DROWSY DETECTION FROM VIDEO DRIVER FACE BASED ON EYE AND MOUTH FEATURES EXTRACTION USING THE CONVOLUTION NEURAL NETWORK METHOD

Akhmad Solikin*, Endang Setyati
Department of Information Technology
Institute of Sains and Teknologi Terpadu Surabaya
*Corresponding email address: solikinakhmad@yahoo.co.id

Abstract — This research was conducted in an effort to minimize the occurrence of road traffic accidents. In this study detected the level of fatigue and sleepiness from the driver's face video based on the extraction of eye and mouth features using the CNN method. The dataset in this study is 300 data with 3 different classes namely drowsiness 100 data, sleepy 100 data and normal 100 data. The number of epochs used in research to achieve high accuracy is as much as 50. In the test results it is known that the validation of accuracy has increased in each of the input layer results.

Keywords: Driver, Drowsiness, Eyes, Mouth, CNN.

I. INTRODUCTION

In driving, drivers often ignore their physical condition, from fatigue to drowsiness. Fatigue or sleepiness, these two factors are the cause of most road traffic accidents. Driving in a sleepy state is a very big problem when we are on the road. Driving in a sleepy state can result in accidents, both minor accidents and accidents that are very severe and even result in death[1-3].

Drowsiness is a concern in safety, especially for drivers who need high concentration where the driver is required to stay focused for a long time. Driver drowsiness is a big influence on road accidents. This can be prevented by using technological capabilities[2-3].

To prevent accidents that are not desired, it is necessary to build a sleepy detection system on the driver. In this case it is expected to be used as a support to help reduce the number of accidents that occur on the road[4].

Convolution Neural Networks (CNN) is a category of Neural Networks that has proven to be very effective in areas such as image recognition and classification. CNN was first developed under the name NeoCognitron by Kunihiko Fukushima, who is a researcher from NHK Broadcasting Science Research Laboratories, Kinuta, Setagaya, Tokyo, Japan. The concept was further developed by Yann LeChun, a researcher from AT&T Bell Laboratories in Holmdel, New Jersey, USA[5].

From the problems that have been studied, in this study it has been proposed that the detection of sleepiness from the video of the driver's face based on the extraction of eye and mouth features using the Convolution Neural Networks method. To produce a more accurate classification of eye and mouth feature extraction this study uses a CNN method model[6].

II. RESEARCH METHODS

In this research the process is divided into two, namely the formation of a dataset and the design of parameters. The formation of the dataset discusses how to prepare data for this study, while the design of parameters discusses what parameters are used when conducting trials in this study.
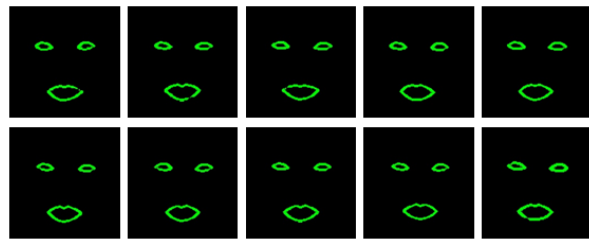
A. Dataset Formation

In this study there are 3 different dataset classes with a total of 300 data. Details in detail in the dataset are: Drowsiness 100 data, Sleepiness 100 data, and Normal 100 data. The sample data can be seen in Figure 1.
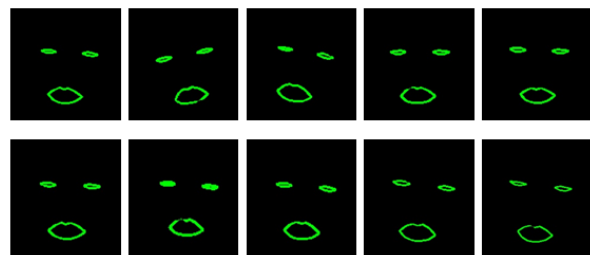
a. Drowsiness
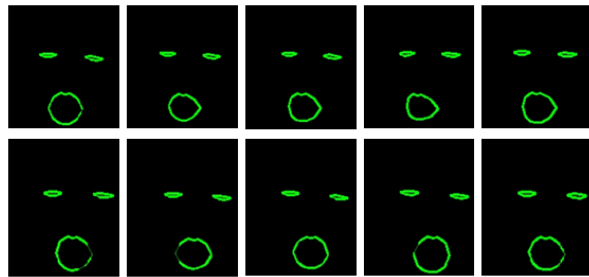
b. Sleepiness

c. Normal

Figure 1. Examples of data from each class

In the RGB image, the drowsiness dataset formation stage is explained in detail as follows:

1. Taking done with a program that has been made to look for faces, then determine the location of the eyes and mouth to avoid interference with the background, the image is changed to black, and only the eyes and mouth have color[7].

2. In collecting data manually done one by one with various conditions between sleepiness and sleepy.

3. Then the image obtained from point 2 is resized and equated with other images. After changing and matching, the image will be saved.

The design of the process of forming a dataset can be seen in Figure 2 below.
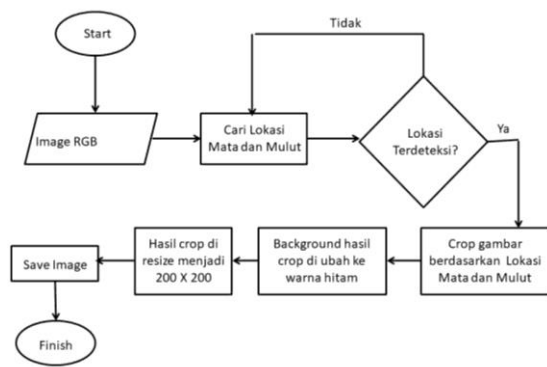

Figure 2. Dataset Formation Process

B. Parameter Design

The design parameters in this study use the CNN model architecture. Convolutional Neural Netwok (CNN) is a scientific field in the field of machine learning that is developing quite rapidly, especially for classifying images[8]. There is also an architectural arrangement of the CNN model shown in Figure 3.
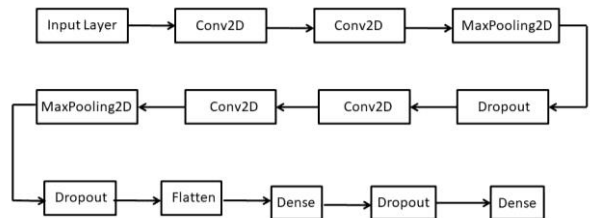

Figure 3. CNN Architecture

Input image in this study is 200x200. The batch size used for training is 32[9]. While the dataset used in this study consists of 3 different classes. Next is to determine the number of epochs or iterations that we use for training. The number of epochs used in research to achieve high accuracy is as much as 50. Then there is a number of filters included in the convolution process. At the convolution stage, 64 filters were used.

Table 1. Testing the Parameters

Epoch	train ACC	train Loss	val ACC	val Loss
1	0,65	0,890426326	0,900000004	0,387284698
2	0,866666667	0,32825255	0,949999988	0,143961372
3	0,9625	0,083055532	0,966666675	0,187872485
4	0,9875	0,050347499	0,949999988	0,116909252
5	0,995833333	0,030732037	0,966666675	0,185755855
6	0,9875	0,04397945	0,916666655	0,371027031
7	0,975	0,069915423	0,949999988	0,092684374
8	0,991666667	0,025796949	0,966666675	0,103127208
9	0,995833333	0,021433797	0,966666675	0,116941646
10	0,991666667	0,0368429	0,966666675	0,182349469
-	-	-	-	-
-	-	-	-	-
-	-	-	-	-
50	1	0,000037556	0,983333341	0,0833439

In table 1 it has been explained that to get high accuracy values, iterated 50 times.

III. RESULTS AND DISCUSSION

At this stage, it is discussed about the results of the training and testing conducted and the output of the models that have been made. After going through the training process, the

program will display the output in the form of accuracy values from the training and validation for each iteration[10]. In the test results it is known that the validation accuracy has increased in each input layer results. To see in detail the results of the trial here the author will show the graph in Figure 4.
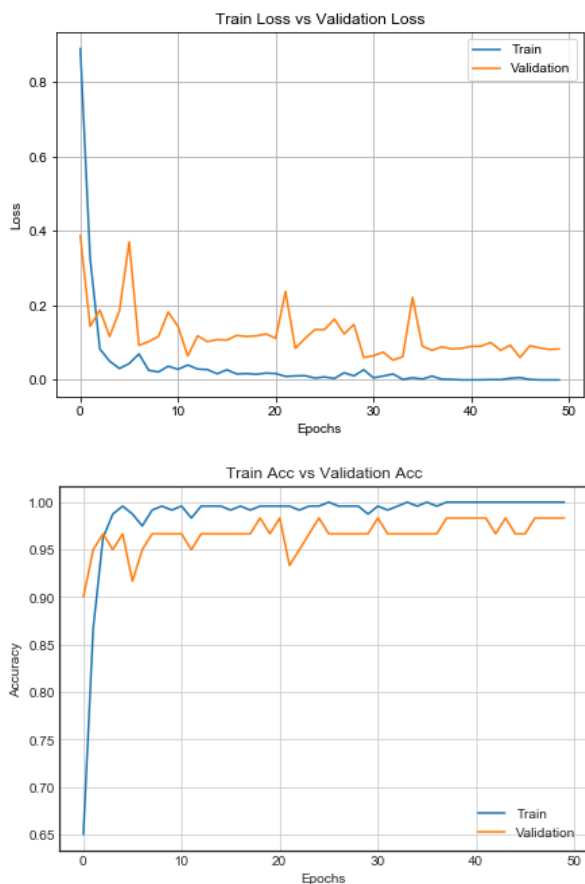

Figure 4. Accuracy and Loss Gradient Graphs

Table 1 explains the average value of accuracy and data loss. Each is explained in detail in table 2.

Table 2. The average value of the trial results

%	Train ACC	Train Loss	Val ACC	Val Loss
Rata-rata	98	0,4	96	12

From the value of the trial results in table 2, a display has been generated in the program as shown in Figure 5. The outputs of the program are listed in 3 different labels, they are drowsiness, sleepiness, and normal. The highest level of validation accuracy is 96%.

Figure 5. Program Trial

IV. CONCLUSION

Classification using Convolutional Neural Networks is able to handle complex features such as drowsiness, sleepiness, and normal labeling in the current study[11]. The process of forming a dataset is done manually one by one with various conditions between drowsiness and sleepiness. From the test results it is known that the validation accuracy has increased in each input layer results. The highest accuracy validation value is 96%.

ACKNOWLEDGMENT

REFERENCE

[1] Wanghua Deng, Ruoxue Wu, "Real-Time Driver-Drowsiness Detection System Using Facial Features", IEEE (2019)

[2] Elis Fitrianingsih, Endang Setyati, Luqman Zaman,"Extraction of Eye and Mouth Features for Drowsiness Face Detection Using Neural Network", KINETIK (2018)

[3] Venkata Rami Reddy Chira, Srinivasulu Reddy Uyyala, Venkata Krishna Kishore Kolli, "Deep CNN: A Mechine Learning Approach For Driver Drowsiness Detection Based on Eye State", IIETA (2019)

[4] Sanghyuk Park, Fei Pan, Sunghun Kang, Chang D Yoo, "Driver Drowsiness detection system based on feature representation learning using various deep networks", KAIST (2019)

[5] I Wayan Suartika E P, Arya Yudhi Wijaya, Rully Soelaiman, "Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101", JURNAL TEKNIK ITS (2016)

[6] Muhammad Zufar, Budi Setiyono, "Convolutional Neural Networks untuk Pengenalan Wajah Secara Real-Time", JURNAL SAINS DAN SENI ITS (2016)

[7] Syamsul Bahri, Rachmat, "Transformasi Citra Biner Menggunakan Metode Thresholding dan Otsu Thresholding", JURNAL SISTEM INFORMASI DAN TEKNOLOGI INFORMASI (2018)

[8] William Sugiarto, Yosi Kristian, Eka Rahayu Setyaningsih, "Estimasi Arah Tatapan Mata Menggunakan Ensemble Convolutional Neural Network", TEKNIKA (2018)

[9] Tri Septianto, Endang Setyati, Joan Santoso, "Model CNN LeNet dalam Rekognisi Angka Tahun pada Prasasti Peninggalan Kerajaan Majapahit", JTSISKOM (2018)

[10] Ade Muhammad Rizki, Nola Marina, "Klasifikasi Kerusakan Bangunan Sekolah Menggunakan Metode Convolutional Neural Network Dengan Pre-trained Model VGG-16", Jurnal Ilmiah Teknologi dan Rekayasa (2019).

[11] Stephen Ekaputra Limantoro, Yosi Kristian, Devi Dwi Purwanto, "Deteksi Pengendara Sepeda Motor Menggunakan Deep Convolutional Neural Networks", SEMANTIKOM (2017)