

IMPLEMENTASI KOMBINASI METODE *MEAN DENOISING* DAN *CONVOLUTIONAL NEURAL NETWORK* PADA *FACIAL LANDMARK DETECTION*

I Putu Agus Eka Darma Udayana¹, I Kadek Dwi Gandika Supartha²

^{1,2}Teknik Informatika, STMIK STIKOM Indonesia

e-mail: agus.ekadarma@gmail.com¹, gandika23@gmail.com²

Abstrak

Facial landmark detection merupakan bagian dari *facial recognition*, bertujuan untuk mengidentifikasi titik fokus pada wajah berdasarkan ciri penampakan bagian wajah yang cenderung menonjol, seperti area mata, hidung, bibir, serta tulang pipi. *Facial landmark detection* sering diimplementasikan pada bidang pengenalan wajah, prediksi pose wajah, rekonstruksi wajah 3 dimensi, serta pengembangan sistem deteksi kelelahan karyawan berdasarkan ekspresi wajah. Seiring bertambahnya ketersediaan citra wajah dan kebutuhan proses komputasi yang cepat, metode *Convolutional Neural Network* (CNN) diimplementasikan pada *facial landmark detection*. Namun beragamnya kualitas citra menyebabkan CNN kurang optimal dalam melakukan deteksi. Oleh karena itu guna mengatasi permasalahan terkait kualitas citra ini, diimplementasikan metode *mean denoising* sebagai upaya peningkatan nilai akurasi CNN dalam melakukan pendeteksian *landmark* wajah. *Dataset* citra wajah diperoleh dari *platform Kaggle*, *LFW-People*, *AFLW200* dan *Female Facial Image Dataset*, dengan total sebanyak 2.050 citra wajah, dan terbagi menjadi 2.000 data latih dan 50 data uji. Berdasarkan hasil pengujian, kombinasi metode CNN dengan *mean denoising* menghasilkan peningkatan akurasi yang lebih baik dalam pengenalan objek pada wajah pada kualitas citra yang heterogen dengan rata-rata akurasi pengujian sebesar 81,33%. Akurasi yang cukup baik ini didapatkan karena citra wajah masukan dilakukan penghilangan *noise* terlebih dahulu sehingga fitur dari citra yang seringkali menyebabkan sistem CNN salah dalam mengidentifikasi objek pada wajah dapat diminimalisir.

Kata kunci: *Facial Landmark Detection*, *Denoising*, CNN

Abstract

Facial landmark detection aims to identify focal points on the face based on the appearance characteristics of the facial features that tend to stand out, such as the eye area, nose, lips, and cheekbones. *Facial landmark detection* is often implemented in the fields of facial recognition, facial pose prediction, 3-dimensional facial reconstruction, and the development of employee fatigue detection systems based on facial expressions. Along with the increasing availability of facial images and the need for fast computation processes, the *Convolutional Neural Network* (CNN) method is implemented in facial landmark detection. However, the diversity of image quality causes CNN to be less than optimal in detecting it. Therefore, to solve this problem related to image quality, the *mean denoising* method was implemented as an effort to increase the accuracy value of CNN in detecting facial landmarks. The facial image dataset was obtained from the Kaggle platform, *LFW-People*, *AFLW200*, and *Female Facial Image Dataset*, with a total of 2,050 face images, and divided into 2,000 training data and 50 test data. Based on the test results, the combination of the CNN method with *mean denoising* gained accuracy in recognizing objects on faces at heterogeneous image quality up to 81.33%. This fairly good accuracy is obtained because the input face image is eliminated first so that the features of the image that often cause the CNN system to incorrectly identify objects on the face can be minimized.

Keywords: *Facial Landmark Detection*, *Denoising*, CNN

Diterima Redaksi: 07-12-2020 | Selesai Revisi: 26-01-2021 | Diterbitkan Online: 30-03-2021

DOI: <https://doi.org/10.23887/janapati.v10i1.29779>

PENDAHULUAN

Facial landmark detection merupakan bagian dari *facial recognition*, yang bertujuan untuk mengidentifikasi titik fokus pada wajah berdasarkan ciri penampakan bagian wajah yang cenderung menonjol, seperti area mata, hidung, bibir, serta tulang pipi [1]. *Facial landmark detection* sering diimplementasikan pada bidang pengenalan wajah, prediksi pose wajah, rekonstruksi wajah 3 dimensi, serta pengembangan sistem deteksi kelelahan karyawan berdasarkan ekspresi wajah [2]–[5].

Sesuai dengan definisinya, ketersediaan *dataset* wajah merupakan modal utama yang diperlukan pada penelitian *facial landmark detection*. Namun karena saat ini hal tersebut bukanlah menjadi tantangan yang berarti, mengingat saat ini sudah masuk pada era *big data*, yang berdampak positif terhadap banyaknya ketersediaan *dataset* wajah, mulai dari beragam bentuk, pose, ekspresi, jenis wajah sesuai jenis kelamin, etnis, dan ras manusia [6].

Meskipun ketersediaan *dataset* yang melimpah memiliki dampak positif, namun hal ini juga memiliki sisi negatif jika dilihat dari sisi proses komputasi. Tentunya banyaknya data yang diproses menyebabkan proses komputasi semakin lama. Kondisi ini tentu menjadi tantangan bagi peneliti untuk mengembangkan beragam jenis algoritme *machine learning* yang tangguh dan memiliki waktu komputasi relatif singkat, terlebih saat ini telah masuk pada tahap pengembangan metode *deep learning*, yang merupakan peningkatan metode *machine learning* konvensional [7]. Salah satu metode *deep learning* yang cukup familiar adalah *Convolutional Neural Network* (CNN).

Implementasi CNN dalam pengembangan sistem pengenalan wajah sudah cukup banyak dilakukan, namun tidak jarang sistem yang dikembangkan gagal dalam melakukan rekognisi, terutama pada kasus citra yang memiliki kualitas sangat rendah [8], [9]. Sehubungan dengan permasalahan tersebut, pada makalah ini dipaparkan kombinasi antara CNN dengan metode peningkatan kualitas citra, yakni *mean denoising* dan *mean denoising color* untuk meningkatkan kualitas citra masukan mereduksi derau, sehingga model pengenalan pola yang dibuat menggunakan metode CNN dapat lebih mudah dalam mengenali *landmark* pada wajah.

STUDI LITERATUR

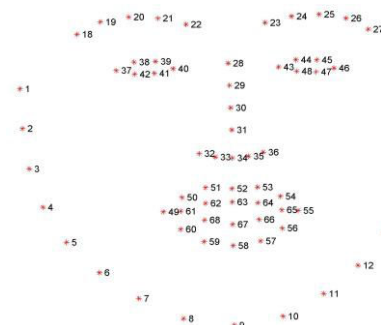
A. State of The Art

Penelitian terkait *Facial Landmark Detection* telah banyak dilakukan oleh beberapa peneliti seperti Xu Zou dkk (2019)

mengimplementasikan metode *Hierarchical Structured Landmark Ensemble* (HSLE) [10]. Pada penelitian lainnya yang dilakukan oleh Jiangjing dkk pada tahun 2017 dijelaskan untuk melakukan deteksi objek pada wajah penulis menggunakan *approach hybrid* antara regresi dan *deep learning* yang menghasilkan reduksi data [11]. Selain itu penelitian lainnya untuk melakukan *facial detection* penulis menggunakan pendekatan adaptif inklusi, sehingga metode yang digunakan dapat mengurangi dampak oklusi dan mendapatkan representasi fitur yang bersih [12]. Penelitian perbandingan lainnya dilakukan analisis terhadap tiga kategori metode *facial landmark* yaitu metode holistik, metode *Constrained Local Model* (CLM), dan metode berbasis regresi. Setiap metode memiliki kelemahan dan keunggulan masing-masing, jadi pada penelitian berikutnya penulis berharap untuk mengkombinasikan metode tersebut untuk mendapatkan hasil yang maksimal [1]. Berdasarkan penelitian pendahulu tersebut, pada penelitian ini ditawarkan model baru melalui hibridisasi metode CNN dengan metode *grayscale enhancement method*, *denoising* dan *denoising color* pada tahap *pre-processing*, yang diharapkan dapat meningkatkan hasil akurasi metode CNN.

B. Facial Landmark Detection

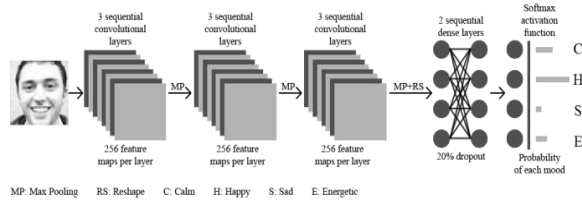
Facial landmark detection adalah salah satu contoh keluaran terstruktur yang bertujuan untuk memprediksi bentuk geometri yang diperoleh dari sebuah data berupa citra wajah [13]. *Facial landmarks* adalah satu set poin penting yang ada pada citra wajah manusia. Jumlah *landmark* bergantung pada dataset atau aplikasi yang digunakan. Salah satu model *facial landmarks* yaitu 68 *landmarks* ditunjukkan pada Gambar 1 [14].



Gambar 1. Penanda Koordinat Bentuk Wajah

C. Convolutional Neural Network

Convolutional Neural Network (CNN) merupakan metode yang diturunkan dari pengembangan dari *Multilayer Perceptron* (MLP) untuk pengolahan data yang bersifat dua dimensi [15]. CNN termasuk dalam rumpun *Deep Neural Network* dikarenakan kedalaman jaringan yang tinggi dan sangat unggul jika diimplementasikan pada data citra. Adapun model CNN dapat dilihat pada Gambar 2.

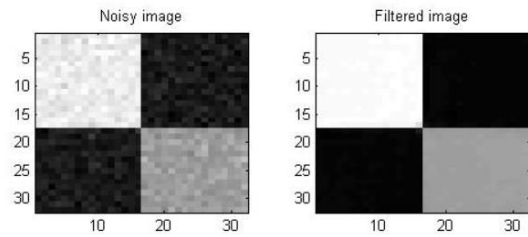


Gambar 2. Konsep Kerja CNN

CNN merupakan metode yang pertama kali dikembangkan dengan nama *NeoCognitron* oleh peneliti asal Jepang bernama Kunihiko Fukushima. Konsep yang dikembangkan oleh Kunihiko Fukushima tersebut kemudian dikembangkan lagi oleh seorang peneliti dari USA atas nama LeChun. LeChun berhasil mengembangkan model awal CNN dengan nama *LeNet* pada penelitian yang membahas pengenalan angka dan tulisan tangan. Penerapan metode CNN semakin hari semakin diminati berkat pada tahun 2012 seorang Alex Krizhevsky berhasil menjuarai kompetisi *ImageNet Large Scale Visual Recognition Challenge 2012* dengan menggunakan metode CNN. Hal ini semakin membuktikan metode CNN sebagai metode klasifikasi objek pada citra terbaik, setelah mengungguli metode *Machine Learning* lainnya seperti SVM.

D. Metode Denoising

Secara umum, wajah merupakan salah satu bagian dari tubuh manusia. Untuk mengamati wajah secara detail kadang harus melakukan pemotongan gambar khusus pada bagian kepala saja, sehingga menyebabkan gambar tidak memiliki kualitas yang bagus dan rentan terhadap masalah *noise*. Untuk mengatasi hal ini, penulis menggunakan algoritma denoising ke dalam sistem sebelum data diproses pada lapisan pertama CNN.



Gambar 3. Proses Denoising

Masalah *Image Denoising* masih menjadi masalah penting di bidang pengolahan citra digital. Sejak penemuan transformasi *Wavelet* dan menghasilkan kinerja yang jauh lebih baik dalam istilah *denoising* citra karena keunggulan metode yang menawarkan fitur tertentu seperti ketersebaran dan struktur multi resolusi dalam kernel. Popularitas transformasi *Wavelet* dalam 20 tahun terakhir berbagai peningkatan algoritme telah memberikan hal baru di lapangan. Pada penelitian ini, berfokus pada penggunaan metode klasik melakukan *denoising* citra yang cenderung mengaburkan tepi tajam citra, merusak garis pada citra, mendistorsi detail halus citra, serta melakukan pengaburan pada derau [16].

F. Pengujian Akurasi

Teknik pengujian akurasi dapat dikatakan sebagai seberapa kedekatan nilai yang dihasilkan dari pengujian dengan nilai yang sebenarnya [17]. Adapun formulasinya dapat dilihat pada Persamaan 1.

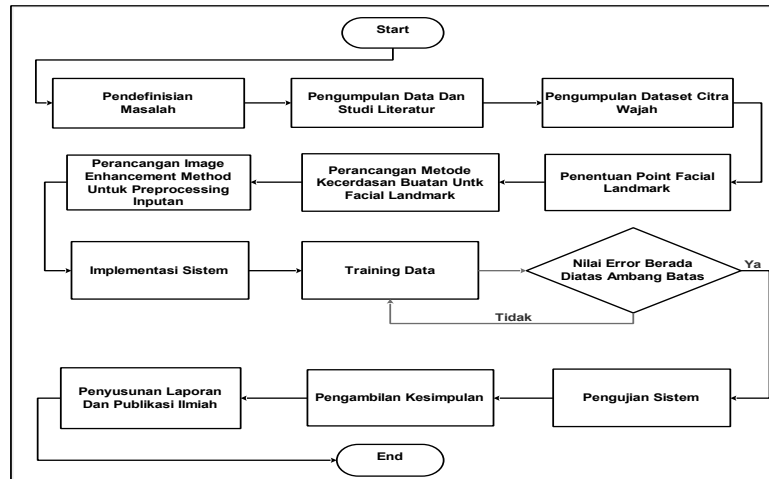
$$Accuracy (\%) = \frac{Correct\ Answer}{Total\ Testing} * 100 \quad (1)$$

Correct Answer adalah satuan yang menyatakan jumlah hasil benar luaran yang dihasilkan sedangkan *total testing* adalah jumlah dari keseluruhan data yang diproses atau diujikan pada sistem [18].

METODE

A. Tahapan Penelitian

Terdapat beberapa tahapan yang dilakukan untuk menyelesaikan penelitian ini, dimana setiap tahapan akan sangat mempengaruhi kinerja dari model yang dihasilkan. Gambar 4 merupakan ilustrasi dari alur penelitian yang dilakukan.

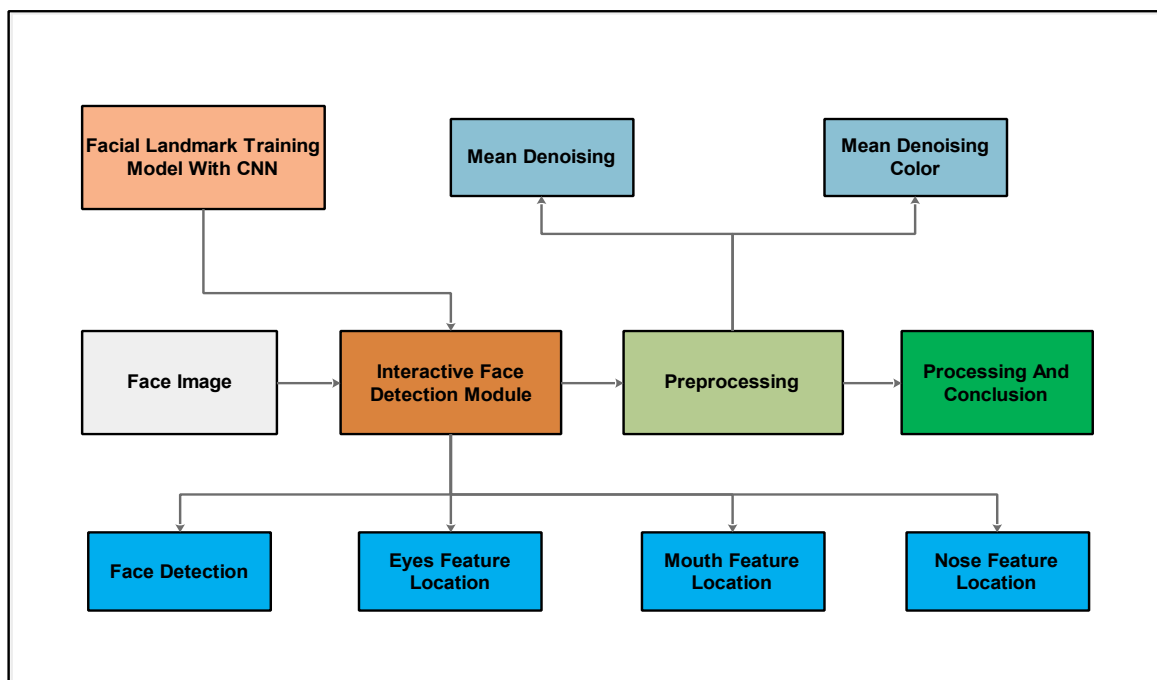


Gambar 4. Tahapan Penelitian

Seperti yang terlihat pada Gambar 4, proses pengembangan sistem dimulai dari pendefinisian masalah dan dilanjutkan dengan pengumpulan data serta studi literatur. *Dataset* yang dikumpulkan dalam penelitian ini adalah berupa data wajah manusia dengan berbagai pose dan ekspresi sehingga nantinya proses *training* data mendapatkan hasil yang bagus ketika mendapatkan masukan data uji dengan beragam bentuk wajah. Setelah mengumpulkan data wajah, pada alur penelitian berikutnya dilaksanakan diskusi untuk menentukan bagian dari wajah apa saja yang akan dicari lokasinya, sehingga sistem dapat secara spesifik

menentukan *facial landmark*. Tidak hanya sampai disitu saja, pada penelitian ini penulis juga berusaha memberikan solusi untuk meningkatkan akurasi sistem dalam menentukan posisi mata, hidung dan mulut wajah. Untuk meningkatkan akurasi tersebut penulis menambahkan metode peningkatan kualitas citra, sehingga sistem nantinya dapat lebih mudah dalam menentukan *facial landmark* dari *input* citra wajah. Tahapan akhir dari penelitian ini adalah dilakukan pengujian dengan melakukan komparasi terhadap metode peningkatan kualitas citra berdasarkan model *training* menggunakan metode CNN.

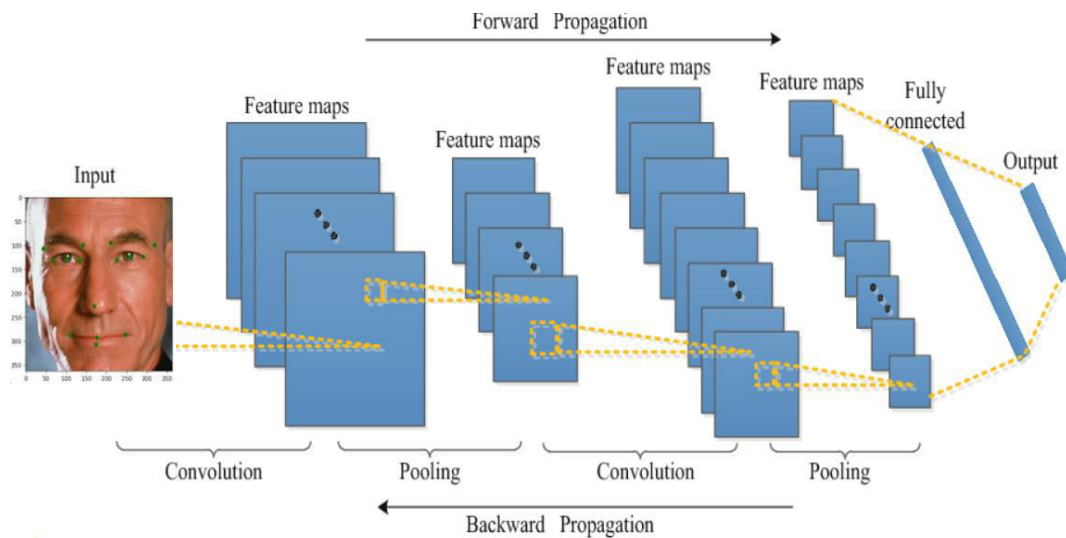
B. Desain Sistem



Gambar 5. Gambaran Umum Sistem

Sistem yang dikembangkan nantinya akan mampu mengenali *landmark* wajah manusia untuk menentukan posisi mata, hidung dan mulut. Tujuan berikutnya dari penentuan posisi tersebut adalah pada penelitian selanjutnya dapat menentukan apakah kondisi mata yang dianalisis sedang tertutup ataupun terbuka yang menandakan seseorang tersebut kelelahan atau mengantuk. Pada penentuan posisi mulut nantinya dapat dianalisa apakah seseorang tersebut sedang menguap atau tidak yang artinya seseorang tersebut sedang dalam keadaan mengantuk.

Seperti terlihat pada Gambar 5 yang masuk pada sistem masuk pada proses *interactive face detection modul* untuk mengetahui posisi dari setiap objek pada wajah berdasarkan hasil data *face training module*. Setelah itu hasil dari proses tersebut akan masuk pada *processing* aplikasi untuk menentukan identitas setiap objek yang dianalisis sampai menentukan keadaan mata dan mulut apakah terindikasi mengalami kelelahan atau tidak. Semua proses tersebut akan penulis kembangkan menggunakan metode CNN seperti terlihat pada Gambar 6.



Gambar 6. Penerapan Metode CNN

Convolution Layer merupakan tahap yang bertujuan untuk melakukan filter terhadap citra wajah yang akan dikenali. Filter tersebut nantinya akan bergeser pada keseluruhan bagian gambar berdasarkan parameter *stride* dan menghasilkan sebuah *output* yang disebut sebagai *activation map* atau *feature map*. Tahap selanjutnya adalah melakukan operasi *pooling* yang terdiri dari sebuah filter dengan *stride* tertentu yang akan bergeser pada seluruh *area activation map* atau *feature map*. Tujuan penggunaan *pooling layer* adalah untuk mengurangi dimensi *feature map*, sehingga mempercepat komputasi karena parameter yang harus di update semakin sedikit dan mengatasi *overfitting*. Tahapan terakhir adalah *fully connected layer* yang bertujuan untuk menghubungkan semua hasil extraction layer

dari citra makanan untuk dilakukan prediksi *fully connected layer* keadaan objek wajah.

1. Pre-processing

Pada proses *pre-processing*, dilaksanakan standarisasi ukuran gambar wajah yang akan dilakukan proses *training dataset* untuk mendapatkan model terbaik dalam mengenali objek gambar wajah pada saat melakukan pengujian sistem. Ukuran gambar yang digunakan dalam penelitian ini adalah 96×96 pixel, sehingga nantinya sistem dapat lebih fokus dalam melakukan operasi untuk menentukan posisi mata, hidung dan mulut.

Seperti terlihat pada Gambar 7, semua gambar yang akan digunakan dalam pengembangan sistem sudah terstandarisasi sehingga nantinya dapat membuat sistem lebih fokus dalam melakukan *processing* citra.

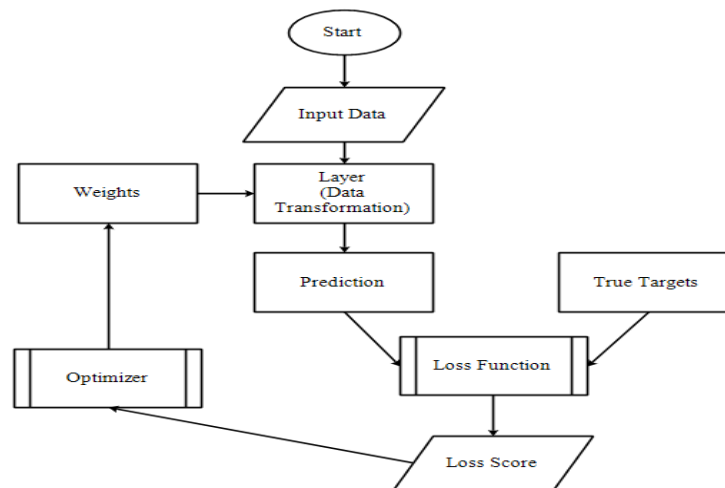


Gambar 7. Standarisasi Dataset

2. Proses Training

Proses *training* pada penelitian ini digunakan untuk mendapatkan akurasi terbaik sistem dalam menentukan posisi mata, hidung dan mulut pada wajah. Fase *training* menggunakan metode CNN ini terdiri dari dua buah proses yaitu *feedforward* dan *backpropagation*. Pada proses *feedforward* dibutuhkan jumlah dan *size* layer yang akan

dibentuk, sedangkan pada proses *backpropagation* pada vektor akan melakukan proses konvolusi dan *max pooling* yang bertujuan untuk mengurangi ukuran gambar serta menduplikasi *neuron*. Untuk membangun kerangka kerja CNN pada tahap ini digunakan *framework machine learning Tensorflow*. *Framework* ini dibangun menggunakan *activation relu*, pemilihan *activation* ini dikarenakan pada proses training sistem tidak akan hanya mencari nilai 0 dan 1, namun akan memprediksi posisi paling mendekati dari mata, mulut dan hidung dengan 0 sampai tak terhingga. Untuk menentukan jumlah *looping* pada proses ini, penulis menggunakan konfigurasi *epochs* sebanyak 25 dengan *batch size* sebesar 50. Penentuan jumlah *epochs* ini dapat berubah sesuai dengan pencapaian nilai *Mean Squared Error* (MSE) yang ditetapkan dalam penelitian, diaman nilai yang harus dicapai dalam pelatihan untuk menjadi model terbaik adalah sebesar 0,001.



Gambar 8. Proses Training Data

Seperti terlihat pada Gambar 8, proses *training* pada penelitian ini akan dihentikan sampai menghasilkan model yang paling optimal, yang pada penelitian ini ambang batas maksimal *loos* dan MSE adalah sebesar 0,01. Model terbaik tersebut nantinya akan digunakan untuk melakukan proses pengujian sistem.

3. Peningkatan Kualitas Citra

Proses peningkatan kualitas citra digunakan untuk mengubah data *input* citra original menjadi citra yang memiliki keutuhan gambar lebih tajam atau detail. Dalam hal ini, metode peningkatan kualitas citra yang

digunakan adalah *mean denoising* dan *mean denoising color* serta merubah semua data *input* ke dalam bentuk *grayscale*. Tujuan dari merubah semua gambar menjadi *grayscale* adalah untuk mengurangi fitur-fitur yang ada pada gambar, sehingga sistem akan lebih mudah dalam melakukan proses *facial landmark*.

C. Pengujian Sistem

Untuk melakukan proses pengujian sistem, penulis menggunakan pengujian akurasi sistem. Pengujian akurasi ini digunakan untuk menilai komparasi penggunaan metode

peningkatan kualitas citra dalam mengenali objek pada wajah. Adapun banyak data yang digunakan dalam setiap pengujian adalah 50 data wajah di luar data *training* dan persamaan yang digunakan untuk menentukan akurasi sistem tertuang pada Persamaan 1.

HASIL DAN PEMBAHASAN

A. Hasil *Training* Metode CNN

Proses *training* yang dilakukan pada penelitian ini telah menghasilkan model terbaik karena memenuhi ambang batas minimal *loss* dan *mse* yang telah ditentukan pada perancangan sistem. Untuk melakukan *training*,

pada penelitian ini menggunakan sebanyak 2.000 data citra wajah. *Dataset* citra wajah yang digunakan dalam proses *training* pada penelitian ini adalah kombinasi dari beberapa *dataset* yang dapat diakses secara bebas oleh publik pada situs [kaggle.com](https://www.kaggle.com) serta cvlab.cse.msu.edu dengan nama *LFW-People*, *AFLW200* dan *Female Facial Image dataset*. Penggabungan tiga *dataset* ini dilaksanakan untuk membuat pembelajaran pada sistem mendapatkan *input* data yang beragam sehingga model yang dihasilkan akan memiliki akurasi tinggi dalam mengenali objek wajah pada kondisi wajah yang beragam.

```

00/30 [.....] - 25s 835ms/step - loss: 0.0363 - mse: 0.0363 - val_loss: 0.0360 - val_mse: 0.0360
Epoch 8/25
00/30 [.....] - 25s 834ms/step - loss: 0.0357 - mse: 0.0357 - val_loss: 0.0354 - val_mse: 0.0354
Epoch 9/25
00/30 [.....] - 25s 834ms/step - loss: 0.0350 - mse: 0.0350 - val_loss: 0.0346 - val_mse: 0.0346
Epoch 10/25
00/30 [.....] - 25s 835ms/step - loss: 0.0341 - mse: 0.0341 - val_loss: 0.0335 - val_mse: 0.0335
Epoch 11/25
00/30 [.....] - 25s 834ms/step - loss: 0.0329 - mse: 0.0329 - val_loss: 0.0321 - val_mse: 0.0321
Epoch 12/25
00/30 [.....] - 25s 836ms/step - loss: 0.0312 - mse: 0.0312 - val_loss: 0.0301 - val_mse: 0.0301
Epoch 13/25
00/30 [.....] - 25s 834ms/step - loss: 0.0287 - mse: 0.0287 - val_loss: 0.0270 - val_mse: 0.0270
Epoch 14/25
00/30 [.....] - 25s 834ms/step - loss: 0.0248 - mse: 0.0248 - val_loss: 0.0220 - val_mse: 0.0220
Epoch 15/25
00/30 [.....] - 25s 835ms/step - loss: 0.0186 - mse: 0.0186 - val_loss: 0.0145 - val_mse: 0.0145
Epoch 16/25
00/30 [.....] - 25s 834ms/step - loss: 0.0109 - mse: 0.0109 - val_loss: 0.0077 - val_mse: 0.0077
Epoch 17/25
00/30 [.....] - 25s 834ms/step - loss: 0.0065 - mse: 0.0065 - val_loss: 0.0055 - val_mse: 0.0055
Epoch 18/25
00/30 [.....] - 25s 834ms/step - loss: 0.0049 - mse: 0.0049 - val_loss: 0.0043 - val_mse: 0.0043
Epoch 19/25
00/30 [.....] - 25s 835ms/step - loss: 0.0038 - mse: 0.0038 - val_loss: 0.0034 - val_mse: 0.0034
Epoch 20/25
00/30 [.....] - 25s 834ms/step - loss: 0.0031 - mse: 0.0031 - val_loss: 0.0028 - val_mse: 0.0028
Epoch 21/25
00/30 [.....] - 25s 834ms/step - loss: 0.0026 - mse: 0.0026 - val_loss: 0.0024 - val_mse: 0.0024
Epoch 22/25
00/30 [.....] - 25s 835ms/step - loss: 0.0023 - mse: 0.0023 - val_loss: 0.0021 - val_mse: 0.0021
Epoch 23/25
00/30 [.....] - 25s 835ms/step - loss: 0.0020 - mse: 0.0020 - val_loss: 0.0019 - val_mse: 0.0019
Epoch 24/25
00/30 [.....] - 25s 835ms/step - loss: 0.0018 - mse: 0.0018 - val_loss: 0.0017 - val_mse: 0.0017
Epoch 25/25
00/30 [.....] - 25s 835ms/step - loss: 0.0016 - mse: 0.0016 - val_loss: 0.0016 - val_mse: 0.0016

```

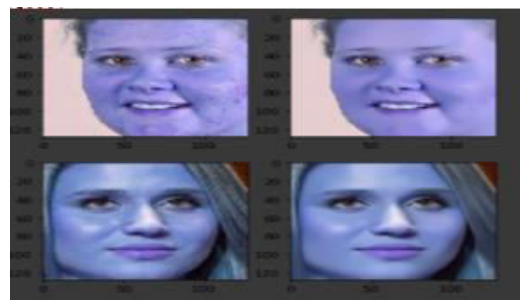
Gambar 9. Hasil *Training* Data

Berdasarkan Gambar 9, proses *training* yang dilakukan oleh sistem menggunakan metode CNN telah menghasilkan *loss* dan MSE sebesar 0,0016. Data ini menunjukkan model sudah layak digunakan untuk melakukan operasi *facial landmark*, karena sudah memenuhi ambang batas maksimal *loss* dan MSE proses *training*.

B. Hasil Peningkatan Kualitas Citra

Peningkatan kualitas citra digunakan untuk mengurangi kesalahan sistem dalam mengenali objek-objek yang terdapat pada wajah citra *input*. Pada percobaan pertama penulis menggunakan metode *Mean Denoising*

untuk melakukan peningkatan kualitas citra, Gambar 10 adalah hasil dari penggunaan metode ini dalam menghilangkan *noise* pada citra wajah.



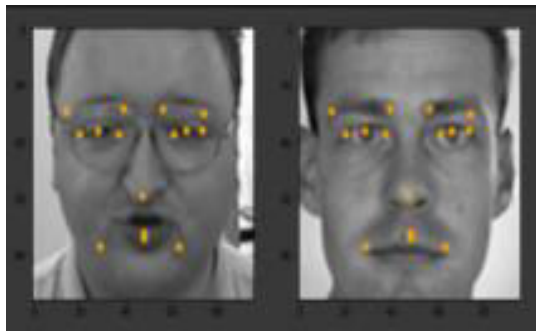
Gambar 10. Mean Denoising

Seperti terlihat pada Gambar 10, secara umum terjadi peningkatan kualitas citra setelah dilakukan koreksi citra pada citra *input*. Peningkatan kualitas citra ini dapat terlihat dari penghilangan *noise* pada citra yang bertujuan untuk membuat citra terlihat lebih tajam sehingga sistem akan lebih mudah dalam mengenali objek yang terdapat pada wajah.

Hampir sama dengan metode sebelumnya, pada metode *Mean Denoising Color input* gambar yang digunakan tidak diubah menjadi *grayscale* terlebih dahulu. Proses ini bertujuan untuk menghilangkan *noise* pada citra dan menstandarkan nilai warna pada citra, sehingga nantinya sistem akan lebih mudah dalam mengenali objek pada wajah.

C. Hasil Dan Analisa Komparasi Metode

Berlatar belakang untuk meningkatkan akurasi sistem yang dikembangkan dengan metode CNN dalam mengenali objek pada wajah. Pengujian ini digunakan untuk membantu model yang dikembangkan mendapatkan akurasi yang lebih baik lagi dengan menambahkan *pre-processing* sebelum citra wajah masuk ke dalam proses CNN.



Gambar 11. Hasil Pengujian Facial Landmark

Berdasarkan hasil komparasi metode peningkatan kualitas citra yang digunakan, citra wajah untuk pengujian sebanyak 50 data gambar yang sebelumnya diolah menggunakan metode *Mean Denoising* memiliki nilai akurasi yang tertinggi, berikut ini adalah pemaparan dari setiap pengujian yang telah dilakukan.

Tabel 1. Pengujian Metode CNN

Facial Landmark	Benar	Salah	Akurasi (%)	Loss (%)
Mata	50	0	100	0
Hidung	22	28	44	56
Mulut	29	21	58	42

Rata-rata	67.33	32.67
-----------	-------	-------

Hasil akurasi pengujian yang didapatkan tanpa menggunakan metode peningkatan kualitas citra sebesar 67,33%.

Tabel 2. Pengujian Kombinasi CNN dan Metode Mean Denoising

Facial Landmark	Benar	Salah	Akurasi (%)	Loss (%)
Mata	50	0	100	0
Hidung	32	18	64	36
Mulut	40	10	80	20
Rata-rata			81.33	18.67

Berdasarkan hasil pengujian pada Tabel 2, nilai akurasi rata-rata yang didapatkan dengan penerapan metode ini adalah sebesar 81,33%. Dari tabel tersebut juga dapat dianalisa bahwa rata-rata *loss* atau kesalahan yang terjadi dengan kombinasi metode ini ketika mengenali objek yang terdapat pada wajah sebesar 27,33%.

Tabel 3. Pengujian Kombinasi Metode CNN dan Mean Denoising Color

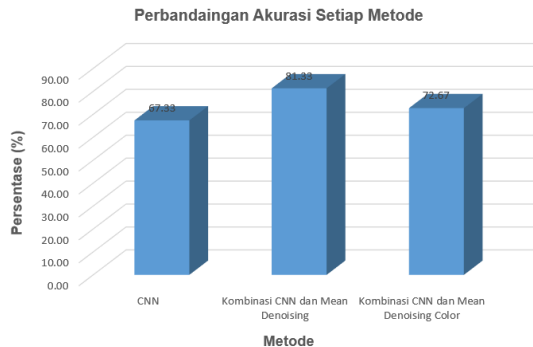
Facial Landmark	Benar	Salah	Akurasi (%)	Loss (%)
Mata	50	0	100	0
Hidung	26	24	52	48
Mulut	33	17	66	34
Rata-rata			72.67	27.33

Berbeda dengan hasil pengujian menggunakan metode sebelumnya, metode *mean denoising color* menghasilkan akurasi sebesar 72,67%.

Tabel 4. Komparasi Metode

Metode	Jumlah Posisi Benar			Rata-Rata Akurasi (%)
	Mata	Hidung	Mulut	
CNN	50	22	29	67.33
CNN dan Mean Denoising	50	32	40	81.33
CNN dan Mean Denoising Color	50	26	33	72.67

Tabel 4 merupakan rekapitulasi dari hasil pengujian dari setiap metode. Untuk lebih jelas dan mudah untuk diamati, Gambar 12 merupakan representasi rekapitulasi tersebut dalam bentuk diagram batang.



Gambar 12. Grafik Komparasi Metode

Ketika dibandingkan, ketiga hasil pengujian tersebut menghasilkan perbandingan yang cukup signifikan. Hasil pengujian akurasi terbaik didapat ketika citra masukan sebelumnya dilakukan peningkatan kualitas citra menggunakan metode *mean denoising*. Akurasi yang cukup baik ini didapatkan karena citra wajah masukan dilakukan penghilangan *noise* terlebih dahulu sehingga fitur dari citra yang seringkali menyebabkan sistem CNN salah dalam mengidentifikasi objek pada wajah dapat diminimalisir. Untuk pengujian menggunakan *mean denoising color* yang hasil pengujiannya dibawah metode *mean denoising* dikarenakan masih banyaknya fitur warna dari citra wajah, sehingga sistem masih agak mengalami kesulitan dalam memprediksi posisi objek pada wajah.

KESIMPULAN

Kesimpulan yang diperoleh dari komparasi metode peningkatan kualitas citra ini adalah metode *Mean Denoising* dapat meningkatkan akurasi dari sistem kecerdasan buatan menggunakan metode CNN dalam mengenali objek yang ada pada wajah atau *facial landmark*. Hasil akurasi yang didapatkan dari pengujian ini adalah kombinasi metode *Mean Denoising* sebagai *pre-processing* gambar wajah sebelum dilakukan pengenalan objek menggunakan metode CNN menghasilkan peningkatan akurasi dibandingkan tanpa melalui tahap *preprocessing* dengan akurasi sebesar 81,33%. Hasil pengujian ini dapat diartikan perpaduan kedua metode ini cocok digunakan untuk proses pengenalan citra ketika dihadapkan pada sumber citra yang kurang tajam atau masih

banyak fitur gambar yang dapat mengganggu proses pengenalan objek.

UCAPAN TERIMAKASIH

Penulis mengucapkan terimakasih kepada LPPM STMIK STIKOM Indonesia yang telah memberikan pendanaan dalam pelaksanaan penelitian ini.

REFERENSI

- [1] Y. Wu and Q. Ji, "Facial Landmark Detection: A Literature Survey," *Int. J. Comput. Vis.*, vol. 127, no. 2, pp. 115–142, 2019, doi: 10.1007/s11263-018-1097-z.
- [2] A. Juhong and C. Pintavirooj, "Face recognition based on facial landmark detection," *BMEiCON 2017 - 10th Biomed. Eng. Int. Conf.*, vol. 2017-Janua, no. December, pp. 1–4, 2017, doi: 10.1109/BMEiCON.2017.8229173.
- [3] Z. Liu, Z. Chen, J. Bai, S. Li, and S. Lian, "Facial pose estimation by deep learning from label distributions," *Proc. - 2019 Int. Conf. Comput. Vis. Work. ICCVW 2019*, pp. 1232–1240, 2019, doi: 10.1109/ICCVW.2019.00156.
- [4] J. Booth *et al.*, "3D Reconstruction of 'In-the-Wild' Faces in Images and Videos," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 8, pp. 1–14, 2018.
- [5] C. Zhang, X. Lu, Z. Huang, S. Xia, and C. Fu, "A Driver Fatigue Recognition Algorithm Based on Spatio-Temporal Feature Sequence," *Proc. - 2019 12th Int. Congr. Image Signal Process. Biomed. Eng. Informatics, CISP-BMEI 2019*, 2019, doi: 10.1109/CISP-BMEI48845.2019.8965990.
- [6] M. J. Baculo and J. Azcarraga, "Emotion Recognition on Selected Facial Landmarks Using Supervised Learning Algorithms," in *Proceedings - 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018*, 2019, pp. 1483–1489, doi: 10.1109/SMC.2018.00258.
- [7] H. Gao and L. Yuan, "Research on key technology of pavement object recognition based on machine learning," *Neural Comput. Appl.*, vol. 32, no. 10, pp. 5483–5493, 2020, doi: 10.1007/s00521-019-04643-6.
- [8] K. Topare, T. Pawar, Y. Patil, and N. Dudhagawali, "IP based drowsiness detector for Real Driving Conditions," no. 20160301, pp. 4–7, 2016.
- [9] S. Junaedi and H. Akbar, "Driver Drowsiness Detection Based on Face Feature and PERCLOS," *J. Phys. Conf. Ser.*, vol. 1090, no. 1, 2018, doi:

- 10.1088/1742-6596/1090/1/012037.
- [10] X. Zou, S. Zhong, L. Yan, X. Zhao, J. Zhou, and Y. Wu, "Learning robust facial landmark detection via hierarchical structured ensemble," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, vol. 2019-October, pp. 141–150, doi: 10.1109/ICCV.2019.00023.
- [11] J. Lv, X. Shao, J. Xing, C. Cheng, and X. Zhou, "A deep regression architecture with two-stage re-initialization for high performance facial landmark detection," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 3691–3700, 2017, doi: 10.1109/CVPR.2017.393.
- [12] M. Zhu, D. Shi, M. Zheng, and M. Sadiq, "Robust facial landmark detection via occlusion-adaptive deep networks," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2019-June, pp. 3481–3491, 2019, doi: 10.1109/CVPR.2019.00360.
- [13] R. Jabbar, K. Al-Khalifa, M. Kharbeche, W. Alhajyaseen, M. Jafari, and S. Jiang, "Real-time Driver Drowsiness Detection for Android Application Using Deep Neural Networks Techniques," *Procedia Comput. Sci.*, vol. 130, pp. 400–407, 2018, doi: 10.1016/j.procs.2018.04.060.
- [14] S. A. Lee, J. Kim, J. M. Lee, Y.-J. Hong, I.-J. Kim, and J. D. Lee, "Automatic Facial Recognition System Assisted-facial Asymmetry Scale Using Facial Landmarks," *Otol. Neurotol.*, vol. 41, no. 8, pp. 1140–1148, 2020, doi: 10.1097/mao.0000000000002735.
- [15] A. Y. Wijaya, R. Soelaiman, J. T. Informatika, and F. T. Informasi, "Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) Pada Caltech 101 Image Classification Using Convolution Neural Network (Cnn) on Caltech 101," vol. 5, no. 1, 2016.
- [16] C. Tian, Y. Xu, Z. Li, W. Zuo, L. Fei, and H. Liu, "Attention-guided CNN for image denoising," *Neural Networks*, vol. 124, pp. 117–129, 2020, doi: 10.1016/j.neunet.2019.12.024.
- [17] K. A. Aryani, D. G. H. Divayana, and I. M. A. Wirawan, "Sistem Pakar Diagnosis Penyakit Jerawat di Wajah dengan Metode Certainty Factor," *J. Nas. Pendidik. Tek. Inform.*, vol. 6, no. 2, p. 96, 2017, doi: 10.23887/janapati.v6i2.11496.
- [18] I. P. A. E. D. U. Udayana and P. G. S. C. Nugraha, "Prediksi Citra Makanan Menggunakan Convolutional Neural Network Untuk Menentukan Besaran Kalori Makanan," *J. Teknol. Inf. dan Komput.*, vol. 6, no. 1, pp. 30–38, 2020.