

## PENERAPAN DEEP CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORK UNTUK MENCIPTAKAN DATA SINTESIS PERILAKU PENGEMUDI DI DALAM KENDARAAN

Michael Stephen Lui<sup>1</sup>, Fitra Abdurrachman Bachtiar<sup>\*2</sup>, Novanto Yudistira<sup>3</sup>

<sup>1,2,3</sup>Universitas Brawijaya, Malang

Email: <sup>1</sup>michaellui@student.ub.ac.id, <sup>2</sup>fitra.bachtiar@ub.ac.id, <sup>3</sup>yudistira@ub.ac.id

<sup>\*</sup>Penulis Korespondensi

(Naskah masuk: 9 Januari 2023, diterima untuk diterbitkan: 26 September 2023)

### Abstrak

Kecelakaan kendaraan adalah salah satu penyebab kematian tertinggi di Indonesia. Salah satu solusi untuk mencegah kecelakaan adalah dengan menggunakan sensor eksternal untuk mendeteksi kondisi jalan. Namun, penyebab utama kecelakaan adalah kelalaian pengemudi ketika mengemudi yang tidak dapat terdeteksi oleh sensor eksternal. Sensor visual dapat mendeteksi perilaku pengemudi di dalam kendaraan. Penggunaan sensor visual memiliki performa yang lebih baik ketika menggunakan metode *deep learning*. Salah satu metode untuk meningkatkan performa metode *deep learning* adalah dengan menggunakan data sintesis hasil model generatif sebagai tambahan data. *Deep Convolutional Generative Adversarial Network* (DCGAN) adalah salah satu model generatif yang menggunakan lapisan konvolusi. DCGAN terdiri dari dua *neural network* bernama *generator* dan *discriminator* yang membentuk hubungan *zero-sum game*. *Generator* menerima masukan berupa gambar asli dengan tambahan *noise* sebagai *input* proses latih secara *unsupervised*, menghasilkan gambar sintesis, sedangkan *discriminator* menerima gambar asli dan gambar sintesis sebagai *input* dan menghitung keaslian gambar yang selanjutnya digunakan sebagai nilai *loss* dengan fungsi *loss Binary Cross Entropy*. Arsitektur DCGAN terdiri dari beberapa *transposed convolutional layer* dengan *batch normalization* dan fungsi aktivasi ReLU dan fungsi aktivasi Tanh sebagai *output layer* pada *generator* dan beberapa *convolutional layer* dengan *batch normalization* dan fungsi aktivasi Leaky ReLU dan fungsi aktivasi Sigmoid sebagai *output layer* pada *discriminator*. Dataset yang digunakan pada penelitian ini adalah dataset ISDDS perilaku umum pengemudi yang dikumpulkan pada skenario simulasi dengan jumlah dua ribu gambar. Hasil pengujian menemukan bahwa nilai *hyperparameter* dapat menghasilkan gambar sintesis perilaku pengemudi di dalam kendaraan yang baik dengan nilai FID sebesar 274,16 pada *learning rate discriminator* pada 0,0001,  $\beta_1$  *discriminator* pada 0,8005, *learning rate generator* pada 0,0017,  $\beta_1$  *generator* pada 0,1138 selama 43 *epoch* dengan menggunakan *optimizer Adam* pada *generator* dan *discriminator*.

**Kata kunci:** teknologi pencegah kecelakaan, *deep learning*, gambar sintesis, model generatif, DCGAN

## IMPLEMENTATION OF DEEP CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORK TO GENERATE SYNTHETIC DATA OF DRIVER ACTIVITY INSIDE VEHICLE

### Abstract

Vehicle crash is one of the leading causes of death in Indonesia. One of the solutions to prevent vehicle crash is by using external sensor to detect road condition. Yet, most crash happened because of driver distraction, which is hard to detect using external sensor. Visual sensor can be used to detect driver activity inside vehicle. Visual sensor that uses deep learning method performs well. One way to increase deep learning method performance is by using additional synthesis data made by generative model. Deep Convolutional Generative Adversarial Network (DCGAN) is a generative model that uses convolution layer. DCGAN consists of two neural networks titled generator and discriminator which create zero-sum game relationship. Generator will receive real image with added noise as input of unsupervised training process, creating synthetic image, while discriminator will receive real image and synthetic image as input and calculate the realness of those image which will be used as loss value with Binary Cross Entropy loss function. The Architecture of DCGAN is composed of multiple transposed convolutional layers with batch normalization and activation function ReLU and activation function Tanh as output layer in generator and multiple convolutional layers with batch normalization and activation function Leaky ReLU and activation function Sigmoid as output layer in discriminator. Dataset used in this

research is primary dataset of common driver activity collected in simulation scenario with the size of two thousand images. Experiment result shows that DCGAN is able to create good image synthesis of driver activity inside vehicle with FID of 274,16 using hyperparameter consisting of learning rate discriminator at 0,0001,  $\beta_1$  discriminator at 0,8005, learning rate generator at 0,0017,  $\beta_1$  generator at 0,1138 for 43 epochs by using Adam optimizer on generator dan discriminator.

**Keywords:** crash prevention technology, deep learning, image synthesis, generative model, DCGAN

## 1. PENDAHULUAN

Kecelakaan kendaraan adalah salah satu penyebab kematian tertinggi di Indonesia. Kementerian Perhubungan Republik Indonesia melaporkan bahwa terdapat 103.645 kasus kecelakaan pada tahun 2021, dengan korban jiwa mencapai 25.266, dan kerugian materi sebesar Rp 246 miliar (Rahayu, 2022). Rata-rata kematian akibat kecelakaan di dunia adalah 27,5 per 100.000 orang (World Health Organization, 2018). Penelitian oleh Née dkk. (2019) menemukan bahwa 57,6% kecelakaan di Australia pada tahun 2000 hingga 2011 disebabkan oleh kelalaian pengemudi dan 25% dari kelalaian pengemudi tersebut akibat gangguan pada pengemudi.

Menurut penelitian oleh Mcdonald dkk. (2018), *Advanced Driver Assistance System* (ADAS) berkontribusi dalam penurunan jumlah kecelekaan. ADAS adalah teknologi pencegah kecelakaan yang menggunakan sensor eksternal untuk mendeteksi kondisi jalan. Namun, penelitian oleh Beanland dkk. (2013) menemukan bahwa penyebab utama kecelakaan adalah kelalaian pengemudi ketika mengemudi. Penelitian pedeteksi perilaku pengemudi adalah penelitian menggunakan sensor visual maupun psikologikal untuk mendeteksi perilaku pengemudi sehingga mampu memberikan bantuan yang lebih efektif dibandingkan dengan teknologi pencegah kecelakaan yang menggunakan sensor eksternal (Koay et al., 2022). Tujuan penelitian pendeteksi perilaku pengemudi adalah untuk mengurangi jumlah kecelakaan yang sangat merugikan banyak pihak.

Penelitian oleh Omerustaoglu dkk. (2020) menggunakan data primer berupa 137.093 gambar perilaku pengemudi natural dan menggunakan Convolutional Neural Network (CNN) dan Long Short-Term Memory (LSTM) untuk mengklasifikasikan kondisi pengemudi. Hasil penelitian mendapatkan akurasi sebesar 0,62 pada CNN dan 0,76 pada LSTM. Salah satu masalah yang dialami pada penelitian tersebut adalah kurangnya variasi data karena yang digunakan hanya memiliki satu objek pengemudi. Penelitian oleh Zhang dkk. (2020) menggunakan data primer berupa 60 jam video perilaku pengemudi yang diambil pada skenario simulasi. Penelitian tersebut berusaha untuk mengklasifikasikan perilaku pengemudi menggunakan beberapa model CNN yaitu MobileNet, MobileNetV2, ShuffleNet, ShuffleNetV2, dan I3D. Hasil penelitian

mendapatkan akurasi yang rendah berkisar antara 0,62 hingga 0,69. Penelitian tersebut memiliki akurasi yang buruk akibat data yang kabur dikarenakan gerakan pengemudi sehingga mengurangi keefektifan model dalam melakukan klasifikasi. Penelitian oleh Hu dkk. (2020) menggunakan data primer berupa 42.816 gambar perilaku pengemudi yang diambil pada kendaraan yang diparkir untuk mengklasifikasikan perilaku pengemudi. Penelitian tersebut menggunakan beberapa metode *machine learning* yaitu Multilayer Perceptron (MLP) dan Support Vector Machine (SVM) yang menghasilkan akurasi 0,71 pada MLP serta 0,75 pada SVM, serta beberapa model CNN yaitu AlexNet dan VGG19 yang menghasilkan akurasi 0,8 pada AlexNet dan 0,84 pada VGG19. Untuk mengurangi kemungkinan *overfitting* akibat kurangnya data, penelitian tersebut menggunakan metode augmentasi yaitu *random crop*, *random noise*, dan *flipping* untuk meningkatkan jumlah data. Penelitian oleh Wagner dkk. (2022) menggunakan data primer yang dikumpulkan menggunakan dua kamera yang menghasilkan gambar dengan total 19.427 gambar hitam putih yang menampilkan perilaku pengemudi secara natural untuk proses pelatihan beberapa model CNN yaitu VGG16 dan VGG19. Hasil penelitian klasifikasi menggunakan kedua model tersebut mendapatkan akurasi yang sangat rendah pada 0,38 dan 0,54. Penelitian tersebut lalu menggunakan model ResNeXt34 dan ResNeXt50 yang lebih kompleks dan menggunakan konsep kardinalitas sehingga mampu menghasilkan akurasi yang tinggi pada 0,81 dan 0,9.

Penelitian-penelitian tersebut memiliki masalah pada proses pelatihan model *deep learning* akibat kurangnya data yang tersedia. Model *deep learning* membutuhkan data dengan jumlah yang besar untuk menghasilkan performa yang maksimal (Goodfellow dkk. 2018, p. 8). Proses augmentasi dapat digunakan untuk meningkatkan jumlah data sehingga mengurangi kemungkinan terjadinya *overfitting*. Salah satu jenis proses augmentasi terbaru adalah dengan cara menciptakan data sintesis dengan menggunakan model generatif (Wu et al., 2020). Berdasarkan penelitian-penelitian sebelumnya, penelitian ini bertujuan untuk menggunakan model generatif untuk menciptakan data sintesis perilaku pengemudi yang merupakan penelitian yang belum ditelusuri sebelumnya. DCGAN adalah salah satu jenis GAN yang menggunakan lapisan konvolusi. DCGAN dipilih

sebagai metode penelitian ini karena model tersebut memiliki hasil yang baik ketika digunakan pada berbagai jenis data seperti yang dapat dilihat pada penelitian oleh Lu dkk. (2022).

Penelitian ini akan melakukan pencarian *hyperparameter* DCGAN yang memiliki nilai FID terbaik sehingga dapat menciptakan gambar sintesis perilaku pengemudi yang berkualitas. Pengaturan *hyperparameter* yang akan dilakukan adalah *epoch*, *optimizer*, *learning rate*, *momentum*, dan  $\beta_1$  pada *generator* dan *discriminator*. Evaluasi yang digunakan adalah Frechet Inception Distance (FID) dan evaluasi manual. FID digunakan karena kekonsistensian dengan inspeksi manusia serta sensitif terhadap perubahan pada distribusi nilai gambar sehingga mampu mendeteksi kekaburan serta artefak yang sering dialami pada gambar sintesis yang buruk (Borji, 2022). Evaluasi manual digunakan karena evaluasi manual adalah salah satu metode evaluasi yang paling intuitif untuk mengukur kualitas gambar sintesis (Borji, 2022).

Berdasarkan masalah-masalah yang telah diuraikan sebelumnya yaitu tidak adanya penelitian sebelumnya yang menggunakan model generatif untuk menciptakan data sintesis perilaku pengemudi yang dapat membantu meningkatkan performa model klasifikasi perilaku pengemudi, maka dilaksanakanlah penelitian ini dengan judul “Penerapan Deep Convolutional Generative Adversarial Network untuk Menciptakan Data Sintesis Perilaku Pengemudi di Dalam Kendaraan.” Hasil penelitian ini diharapkan dapat menjadi solusi kurangnya data perilaku pengemudi dengan menghasilkan data sintesis yang berkualitas.

## 2. PENELITIAN SEBELUMNYA

Penelitian yang dilakukan oleh Douarre dkk. (2019) bertujuan untuk meningkatkan performa metode segmentasi infeksi apel dengan menggunakan gambar sintesis. Data yang digunakan penelitian ini adalah 11 gambar IR dengan dimensi  $2592 \times 1944$ , dengan pada setiap gambar terdapat beberapa objek berupa daun yang telah dianotasi sebagai terinfeksi atau tidak terinfeksi. Penelitian ini menggunakan SegNet untuk melakukan segmentasi, DCGAN untuk menciptakan gambar sintesis infeksi apel, dan metrik F1 *score* untuk mengukur performa metode segmentasi. SegNet menggunakan *hyperparameter* awal *learning rate* sebesar 0,0001, *weight decay* sebesar 0,0005. *Hyperparameter* tersebut lalu disesuaikan menggunakan *grid search* untuk mencegah *overfitting* karena ketidakseimbangan kelas pada data. DCGAN menggunakan *hyperparameter default* yaitu *learning rate* 0,0002, *optimizer* Adam, dan  $\beta_1$  sebesar 0,5. Pelatihan dilakukan hingga nilai *loss* dan gambar sintesis stabil selama sekitar seratus ribu iterasi. Hasil penelitian menemukan bahwa gambar sintesis dapat meningkatkan performa metode segmentasi dengan F1 *score* awal 0,5 menjadi 0,6.

Penelitian yang dilakukan oleh Bautista dkk. (2022) bertujuan untuk mengukur performa DCGAN pada gambar *ultrasound*. Data yang digunakan adalah gambar *ultrasound* yang telah melalui proses *resize* menjadi  $64 \times 64$ . DCGAN menggunakan *hyperparameter default* yaitu *learning rate* 0,0002, *optimizer* Adam, dan  $\beta_1$  sebesar 0,5. DCGAN digunakan untuk membuat delapan ratus gambar sintesis dengan proses pelatihan berjalan selama 1.000 *epoch*. Penelitian tersebut menggunakan evaluasi manual dengan hasil sebanyak 56,14% gambar sintesis tidak dapat dibedakan dari gambar asli. Hasil penelitian tersebut menunjukkan bahwa gambar sintesis hasil DCGAN mampu menggantikan gambar asli *ultrasound*.

Penelitian yang dilakukan oleh Adedeji dkk. (2022) bertujuan untuk membandingkan akurasi model klasifikasi VGG16 dan Resnet50 pada gambar satelit dengan proses augmentasi geometrik, DCGAN, dan WGAN-GP. Data yang digunakan adalah 27.000 gambar satelit dengan dimensi  $64 \times 64$  yang terbagi menjadi 10 kelas. Proses augmentasi geometrik yang digunakan adalah *random horizontal flip*, *random vertical flip*, dan *random rotation*. DCGAN menggunakan *hyperparameter default* yaitu *learning rate* 0,0002, *optimizer* Adam, dan  $\beta_1$  sebesar 0,5. WGAN-GP menggunakan *hyperparameter* yang sama dengan DCGAN. VGG16 dan Resnet50 yang digunakan adalah model *pretrained*. Gambar sintesis hasil augmentasi digunakan untuk menggantikan 10% dari gambar asli. Penelitian tersebut menemukan bahwa penggunaan DCGAN untuk menggantikan gambar asli tanpa mengalami penurunan akurasi yang signifikan, yaitu VGG16 sebesar 0,03 dan Resnet50 sebesar 0,23.

Penelitian yang dilakukan oleh Curtó dkk. (2017) bertujuan untuk mengembangkan arsitektur baru yang mampu mengolah masukan resolusi tinggi secara stabil untuk meningkatkan kualitas gambar sintesis yang dihasilkan oleh DCGAN. Arsitektur HDCGAN menggunakan fungsi aktivasi *Scaled Exponential Linear Units* (SELU) untuk menggantikan fungsi aktivasi ReLU dan Leaky ReLU yang berada pada arsitektur DCGAN untuk meningkatkan stabilitas model ketika menggunakan masukan dengan resolusi yang tinggi. Penelitian menemukan bahwa pada data CelebA, DCGAN menghasilkan nilai FID sebesar 16,3 dan MS-SSIM sebesar 0,28, sedangkan HDCGAN menghasilkan nilai FID sebesar 8,44 dan MS-SSIM sebesar 0,19.

Penelitian yang dilakukan oleh Kim (2020) bertujuan untuk mengembangkan arsitektur baru bernama BoolGAN yang merupakan pengembangan dari DCGAN. BoolGAN menghasilkan gambar sintesis yang lebih baik dengan menambahkan *convolutional layer* di akhir *generator* untuk mengurangi *noise*, *dropout layer* di akhir *discriminator* untuk memberikan efek *stochastic*, dan Wasserstein *loss* untuk mengurangi

kemungkinan terjadinya *mode collapse*. BoolGAN menghasilkan proses pelatihan yang lebih stabil dan mampu mengurangi noise pada gambar sintesis. Pelatihan dilaksanakan selama 50 *epoch* dengan *learning rate* 0,0002, *optimizer* Adam,  $\beta_1$  sebesar 0,5, dan *dropout* 0,2. Penelitian menemukan bahwa BoolGAN memiliki nilai FID yang lebih baik pada 165,96 dibandingkan dengan nilai FID hasil DCGAN yaitu 195,92.

Penelitian-penelitian sebelumnya menunjukkan bahwa DCGAN dapat digunakan untuk menghasilkan data sintesis yang berkualitas. Metode evaluasi DCGAN dapat berupa evaluasi manual seperti yang dapat dilihat pada penelitian oleh Bautista dkk. (2022), menjadi metode augmentasi data seperti yang dapat dilihat pada penelitian oleh Douarre dkk. (2019) dan Adedeji dkk. (2022), dan FID seperti yang dapat dilihat pada penelitian oleh Kim (2020). Pada penelitian ini DCGAN diimplementasikan untuk menciptakan data sintesis perilaku pengemudi yang belum pernah dilakukan. Proses pelatihan DCGAN pada penelitian ini menggunakan *dataset* primer ISDDS dengan jumlah data sebanyak dua ribu. Penelitian ini juga mencari *hyperparameter* performa terbaik untuk menciptakan data perilaku pengemudi yang berkualitas. Evaluasi yang digunakan pada penelitian ini adalah FID dan evaluasi manual.

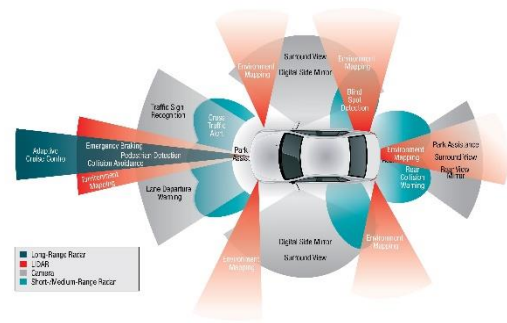
### 3. LANDASAN TEORI

#### 3.1. Advanced Driving Assistance System (ADAS)

*Advanced Driving Assistance System* (ADAS) adalah sistem yang membantu pengemudi untuk menghindari kecelakaan kendaraan. Kendaraan yang telah menggunakan sistem ini memiliki fitur seperti *Forward Collision-Avoidance Assist* (FCA), *Lane Keeping Assist* (LKA), dan *Intelligent Speed Limit Assist* (ISLA) yang menggunakan sensor eksternal untuk mendeteksi kondisi jalan. Namun, sistem tersebut kurang efektif karena penyebab utama kecelakaan adalah pengemudi yang tidak fokus (Hari & Sankaran, 2021). Penelitian pendeteksi perilaku pengemudi bertujuan untuk mengurangi jumlah kecelakaan dengan menggunakan sensor visual. Sensor visual dapat mendeteksi gangguan yang dialami pengemudi sehingga lebih efektif dalam pencegahan kecelakaan. Ilustrasi fitur pada kendaraan yang menggunakan ADAS dapat dilihat pada Gambar 1.

#### 3.2. Safety Driving

*Safety driving* adalah suatu kondisi dimana pengemudi tidak mengalami gangguan sehingga dapat menyetir secara aman. Gangguan pada pengemudi dapat berupa gangguan manual, gangguan visual, dan gangguan kognitif (Dingus et al., 2016).



Gambar 1. Ilustrasi fitur pada kendaraan yang menggunakan ADAS  
(Sumber: Wendt & Cook, 2018)

Gangguan manual adalah gangguan secara fisik seperti makan dan minum. Gangguan visual adalah gangguan yang berdampak pada mata pengemudi seperti mengirim pesan pada handphone dan mengoperasikan radio. Gangguan kognitif adalah gangguan yang membuat pengemudi tidak fokus seperti berbicara dengan penumpang atau menelepon. Sistem pendeteksi perilaku pengemudi adalah sistem yang menggunakan sensor visual maupun psikologikal (Koay et al., 2022). Sistem tersebut mendeteksi dan memperingati apabila pengemudi sedang mengalami gangguan dengan tujuan untuk mengurangi jumlah kecelakaan akibat gangguan-gangguan tersebut.

#### 3.3. Deep Convolutional Generative Adversarial Network (DCGAN)

*Deep Convolutional Generative Adversarial Network* (DCGAN) adalah perkembangan dari GAN dimana DCGAN menggunakan lapisan konvolusi pada arsitekturnya. DCGAN terdiri dari dua *neural network* bernama *generator* dan *discriminator* yang membentuk hubungan *zero-sum game*. *Discriminator* berlatih untuk memaksimalkan performa klasifikasi gambar asli, sedangkan *generator* berlatih untuk menciptakan gambar sintesis yang tidak dapat dibedakan dari gambar asli. *Generator* akan menerima gambar asli dan *noise* sebagai *input* proses latih secara *unsupervised*, menghasilkan gambar sintesis, sedangkan hasil *discriminator* akan digunakan sebagai nilai *loss* dengan fungsi *loss Binary Cross Entropy* (1).

$$L(x) = -\frac{1}{n} \sum_{i=1}^n x \log \hat{x} + (1-x) \log(1-\hat{x}) \quad (1)$$

dengan  $n$  adalah jumlah kelas,  $x$  adalah *input*. Hasil akhir *discriminator* akan digunakan untuk menghitung nilai *loss* (2)

$$\min_G \max_D L_D(D, G) = \log D(x) + \log(1 -$$

$$D(G(z)) \quad (2)$$

dengan  $G$  adalah *generator*,  $D$  adalah *discriminator*,  $x$  adalah gambar asli, dan  $z$  adalah *noise*. Nilai *loss*

akan digunakan untuk proses pelatihan *generator* dan *discriminator*.

Pada DCGAN, *generator* menggunakan fungsi aktivasi ReLU (3) untuk seluruh lapisan kecuali pada *output layer* yang menggunakan fungsi aktivasi Tanh (4). Alasan utama penggunaan kedua fungsi aktivasi tersebut adalah untuk mempercepat proses pelatihan model (Radford et al., 2015). *Discriminator* DCGAN menggunakan fungsi aktivasi LeakyReLU dengan *slope* 0,2 (5) pada seluruh lapisan kecuali pada *output layer* yang menggunakan fungsi aktivasi sigmoid (6).

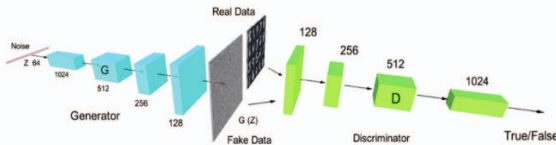
$$f(x) = \max(0, x) \quad (3)$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (4)$$

$$f(x) = \max(sx, x) \quad (5)$$

$$f(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

dengan  $x$  adalah masukan,  $s$  adalah *slope*, dan  $e$  adalah bilangan euler.



Gambar 2. Struktur DCGAN

## 4. METODOLOGI PENELITIAN

### 4.1. Pengumpulan Data

Data yang digunakan pada penelitian ini merupakan *dataset* primer ISDDS yang terdiri dari dua ribu gambar yang diambil dari 50 subjek yang menampilkan perilaku-perilaku umum pengemudi. *Dataset* ini dikumpulkan di laboratorium Teknologi Pembelajaran Fakultas Ilmu Komputer Universitas Brawijaya sejak 9 Agustus 2022 hingga 7 Oktober 2022. *Dataset* ini dibagi menjadi 10 kelas yaitu: *safety driving*, *texting-right*, *texting-left*, *talking on the phone-right*, *talking on the phone-left*, *operating the radio*, *drinking*, *reaching behind*, *hair and makeup*, dan *talking to passenger*. Gambar pada *dataset* ini diambil pada kondisi simulasi dengan menggunakan kamera.

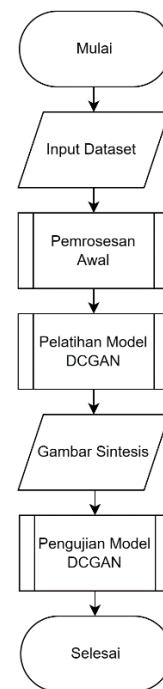
Pemrosesan awal yang dilakukan adalah normalisasi, *crop*, dan *resize*. Gambar akan melalui tahap *crop* terlebih dahulu untuk menghilangkan bezel hitam pada bagian kiri dan kanan gambar. *Tree-structured Parzen Estimator* (TPE) adalah sebuah metode *sampling* yang digunakan untuk memilih nilai *hyperparameter* yang akan diuji. TPE melakukan pemodelan  $p(x|y)$  dengan menggunakan parameter pada *Deep Belief Network* (DBN) dan

mengganti distribusi parameter tersebut dengan densitas non-parameter (Bergstra et al., 2011).

Hasil *crop* adalah gambar dengan dimensi  $930 \times 720$ . Tahap selanjutnya adalah *resize* untuk menghasilkan gambar dengan dimensi  $64 \times 64$ . Tujuan dilakukannya *resize* adalah untuk mempermudah komputasi sehingga mempercepat proses pelatihan. Gambar lalu akan melalui tahap normalisasi dengan nilai *mean* dan standar deviasi 0,5. Normalisasi dilakukan untuk mempermudah komputasi dengan mengubah persebaran data menjadi bernilai antara 0 dengan 1.

### 4.2. Perancangan Sistem

Implementasi metode DCGAN menggunakan bahasa pemrograman Python dengan *library* PyTorch. Sistem akan diimplementasikan sesuai dengan diagram pada Gambar 3.



Gambar 3. Diagram alir sistem.

### 4.3. Perancangan Pengujian

Pada penelitian ini akan dilakukan dua jenis pengujian yaitu pengujian *hyperparameter* dan pengujian dengan evaluasi manual. Pengujian *hyperparameter* pada penelitian ini bertujuan untuk menemukan pengaturan *hyperparameter* yang menghasilkan gambar sintesis yang berkualitas. Pengujian dengan evaluasi manual bertujuan untuk mengukur kualitas gambar sintesis secara subjektif.

Pengujian *hyperparameter* dilaksanakan dengan menggunakan bantuan *library* Optuna. Perancangan pengujian pengaturan *hyperparameter* DCGAN pada penelitian ini terdiri dari: *optimizer*, *learning rate*, dan *momentum* atau  $\beta_1$  untuk *generator* maupun *discriminator*. Optuna

menggunakan metode *sampling* TPE untuk melakukan pemilihan *hyperparameter* yang digunakan pada proses pelatihan. TPE melakukan pemodelan dengan menggunakan parameter pada DBN dan mengganti distribusi parameter tersebut dengan densitas non-parameter. Percobaan *hyperparameter* dilaksanakan sebanyak 100 kali.

Pengujian evaluasi manual dilaksanakan untuk mengukur kualitas gambar sintesis secara subjektif. Pengumpulan pendapat dilakukan dengan menggunakan survei. Survei terdiri dari 10 pertanyaan yang meminta responden untuk menilai kualitas gambar sintesis dibandingkan dengan gambar asli. Pada setiap pertanyaan, responden menilai kualitas gambar sintesis dengan beberapa pilihan jawaban yaitu: baik, netral, dan buruk

#### 4.5. Metode Evaluasi

Metode evaluasi yang digunakan pada penelitian ini adalah *Frechet Inception Distance* (FID). FID adalah metode evaluasi yang paling sering digunakan untuk mengukur kualitas gambar hasil sintesis model generatif (Bautista et al., 2022). FID merupakan perkembangan dari *Inception Score* (IS) yang hanya mengevaluasi distribusi gambar sintesis. FID membandingkan distribusi gambar asli yang digunakan untuk proses pelatihan *generator* dan gambar sintesis (10). FID menggunakan bobot yang didapatkan dari Inception V3 yang dilatih menggunakan ImageNet.

$$FID = |\mu_1 - \mu_2| + Tr(\sigma_1 + \sigma_2 - \sqrt{\sigma_1 \sigma_2}) \quad (10)$$

dengan  $\mu$  adalah rata-rata dan  $\sigma$  adalah variansi.

## 5. HASIL DAN PEMBAHASAN

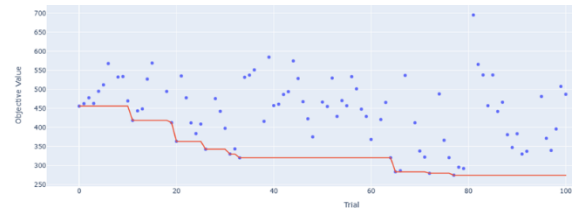
Cuplikan hasil pengujian dengan bantuan *library* Optuna sebanyak 100 kali dapat dilihat pada Tabel 1.

Tabel 1. Cuplikan hasil pengujian *hyperparameter*.

trial	FID	Dlr	Dm	Glr	Gm	Epoch
1	455,68	0,003	0,1	0,008	0,4	19
...	...	...	...	...	...	...
69	274,16	0,0001	0,8	0,002	0,1	43
...	...	...	...	...	...	...
100	545,99	0,0002	0,6	0,002	0,2	38

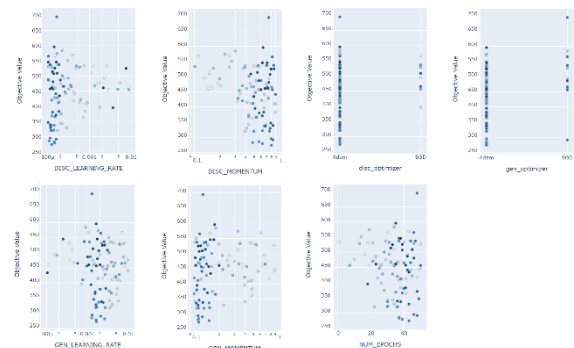
dengan Dlr adalah *discriminator learning rate*, Dm adalah *discriminator momentum*, Glr adalah *generator learning rate*, dan Gm adalah *generator momentum*.

Gambar 4 menunjukkan riwayat FID dan FID terbaik pada pengujian. Nilai FID terlihat tidak memiliki pola.



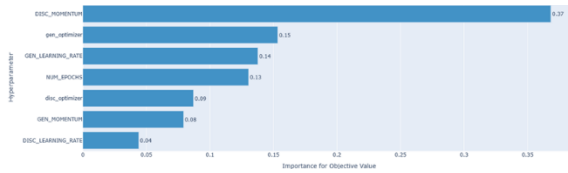
Gambar 4. Riwayat hasil pengujian.

Gambar 5 menunjukkan persebaran nilai *hyperparameter* yang digunakan pada saat pengujian. Warna titik yang lebih gelap mengindikasikan nomor percobaan yang lebih besar. Pengujian lebih sering menggunakan *optimizer* Adam pada *generator* dan *discriminator* dibandingkan dengan *optimizer* SGD. Pada *hyperparameter learning rate discriminator*, terlihat bahwa nilai yang digunakan pada percobaan terbaru berada pada sekitar 0,0001. *Hyperparameter* lain seperti  $\beta_1$  *discriminator* memiliki nilai yang lebih tersebar diantara 0,4 hingga 0,9. *Hyperparameter learning rate generator* memiliki persebaran yang lebih luas dibandingkan dengan *learning rate discriminator* dengan persebaran berada pada jangka antara 0,001 hingga 0,005. Sebaliknya, persebaran nilai  $\beta_1$  *generator* lebih rapat dibandingkan dengan nilai  $\beta_1$  *discriminator* dengan persebaran berpusat pada nilai 0,12. *Epoch* yang digunakan tersebar pada jangka 30 hingga 50 *epoch*.



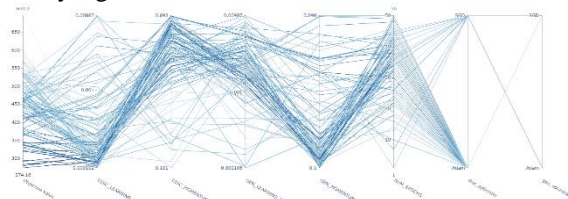
Gambar 5. Persebaran nilai *hyperparameter* terhadap nilai FID.

Gambar 6 menunjukkan ukuran kepentingan *hyperparameter* terhadap nilai FID yang diukur menggunakan metode *Functional Analysis of Variance* (fANOVA). Berdasarkan gambar tersebut, dapat dilihat bahwa  $\beta_1$  *discriminator* merupakan *hyperparameter* terpenting untuk menghasilkan nilai FID yang baik dengan pengaruh yang tinggi dibandingkan dengan *hyperparameter* lain. *Hyperparameter* lain yaitu *learning rate generator*, *optimizer* yang digunakan pada *generator*, jumlah *epoch* yang digunakan, *optimizer* yang digunakan pada *discriminator*,  $\beta_1$  *generator*, dan *learning rate discriminator* memiliki pengaruh yang kecil terhadap nilai FID.



Gambar 6. Persebaran nilai hyperparameter terhadap nilai FID.

Gambar 7 menunjukkan hubungan antar *hyperparameter* terhadap nilai FID. Warna garis yang lebih gelap mengindikasikan nomor percobaan yang lebih besar. Pada hubungan antara  $\beta_1$  *discriminator* dengan *learning rate discriminator* dapat dilihat bahwa semakin tinggi  $\beta_1$  *discriminator* maka dibutuhkan *learning rate discriminator* yang rendah untuk mendapatkan nilai FID yang baik. Sebaliknya,  $\beta_1$  *generator* membutuhkan nilai yang rendah didampingi dengan *learning rate generator* yang tinggi dibutuhkan untuk mendapatkan nilai FID yang baik.



Gambar 7. Hubungan antar hyperparameter terhadap nilai FID.

Hasil DCGAN dengan *hyperparameter* yang menghasilkan nilai FID terbaik yaitu pada percobaan ke-69 yaitu dengan nilai FID sebesar 274,16 dengan menggunakan *optimizer* Adam pada *generator* dan *discriminator*, *hyperparameter learning rate discriminator* pada 0,0001,  $\beta_1$  *discriminator* pada 0,8005, *learning rate generator* pada 0,0017,  $\beta_1$  *generator* pada 0,1138 selama 43 *epoch*.



Gambar 8. Perkembangan hasil gambar sintesis DCGAN.

Gambar 8 menunjukkan perkembangan hasil gambar sintesis DCGAN setiap 5 *epoch*. Pada *epoch* pertama gambar sintesis memiliki warna keunguan, dengan gambar yang tidak terlalu jelas. Setelah itu pada *epoch* ke-5 terlihat sosok manusia namun masih terlihat buram. Gambar sintesis selanjutnya terlihat lebih jelas namun masih belum menghasilkan gambar yang jernih. Meskipun pada *epoch* terakhir yaitu *epoch* ke-43 gambar sintesis masih belum terlihat jernih.

Tabel 2 berisi cuplikan hasil evaluasi DCGAN yang terdiri dari nilai *loss generator*, nilai *loss discriminator*, dan FID pada setiap. Pada Tabel 2 dapat dilihat bahwa *epoch* dengan nilai FID terbaik berada pada *epoch* ke-17 dengan nilai FID 190,53.

Namun, dapat dilihat bahwa hasil gambar sintesis pada Gambar 8 *epoch* ke-15 dan ke-20 tidak memiliki perbedaan yang signifikan dari hasil gambar sintesis pada *epoch* ke-43.

Tabel 2. Cuplikan hasil evaluasi

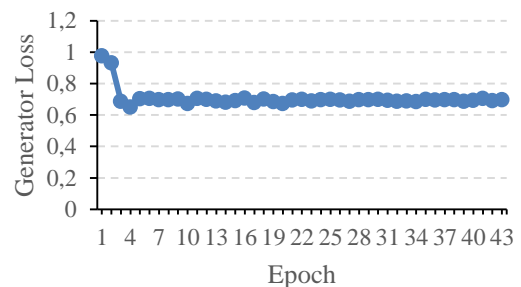
epoch	FID	D loss	G loss
1	209,05	0,48	0,98
...	...	...	...
17	190,53	0,69	0,68
...	...	...	...
43	274,16	0,69	0,69

Tabel 3 berisi hasil evaluasi manual. Pada Tabel 3 dapat dilihat bahwa gambar nomor 10 memiliki penilaian baik sebanyak yaitu 4, dengan gambar nomor 5, 8, dan 9 memiliki penilaian baik sejumlah 3. Hal ini memiliki kemiripan dengan nilai FID pada Tabel 3 dimana nilai FID terbaik berada pada *epoch* ke-17 yang dekat dengan *epoch* ke-20 dimana gambar nomor 5 dihasilkan. Total penilaian baik adalah 20 dari 50, penilaian buruk adalah 18 dari 50, dan penilaian netral adalah 12 dari 20.

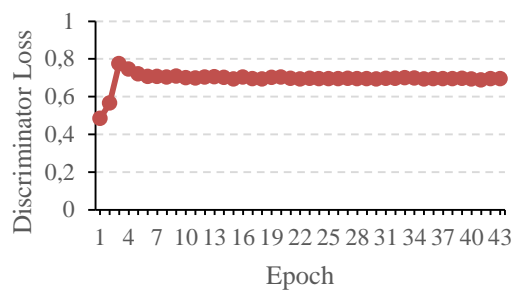
Tabel 3. Cuplikan hasil evaluasi

Gambar	Penilaian		
	Baik	Buruk	Netral
1	0	5	0
2	0	4	1
3	2	3	0
4	2	1	2
5	3	0	2
6	2	2	1
7	1	2	2
8	3	1	1
9	3	0	2
10	4	0	1

Hasil analisis evaluasi manual memiliki kemiripan dengan nilai FID, dengan gambar sintesis pada *epoch* 20 memiliki banyak jawaban baik pada evaluasi manual, dan nilai FID yang paling rendah pada proses pelatihan. Selain itu, evaluasi manual memberikan banyak jawaban baik juga pada kedua gambar sintesis *epoch* terakhir yaitu ke-43. Dapat dilihat pada Lampiran C bahwa gambar sintesis pada kedua *epoch* tersebut memiliki kesamaan yaitu gambar yang relatif jelas dan terlihat memiliki variasi warna baju.

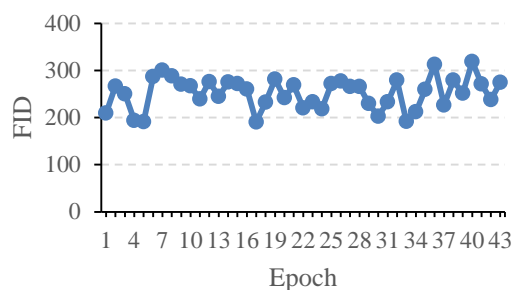


Gambar 9. Nilai loss generator terhadap epoch



Gambar 10. Nilai loss discriminator terhadap epoch

Pada Gambar 9 dan Gambar 10 dapat dilihat bahwa nilai *loss generator* dan *discriminator* mencapai *equilibrium* pada *epoch* ke-5. Gambar 14 menunjukkan bahwa nilai FID memiliki tren untuk naik seiring dengan besarnya *epoch*.



Gambar 11. Nilai FID terhadap epoch

Pada Gambar 11 dapat dilihat bahwa nilai FID memiliki tren untuk naik seiring dengan besarnya *epoch*. Selain itu, perbedaan nilai FID per *epoch* terlihat tajam, menunjukkan bahwa nilai FID tidak stabil. Nilai FID pada *epoch* terakhir juga bukan nilai terbaik, dengan nilai FID terbaik terlihat berada pada *epoch* ke-17.

## 6. KESIMPULAN

Berdasarkan hasil pengujian *hyperparameter*, dapat disimpulkan bahwa nilai *hyperparameter* dapat menghasilkan gambar sintesis perilaku pengemudi didalam kendaraan yang baik pada *learning rate discriminator* pada 0,0001,  $\beta_1$  *discriminator* pada 0,8, *learning rate generator* pada 0,002,  $\beta_1$  *generator* pada 0,1 selama 20 *epoch* dengan menggunakan *optimizer Adam* pada *generator* dan *discriminator*. *Hyperparameter* pada *discriminator* memiliki pengaruh yang lebih tinggi dibandingkan dengan *hyperparameter* pada *generator*, serta momentum/ $\beta_1$  memiliki hubungan terbalik dengan *learning rate*. Berdasarkan hasil pengujian *hyperparameter*, hasil pelatihan DCGAN, dan evaluasi hasil gambar sintesis DCGAN, dapat disimpulkan bahwa nilai FID masih belum stabil dengan jumlah data yang tersedia. Berdasarkan hasil evaluasi manual, dapat disimpulkan bahwa hasil gambar sintesis DCGAN pada *epoch* ke-20, 40, dan 43 memiliki kualitas yang relatif baik. Berdasarkan

hasil metode klasifikasi, dapat disimpulkan bahwa hasil gambar sintesis DCGAN dapat digunakan sebagai data sintesis.

Beberapa saran yang dapat diberikan untuk pengembangan pada penelitian selanjutnya adalah dengan cara meningkatkan jumlah data dan melakukan pencarian *hyperparameter* dengan rentang yang lebih luas dan percobaan yang lebih banyak untuk menemukan *hyperparameter* dengan nilai FID yang stabil sehingga dapat menghasilkan gambar sintesis yang lebih berkualitas.

## DAFTAR PUSTAKA

- RAHAYU, 2022. 2020-2021, *Kerugian Materi akibat Kecelakaan Capai Rp 246 Miliar*. <https://money.kompas.com/read/2022/03/23/151100926/2020-2021-kerugian-materi-akibat-kecelakaan-capai-rp-246-miliar->
- BAUTISTA, T., MATTHEW, J., KERDEGARI, H., PERALTA, L., & XOCHICALE, M., 2022. *Empirical Study of Quality Image Assessment for Synthesis of Fetal Head Ultrasound Imaging with DCGANs*. <https://github.com/budai4medtech/>
- BEANLAND, V., FITZHARRIS, M., YOUNG, K. L., & LENNÉ, M. G., 2013. Driver inattention and driver distraction in serious casualty crashes: Data from the Australian National Crash In-depth Study. *Accident Analysis & Prevention*, 54, 99–107. <https://doi.org/10.1016/J.AAP.2012.12.043>
- BERGSTRA, J., BARDENET, R., BENGIO, Y., & KÉGL, B., 2011. Algorithms for Hyper-Parameter Optimization. *Advances in Neural Information Processing Systems* 24.
- BORJI, A., 2022. Pros and cons of GAN evaluation measures: New developments. *Computer Vision and Image Understanding*, 215. <https://doi.org/10.1016/J.CVIU.2021.103329>
- DINGUS, T. A., GUO, F., LEE, S., ANTIN, J. F., PEREZ, M., BUCHANAN-KING, M., HANKEY, J., & DESIGNED, J. H., 2016. *Driver crash risk factors and prevalence evaluation using naturalistic driving data*. 113(10), 2636–2641. <https://doi.org/10.15787/VTT1VC7C>
- GOODFELLOW, I., BENGIO, Y., & COURVILLE, A., 2018. *Deep learning*. MIT Press. [https://books.google.com/books/about/Deep\\_Learning.html?id=omivDQAAQBAJ](https://books.google.com/books/about/Deep_Learning.html?id=omivDQAAQBAJ)
- HARI, C. V., & SANKARAN, P., 2021. Driver distraction analysis using face pose cues. *Expert Systems with Applications*, 179, 115036. <https://doi.org/10.1016/J.ESWA.2021.115036>
- HU, Y., LU, M., & LU, X., 2020. Feature refinement for image-based driver action recognition via multi-scale attention convolutional neural network. *Signal Processing: Image Communication*, 81.

- <https://doi.org/10.1016/J.IMAGE.2019.115697>
- KOAY, H. V., CHUAH, J. H., CHOW, C.-O., & CHANG, Y.-L., 2022. Detecting and recognizing driver distraction through various data modality using machine learning: A review, recent advances, simplified framework and open challenges (2014–2021). *Engineering Applications of Artificial Intelligence*, 115, 105309. <https://doi.org/10.1016/J.ENGAPPAL.2022.105309>
- LU, Y., CHEN, D., OLANIYI, E., & HUANG, Y., 2022. Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review. *Computers and Electronics in Agriculture*, 200, 107208. <https://doi.org/10.1016/J.COMPAG.2022.107208>
- MCDONALD, A., CARNEY, C., & MCGEHEE, D. V., 2018. *Title Vehicle Owners' Experiences with and Reactions to Advanced Driver Assistance Systems*. [www.aaafoundation.org](http://www.aaafoundation.org)
- NÉE, M., CONTRAND, B., ORRIOLS, L., GIL-JARDINÉ, C., GALÉRA, C., & LAGARDE, E., 2019. Road safety and distraction, results from a responsibility case-control study among a sample of road users interviewed at the emergency room. *Accident Analysis & Prevention*, 122, 19–24. <https://doi.org/10.1016/J.AAP.2018.09.032>
- OMERUSTAOGLU, F., SAKAR, C. O., & KAR, G., 2020. Distracted driver detection by combining in-vehicle and image data using deep learning. *Applied Soft Computing*, 96, 106657. <https://doi.org/10.1016/J.ASOC.2020.106657>
- RADFORD, A., METZ, L., & CHINTALA, S., 2015. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *4th International Conference on Learning Representations*, *ICLR 2016 - Conference Track Proceedings*. <https://doi.org/10.48550/arxiv.1511.06434>
- WAGNER, B., TAFFNER, F., KARACA, S., & KARGE, L., 2022. Vision Based Detection of Driver Cell Phone Usage and Food Consumption. *IEEE Transactions on Intelligent Transportation Systems*, 23(5), 4257–4266. <https://doi.org/10.1109/TITS.2020.3043145>
- WORLD HEALTH ORGANIZATION., 2018. *Global status report on road safety 2018: Summary*. <https://www.who.int/publications/i/item/WHO-NMH-NVI-18.20>
- WU, Q., CHEN, Y., & MENG, J., 2020. Dcgan-based data augmentation for tomato leaf disease identification. *IEEE Access*, 8, 98716–98728. <https://doi.org/10.1109/ACCESS.2020.2997001>
- ZHANG, C., LI, R., KIM, W., YOON, D., & PATRAS, P., 2020. Driver behavior recognition via interwoven deep convolutional neural nets with multi-stream inputs. *IEEE Access*, 8, 191138–191151. <https://doi.org/10.1109/ACCESS.2020.3032344>
- ZHANG, Y. D., & SANGAIAH, A. K., 2022. Cognitive systems and signal processing in image processing. In *Cognitive Systems and Signal Processing in Image Processing*. Elsevier. <https://doi.org/10.1016/C2020-0-02155-9>

*Halaman ini sengaja dikosongkan*