang kemudian dilakukan *upscaling* dengan model Real-ESRGAN_x2 untuk memperoleh citra SR dengan resolusi 1280x720, pada *upscaling* dan *downscaling* dengan faktor 4 citra akan di *downscaling* untuk memperoleh citra LR dengan resolusi 320x180 yang kemudian dilakukan *upscaling* dengan model Real-ESRGAN_x4 untuk memperoleh citra SR dengan resolusi 1280x720. Pada *upscaling* dan *downscaling* dengan faktor 8 citra akan di *downscaling* untuk memperoleh citra LR dengan resolusi 160x90 yang kemudian dilakukan *upscaling* dengan model Real-ESRGAN_x8 untuk memperoleh citra SR dengan resolusi 1280x720.

TABEL 4
RESOLUSI CITRA LR DAN SR

Faktor	Resolusi LR	Resolusi SR
x2	(1280x720)/2 **640x360**	(640x360)x2 **1280x720**
x4	(1280x720)/4 **320x180**	(320x180)x4 **1280x720**
x8	(1280x720)/8 **160x90**	(160x90)x8 **1280x720**

Dari proses tersebut didapatkan citra dengan kualitas yang bervariasi dibuktikan dengan menghitung skor PSNR dan SSIM. Berdasarkan (Tabel 5) citra *ground truth* yang dianggap sebagai citra asli memilki skor PSNR maksimal yaitu 100 db yang menunjukkan bahwa citra di rekonstruksi memilki kesamaan sempurna dan skor maksimal SSIM yaitu 1 yang menunjukkan kesamaan struktural antara citra yang direkonstruksi dan asli, citra dengan faktor *downscaling* dan *upscaling* 2 menghasilkan skor PSNR 27.98 dB dan SSIM 0.88, citra dengan faktor *downscaling* dan *upscaling* 4 menghasilkan skor PSNR 24.16 dB dan SSIM 0.77, citra dengan faktor *downscaling* dan *upscaling* 8 menghasilkan skor PSNR 18.60 dB dan SSIM 0.59.

Berdasarkan hasil tersebut dapat disimpulkan bahwa semakin tinggi faktor pengkalian, PSNR dan SSIM-nya semakin rendah. Ini menunjukkan bahwa semakin besar faktor *downscaling*, semakin buruk kualitas restorasi citra. Dari tabel tersebut skenario faktor x2 menghasilkan gambar dengan kualitas yang lebih baik dibandingkan dengan skenario faktor pengkalian x4 dan x8.

TABEL 5
SKOR KUALITAS CITRA

Faktor	PSNR	SSIM
GT	100 [dB]	1.00 [u]
x2	27.98 [dB]	0.88 [u]
x4	24.16 [dB]	0.77 [u]
x8	18.60 [dB]	0.59[u]

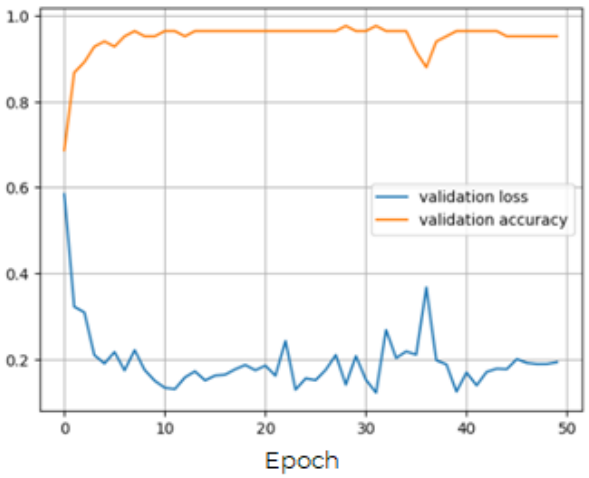
C. Modelling

Setelah dilakukan tahap *preprocessing*, Pada tahap ini dilakukan pemodelan menggunakan teknik *transfer learning* dengan *pre-trained model* MobileNetV2 [10] untuk arsitektur model dapat dilihat pada (Tabel 6). Model yang dilatih memilki total 2.866.978 parameter dengan 608.994 parameter yang dapat dilatih dan 2.257.984 parameter yang tidak dilatih atau bobotnya dibekukan pada tahap *training*.

TABEL 6
ARSITEKTUR MODEL

Layer	Output shape	Parameter
Mobilenetv2 (Functional)	(None, 23, 40, 1280)	2257984
Conv2D (Conv2D)	(None, 21, 38, 32)	368672
Conv2D_1 (Conv2D)	(None, 19, 36, 64)	18496
Conv2D_2 (Conv2D)	(None, 17, 34, 128)	73856
Flatten (Flatten)	(None, 73984)	0
Dense (Output)	(None, 2)	147970
Total params	: 2.866.978	
Trainable params	: 608.994	
Non-trainable params	: 2.257.984	

Model tersebut kemudian dilatih menggunakan data Ground Truth (GT) dengan *batch size* 8, dan *optimizer* Adam dengan *learning rate default* yaitu 0.001 dan jumlah *epoch* 50. Pada tahap *training* model GT didapatkan *validation accuracy* 97.59% pada *epoch* ke-34. Pada (Gambar 4) tertera skor akurasi dan *loss* model GT pada tiap epoch dimana pada data *validation* model menunjukkan kecepatan *convergence* yang cepat yaitu berhasil menghasilkan akurasi diatas 90% hanya dalam 10 *epoch* pertama. Hal ini sangat dipengaruhi oleh penggunaan *pre-trained model* menggunakan MobileNetV2 karena model tersebut telah dilatih sebelumnya pada data dengan jumlah yang relatif besar untuk mengenali pola yang ada pada gambar seperti *edges* dan *texture*. Setelah epoch ke-10 didapatkan fluktuasi akurasi pada model, hal ini merupakan hal yang normal dan dapat disebabkan oleh berbagai faktor mulai dari faktor alami seperti variasi data *training* atau *noise* acak yang terdapat dalam data ataupun faktor pemilihan *hyperparameter* seperti jumlah *epoch*. Pada penelitian yang dilakukan oleh Afaq dan Rao pada tahun 2020 didapatkan bahwa jumlah *epoch* yang terlalu besar dapat menyebabkan model mengalami *overfitting*, selain itu juga didapatkan bahwa pemilihan jumlah *epoch* yang optimal tidak dapat ditentukan secara *universal* karena sangat dipengaruhi faktor seperti karakteristik *dataset* [11]. Permasalahan lain yaitu secara *default* model yang dihasilkan pada tahap *modelling* adalah model yang berdasarkan pada *epoch* terakhir, untuk mendapatkan hasil yang optimal terkadang cara ini kurang efisien karena performa model pada *epoch* terakhir terkadang bukanlah performa terbaik mengingat pada proses *training* kinerja model mengalami fluktuasi. Berdasarkan masalah tersebut ditemukan teknik pemodelan menggunakan *callback* metode ini memungkinkan kita melakukan *early stopping* yaitu pemberhentian proses *training* berdasarkan tingkatan performa yang diinginkan ataupun penyimpanan model terbaik dari keseluruhan proses *training*.



Gambar 4 Akurasi dan *Loss* model GT

D. *Testing*

Pada tahap ini dilakukan pengujian pada model yang telah dihasilkan dengan menghitung jumlah citra yang diprediksi benar dan salah. Data yang digunakan adalah data *testing* yang berjumlah 82 data atau 10% dari total data keseluruhan. Dari (Tabel 7) dapat dilithat hasil dari tahap *testing* pada tiap label L1 dan L2. Ada 4 skenario *testing* yang dilakukan berdasarkan variasi yang didapatkan dari kualitas citra yaitu skenario *ground truth* (GT), citra dengan *downscaling* dan *upscaling factor* 2 (X2), 4 (X4), dan 8 (X8). Pada skenario *ground truth* model berhasil mengklasifikasi 42 data *accident* dan 35 data *non accident* dengan total data yang diklasifikasi dengan benar ada 77. Pada skenario *downscaling* dan *upscaling* dengan *factor* 2 (X2) model berhasil mengklasifikasikan 41 data *accident* dan 27 data *non accident* dengan total data yang diklasifikasi dengan benar ada 68. Pada skenario *downscaling* dan *upscaling* dengan *factor* 4 (X4) model berhasil mengklasifikasikan 31 data *accident* dan 31 data *non accident* dengan total data yang diklasifikasi dengan benar ada 62. Pada skenario *downscaling* dan *upscaling* dengan *factor* 8 (X8) model berhasil mengklasifikasikan 38 data *accident* dan 26 data *non accident* dengan total data yang diklasifikasi dengan benar ada 64. Pada tahap ini ditemukan bahwa kegagalan dalam mengidentifikasi objek disebabkan oleh citra yang menampilkan objek secara berkerumunan.

TABEL 7
PERBANDINGAN HASIL TESTING

Test	Label				Total Benar	Total Salah
	L1		L2			
	Benar	Salah	Benar	Salah		
GT	42	4	35	1	77	5
X2	41	5	27	9	68	14
X4	31	15	31	5	62	20
X8	38	8	26	10	64	18

E. Evaluation

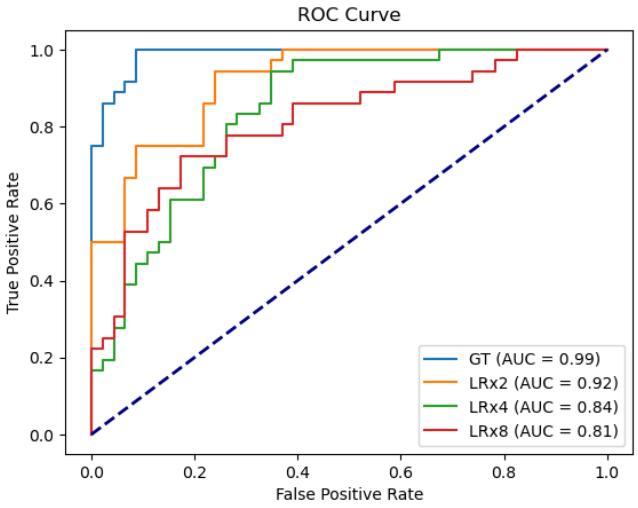
Pada tahap ini dilakukan evaluasi kinerja model secara keseluruhan pada tiap skenario dengan menghitung skor metrik seperti akurasi, *precision*, *recall*, skor AUC dan waktu *inference* yang dibutuhkan model.

TABEL 8
HASIL EVALUASI PENGUJIAN

Skenario	**GT**	**X2**	**X4**	**X8**
Inference Time	3.33s			
PSNR	100 db	27.98 db	24.16 db	18.60 db
SSIM	1.00 u	0.88 u	0.77 u	0.59 u
Akurasi	94%	83%	76%	78%
Precision	94%	83%	77%	78%
Recall	94%	82%	77%	77%
AUC *score*	99%	92%	84%	81%

Berdasarkan (Tabel 8) hasil evaluasi pengujian dibagi berdasarkan skenario uji. *Ground Truth* (GT) adalah referensi atau kebenaran dalam pengujian ini, yang digunakan sebagai titik pembanding. Selanjutnya, terdapat tiga skenario uji, yaitu X2, X4, dan X8, yang mengacu pada faktor *upscaling* dan *downscaling*. Hasil pengujian terdiri dari beberapa metrik. Waktu inferensi menunjukkan berapa lama model membutuhkan waktu untuk melakukan prediksi atau inferensi berdasarkan data yang diberikan. *Peak Signal-to-Noise Ratio* (PSNR) mengukur kualitas rekonstruksi atau restorasi citra, dengan nilai yang lebih tinggi menunjukkan kualitas yang lebih baik. *Structural Similarity Index* (SSIM) adalah metrik yang mengukur sejauh mana citra yang direkonstruksi mirip dengan citra asli, dengan 1 menunjukkan kesamaan sempurna. Akurasi dan presisi, *recall*, dan AUC *score* merupakan metrik yang menggambarkan ketepatan prediksi atau klasifikasi model terhadap data pengujian, dengan penekanan pada aspek yang berbeda. Pada (Tabel 7) dapat diketahui bahwa skenario pengujian GT memiliki performa klasifikasi tertinggi dengan skor akurasi 94%, presisi 94%, *recall* 94% dan skor AUC 94%. Hal ini karena skenario pengujian GT dilakukan pada data yang sebenarnya, dibuktikan dengan skor PSNR 100 dB dan SSIM 1. Namun, ketika melihat hasil model *testing* pada skenario X2, X4, dan X8, terlihat adanya penurunan kinerja dibandingkan dengan skenario GT. Dalam skenario X2 yang memiliki skor kualitas citra dengan PSNR 27.98 dB, SSIM 0.88, terjadi penurunan akurasi menjadi 83%. Demikian juga dengan skenario X4 dan X8, di mana mengalami penurunan kualitas citra lebih lanjut pada PSNR, SSIM, dan akurasi. Selain itu, terlihat juga perbedaan dalam presisi dan *recall* antara skenario X4 dan X8 dengan skenario GT. Presisi dan *recall* pada skenario GT memiliki nilai yang cukup tinggi sebesar 94%, sementara pada skenario X4 dan X8, nilai presisi dan recall hanya berkisar antara 76% hingga 78%. Hal ini menunjukkan bahwa terdapat kehilangan informasi yang signifikan dalam proses pengukuran pada skenario X4 dan X8.

Selain itu, juga dilakukan analisis performa dengan metode ROC Curve (Gambar 5). Berdasarkan hasil uji skenario yang dilakukan, dapat dilihat bahwa pada skenario uji GT (Ground Truth) menghasilkan skor AUC yang tertinggi yaitu sebesar 0.99. AUC (Area Under the Curve) merupakan metrik evaluasi yang digunakan untuk mengukur kinerja model klasifikasi, dimana semakin tinggi nilai AUC, semakin baik model dalam membedakan kelas-kelas yang berbeda. Oleh karena itu, skor AUC sebesar 0.99 menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam melakukan klasifikasi. Selanjutnya, skenario LRX2 memiliki skor AUC sebesar 0.92. Skor ini lebih rendah dibandingkan dengan model GT, namun tetap menunjukkan kinerja yang baik. Skenario LRX2 memiliki kualitas citra dibawah citra GT. Meskipun demikian, hasil yang diperoleh masih memberikan indikasi bahwa model ini mampu membedakan kelas-kelas dengan baik dengan sedikit penurunan performa. Selanjutnya, skenario LRX4 memiliki skor AUC sebesar 0.84. Meskipun skor ini lebih rendah dibandingkan dengan skenario sebelumnya performa model pada skenario LRX4 masih memiliki kinerja yang cukup baik. Terakhir, skenario LRX8 memiliki skor AUC terendah yaitu 0.81. Skor ini menunjukkan bahwa model pada skenario uji LRX8 memiliki kinerja yang paling rendah dibandingkan dengan skenario-skenario sebelumnya. Hal ini karena skenario uji LRX8 menggunakan kualitas citra terendah. Secara keseluruhan, hasil ini menunjukkan bahwa model pada skenario uji GT memiliki kinerja terbaik dalam membedakan kelas-kelas pada dataset yang digunakan, diikuti oleh skenario LRX2, LRX4, dan LRX8.



Gambar 5 ROC Curve

Berdasarkan hasil tersebut dapat diketahui bahwa kualitas citra yang dihitung berdasarkan nilai PSNR dan SSIM pada skenario GT, X2, X4, X8 berkorelasi positif pada performa

klasifikasi seperti Akurasi, *precision, recall,* dan AUC *score*. Semakin tinggi kualitas citra atau skor PSNR dan SSIM-nya, semakin tinggi pula performa klasifikasinya.

IV. KESIMPULAN

Pada penelitian ini telah dilakukan penelitian dimulai dari tahap *data acquisition* diikuti *preprocessing, modeling, testing,* dan tahap *evaluation* dengan menggunakan model MobileNetV2. Penelitian dilakukan pada 989 citra CCTV jalan raya yang kemudian dilakukan proses *downscaling* dan *upscaling* untuk menghasilkan data yang bervariasi yaitu X2, X4, dan X8 untuk pengujian model. Pada penelitian didapatkan model MobileNetV2 berhasil menghasilkan *validation accuracy* 97%, *testing accuracy 94%* dan *average inference time* 3.33 detik pada skenario uji GT. Dari empat skenario pengujian, didapatkan bahwa kualitas citra yang dihitung berdasarkan nilai PSNR dan SSIM sangat berkorelasi positif pada performa klasifikasi seperti Akurasi, *precision, recall,* dan AUC *score*. Semakin tinggi kualitas citra atau skor PSNR dan SSIM-nya, semakin tinggi pula performa klasifikasinya.

DAFTAR PUSTAKA

[1] P. N. Megasari, "Polri Catat 6.707 Kasus Kecelakaan Sepanjang 2022, 452 Orang Tewas," *detikNews*, 2022. https://news.detik.com/berita/d-6325521/polri-catat-6707-kasus-kecelakaan-sepanjang-2022-452-orang-tewas (accessed Apr. 25, 2023).

[2] S. Ghosh, S. J. Sunny, and R. Roney, "Accident Detection Using Convolutional Neural Networks," in *2019 International Conference on Data Science and Communication (IconDSC)*, 2019, pp. 1–6. doi: 10.1109/IconDSC.2019.8816881.

[3] C. B. Ng and W. H. Lo, "Effect of Image Distortion on Facial Age and Gender Classification Performance of Convolutional Neural Networks," in *IOP Conference Series: Materials Science and Engineering*, Institute of Physics Publishing, 2019. doi: 10.1088/1757-899X/495/1/012029.

[4] H. Triana and U. Enri, "Penerapan Deep Learning Pada Kamera Pengawas Jalan Raya Dalam Mendeteksi Kecelakaan," 2022.

[5] J. Pardede and H. Hardiansah, "Deteksi Objek Kereta Api menggunakan Metode Faster R-CNN dengan Arsitektur VGG 16," *MIND Journal*, vol. 7, no. 1, pp. 21–36, Jun. 2022, doi: 10.26760/mindjournal.v7i1.21-36.

[6] C. A. Hartanto and A. Wibowo, "Development of Mobile Skin Cancer Detection using Faster R-CNN and MobileNet v2 Model," in *2020 7th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, 2020, pp. 58–63. doi: 10.1109/ICITACEE50144.2020.9239197.

[7] A. G. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," Apr. 2017, [Online]. Available: http://arxiv.org/abs/1704.04861

[8] R. Indraswari, R. Rokhana, and W. Herulambang, "Melanoma Image Classification Based on MobileNetV2 Network," in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 198–207. doi: 10.1016/j.procs.2021.12.132.

[9] A. Horé and D. Ziou, "Image Quality Metrics: PSNR vs. SSIM," in *2010 20th International Conference on Pattern Recognition*, 2010, pp. 2366–2369. doi: 10.1109/ICPR.2010.579.

[10] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Jan. 2018, [Online]. Available: http://arxiv.org/abs/1801.04381

[11] S. Afaq and S. Rao, "Significance of Epochs on Training A Neural Network," 2020, [Online]. Available: www.ijstr.org