# Detecting Sleep Behaviors In Class Using OpenCV Programming

Research Report

**Made by:**
Aleta Josephine Lukman/XII.1 - 02
Felisha Angeline/XII.5 - 09

**SMA SANTA LAURENSIA**
Jl. Sutera Utama No. 1 Alam Sutera, Pakulonan, Serpong Utara,
Tangerang Selatan, Banten 15325
2024

**ABSTRACT**

More than 2 out of 3 high school students don't get enough sleep which increases the risk of health problems including obesity, diabetes, poor mental health, attention and behaviour issues, and poor academic performance. Due to insufficient sleep, an increasing number of high school students sleeping in class happens worldwide which leads to poor scores that affect their GPAs. This research aims to create a program that would be able to detect students that fall asleep in class. It is created using python programming language by using OpenCV library in Visual Studio Code. The result from this program is then processed and analyzed using a confusion matrix that provides the accuracy, precision, recall, and F1-Score between 3 different dataset sample groups. Upon analyzing the accuracy of the program, it was found that as more dataset samples are taken, the more accurate the programme is on detecting sleep behaviours. However, the accuracy percentage of the programme run is still insufficient which indicates that the programme is prone to error and have not been developed to be effective enough to detect live sleeping behaviours.

Keywords: *sleep, program, OpenCV*

## FOREWORD

Dear Readers,

With respect, we would like to extend our deepest gratitude for your time and attention in reviewing this research report. This study is driven by the urgent need to discover alternative solutions to reduce the number of students sleeping in class. This led us to make a report with the title “Detecting Sleep Behaviors In Class Using OpenCV Programming”

We would also like to express our deepest gratitude to Mr. Liem Cia Ming and Mr. Budhy Mulyana for their guidance throughout our process, from the planning phase of the research to finalizing the report. We also wish to thank all parties involved: teachers, friends, and parents who have supported and prayed for us, allowing our research to proceed well. This research is far from perfect, and we welcome any feedback in the form of criticism and suggestions that can help improve this study.

We sincerely apologize for any shortcomings in this research report, whether in content, explanation of presentation. Finally we hope that this research can serve as available source of knowledge and information, as well as a programmatic tool to assess an individual’s sleep condition effectively

Sincerely,

Aleta Josephine Lukman
&
Felisha Angeline

# TABLE OF CONTENTS

## TABLE OF FIGURES

# CHAPTER I
# INTRODUCTION

## 1.1 Background

According to the American Physical Therapy Association, more than 2 out of 3 high school students aren't getting enough sleep. The US Centers for Disease Control and Prevention (CDC) warns that insufficient sleep can increase the risk for a host of health problems including obesity, diabetes, injuries, poor mental health, attention and behaviour problems, and poor academic performance. The American Academy of Sleep Medicine found that among middle school students, 57.8% reported insufficient sleep, with nearly 12% reporting sleeping fewer than 6 hours a night. While among high school students, 72,7% reported insufficient sleep with about 20% reporting sleeping fewer than 6 hours a night.

Students sleeping in class has been a daily occurrence in every single school in the world. This has been a concern for schools, especially teachers, who felt like they were disrespected and are not appreciated when students do not give their full attention to the teacher. Furthermore, students getting caught sleeping in class has led them to have poor behaviour scores, affecting their GPA. However, this could not be perceived as students being lazy or being disrespectful in class. Instead, this is a concern that shouldn't be overlooked. The reason why a student is sleeping in class, whether it is on purpose or not, is mostly caused by the lack of sleep students get the night before.

Adequate sleep contributes to a student's overall health and well-being. Students should get the proper amount of sleep at night to help stay focused, improve concentration, and improve academic performance. There may be a ton of personal reasons as to why students aren't getting enough sleep. Usually, they would say that it is due to the pile of homeworks and tests that are given by school. However, it is inevitable that the amount of time they're spending on their devices, whether it is playing games or chatting with friends online, also contributes to this issue. Aside from these general problems, each and every one of them would have their own personal reason. Some of them might have psychological issues or family issues, which could lead them to have insomnia at night. These are very concerning as it does not only affect their personal lives, but also affecting their academic performance at school.

Regardless of these supporting arguments, we would like to analyse and identify sleeping behaviours in class through programming. This programme will be able to identify students who are falling asleep in class through a connected CCTV camera, which will then capture and send the data to the destination file, where all of the data is stored. The data will then be received by the teachers and they would be able to see the students who are detected by the programme.

## 1.2 Problem Statement

Students sleeping in class has been a daily occurrence in every single school in the world. This has been a concern for schools, especially teachers, who felt like they were disrespected and are not appreciated when students do not give their full attention to the teacher. Furthermore, students getting caught sleeping in class has led them to have poor behaviour scores, affecting their GPA.

## 1.3 Research Questions

1. Can the programme detect students sleeping in class?
2. Is the programme effective enough to detect students sleeping in class?

## 1.4 Purpose of Research

1. To create a program that will be able to help detect students that are asleep in class
2. To analyse the accuracy of the programme on detecting students sleeping in class

## 1.5 Benefits of Research

This programme is expected to detect students sleeping in class. This expectation is related to the concern that may arise regarding students sleeping in class. By using this program, students falling asleep in class would be easier to detect and get alerted about it. We hope that the detection and alert will give more attention towards the students who might have trouble sleeping at home thus end up sleeping at school.

# CHAPTER II
# LITERATURE REVIEW

## 2.1 Basic Theory

### 2.1 OpenCV

OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. It aims to provide a shared framework for computer vision applications and to enhance the use of machine perception. The library primarily focuses on image processing and video analysis, featuring capabilities like face and object detection. OpenCV is written natively in C++, and also supports other programming languages such as Python and Java.

With over 2,500 optimised algorithms, OpenCV includes a wide set of both classic and state-of-the-art computer vision and machine learning algorithms. These algorithms can be utilised for various tasks, such as detecting and recognizing faces, identifying objects, classifying actions in videos, tracking camera movements and moving objects, retrieving similar images from a database, and monitoring eye movements.

### 2.2 Python Programming Language

Python is a general-purpose open source computer programming language, optimised for quality, productivity, portability, and integration. It is used by hundreds of thousands of developers around the world, in areas such as Internet scripting, systems programming, user interfaces, product customization, and more. Although general-purpose, Python is often called a scripting language because it makes it easy to utilise and direct other software components. (Mark Lutz, 2001). Python demonstrates most of these features common to lots of other languages, which makes it easier to learn every other language after learning python.

### 2.3 Machine Learning

Machine learning is the field of study that enables computers to learn and make decisions without being explicitly programmed (Arthur Samuel, 1959).It focuses on enabling machines to process and analyse data more efficiently. When data is too complex or extensive for manual interpretation, machine learning becomes particularly valuable. The goal of machine learning is to enable systems to learn from data and improve their performance over time.

Numerous studies have explored methods for machines to learn autonomously without direct programming. Machine learning relies on a range of algorithms to address different data problems. The choice of algorithm is influenced by the specific problem, the number of variables, and the most suitable model for the task. Machine Learning is widely being used in approximately every sector, including healthcare, marketing, finance, infrastructure, automation, etc. There are some important real-world examples of machine learning, such as medical diagnosis, traffic forecasting, marketing strategies, speech recognition, virtual assistants, and image recognition.

Machine learning algorithms characteristically fall into one of three learning types:

### 2.3.1 Supervised Learning

Supervised learning is applicable when a machine has sample data, i.e., input as well as output data with correct labels. These labels help evaluate the accuracy of the model. This approach leverages past experiences and labelled examples to predict future events. Initially, it analyses a known training dataset and then develops a function to forecast output values. During this process, the model also identifies and corrects errors through algorithms. Common techniques include linear and logistic regression, decision trees, and support vector machines, all aimed at creating models that accurately predict outcomes based on given input data.

### 2.3.2 Semi-supervised Learning

Semi-supervised learning combines both labelled and unlabeled data. This approach is useful when labelling data is time-consuming or expensive. By using a small amount of labelled data alongside a larger amount of unlabeled data, the model can improve its performance. For example, a classifier might first be trained on labelled data, then use that knowledge to label the unlabeled data, retraining the model for better accuracy. This method can significantly enhance the model's ability to classify new data effectively.

### 2.3.3 Unsupervised Learning

Unsupervised learning does not use labelled and classified data. Instead, the algorithm looks for patterns and groupings in the data without knowing the "right answers." It tries to identify

natural relationships within the dataset, such as grouping similar items together. Although not commonly used, unsupervised learning helps in exploring the data and can draw inferences from datasets to describe hidden structures from unlabeled data. Common methods include clustering algorithms like k-means clustering, which sorts data based on shared characteristics, such as size or shape.

### 2.4 Confusion Matrix

When data is obtained from a model, it needs to be processed to see the performance of the model that was used. Confusion matrix acts as a performance measurement for machine learning. It is used to test recall, precision, accuracy, and F1-score of the model. Usually, a confusion matrix is seen in a table with 4 different combinations of predicted and actual values.

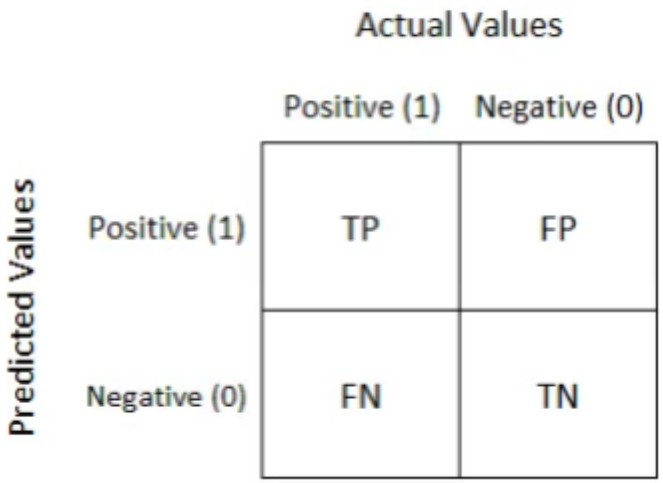


Figure 2.1 Confusion Matrix

TP can also be known as “true positive” when the prediction in positive is true. TN, also known as “true negative” is when the prediction is negative and it’s true. FP (“false positive”), is the first type of error when the positive is predicted but it is false. While FN (“false negative”) is the second type of error which is when the negative is predicted and it is false. Confusion matrix is then calculated for the accuracy, recall, precision, and F1-score of the model. The first formula to appear is $Accuracy = \frac{TP+TN}{Total}$ where this plays the part of measuring the performance of the model. After accuracy, there is the formula for $Recall = \frac{TP}{TP+FN}$ which basically means that from all the positive

classes, how many are predicted correctly. Then there is $Precision = \frac{TP}{TP+FP}$ where this measures the accuracy of a model's positive predictions. It can be said that from all the classes that were predicted positively, how many are actual. Finally there is the F1-score where it is used to evaluate the overall performance of the model. The formula is $F1 - Score = \frac{2 \times precision \times recall}{precision + recall}$ where it compares the precision and recall.

### 2.5 Indicators of Sleeping

The human body undergoes various changes during sleep, which is evident by several physical indicators. One significant indicator is body temperature, recognised as a "gold standard" marker of the human circadian rhythm and closely linked to sleep regulation (Togo et al., 2007). The sleep/wake cycle in humans is generally synchronised with the circadian rhythm of body temperature. This synchronisation occurs partly because the urge to sleep peaks during the decline in body temperature, and partly because alertness increases once body temperature reaches its lowest point (Borbély, 1982; Dijk and Czeisler, 1995).

Additionally, an experiment conducted by Kusunoki Mikie (1985) identified several sleep parameters that determine the comfort of human sleep. The study focused on ten sleep parameters derived from data on body movements and found that the frequency of body movements per hour (FB), mean rest period time (MR), and maximum rest period time (MAR) are the most effective indicators of sleep comfort. Therefore, the physical indicators of body movement frequency tend to decrease as individuals enter deeper stages of sleep.

# CHAPTER III
# METHODOLOGY

## 3.1 Time and Place

- Time: September - November
- Place: St. Laurensia Alam Sutera

## 3.2 Materials

- Laptop/Desktop Computer
- Webcam
- Python installed in computer
- OpenCV Library
- Dataset Samples

## 3.3 Flowchart

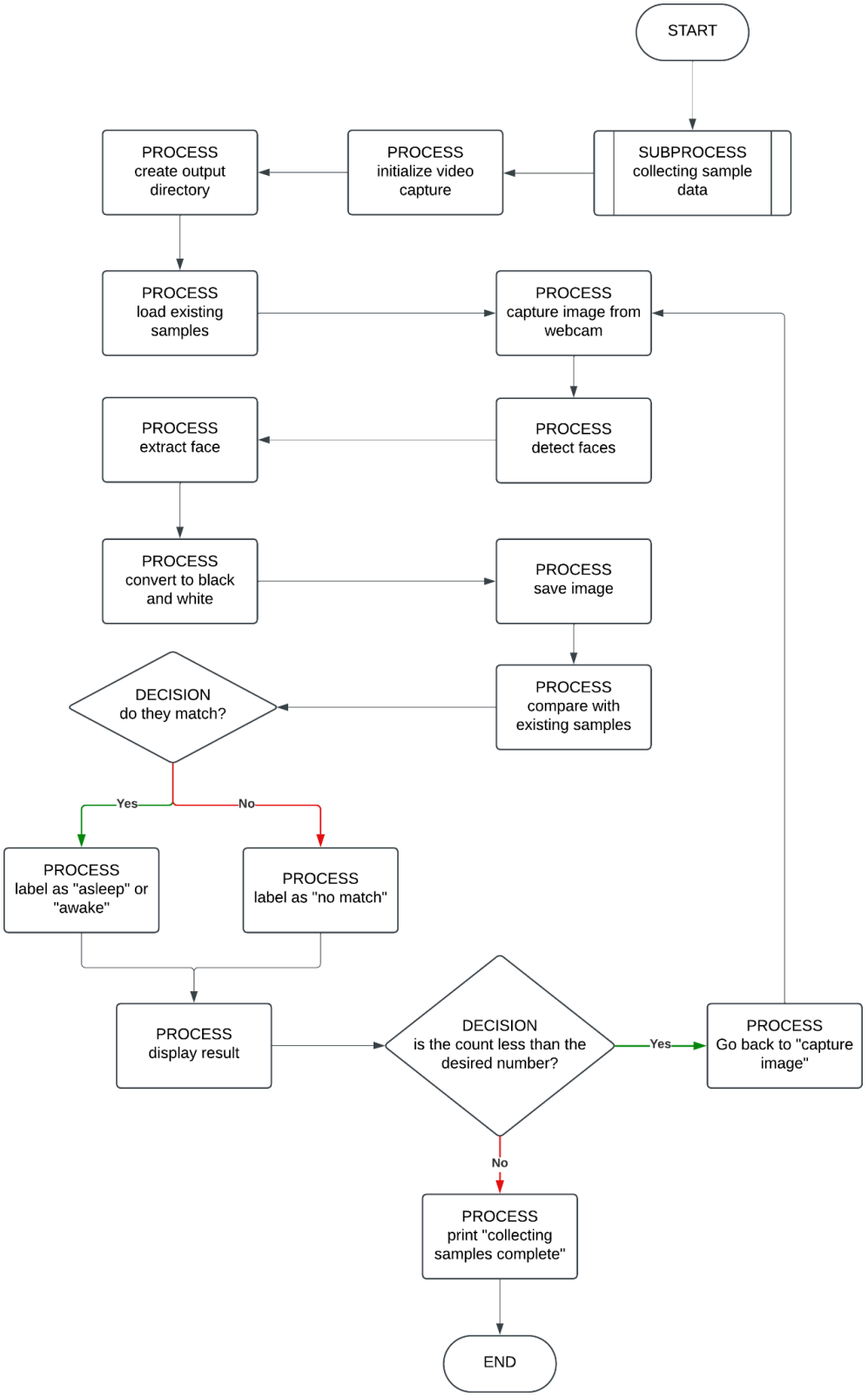


Figure 3.1 Program Flowchart

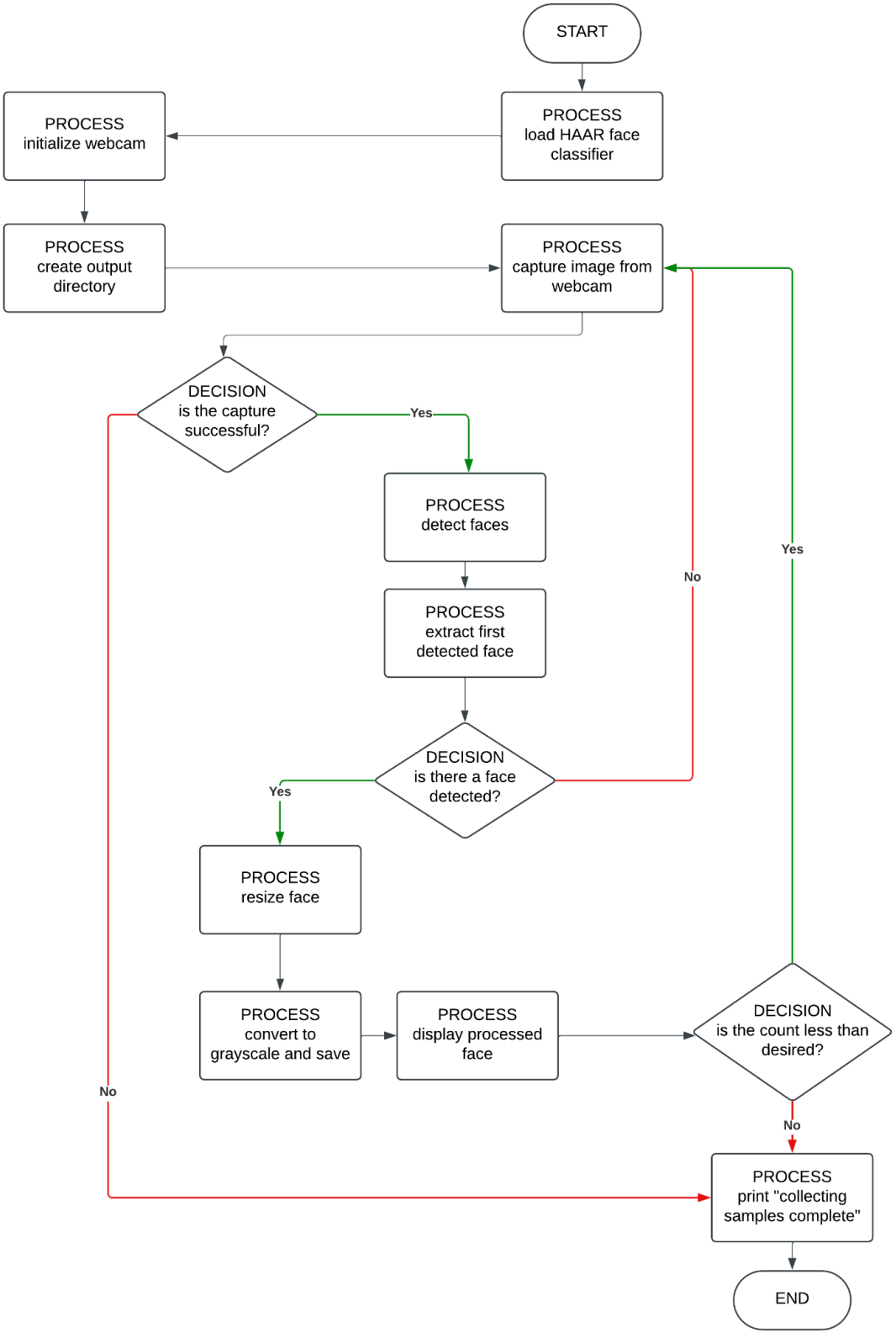


Figure 3.2 Subprocess Collecting Sample Data

### 3.4 Procedure

1. Create a program using Python and OpenCV to capture images (dataset) from a webcam.
2. Set the output directory according to the desired category (create subdirectories for "asleep" and "awake").
3. Define the desired count of images to capture for each category (e.g., 100 for "asleep" and 100 for "awake").
4. Execute the program to capture and save the images based on user input.
5. Create a new program using Python and OpenCV library to capture images from the webcam for prediction.
6. Preprocess the captured image to match the input requirements of your model.
7. Use the pre-trained model to predict the label of the captured image.
8. Compare the predicted label with the actual labels in the dataset.
9. Collect predicted and true labels for evaluation.
10. Use a confusion matrix to assess the model's performance.
11. Visualize the confusion matrix to understand the classification results.

# CHAPTER IV
# DATA AND ANALYSIS

## 4.1 Data

### 4.1.1 Programme Result

#### 4.1.1.1 Successful

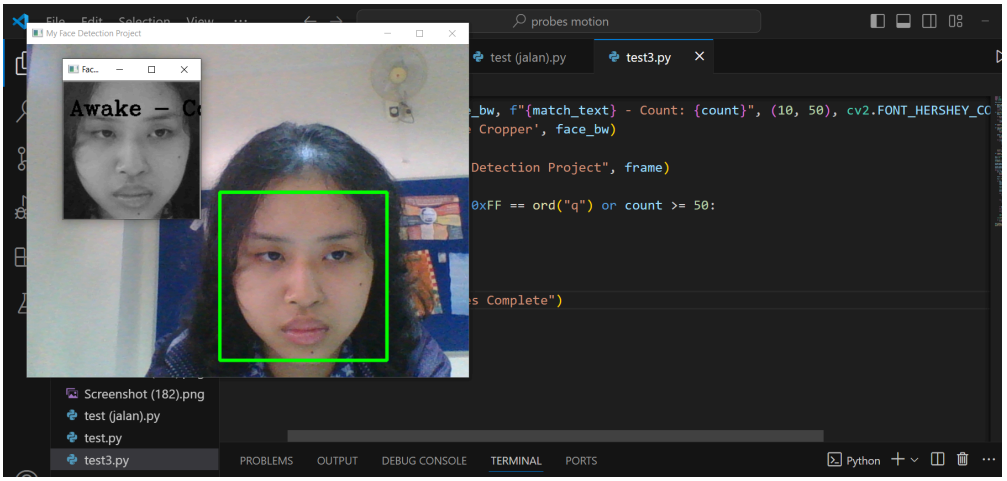


Figure 4.1 Successful Awake

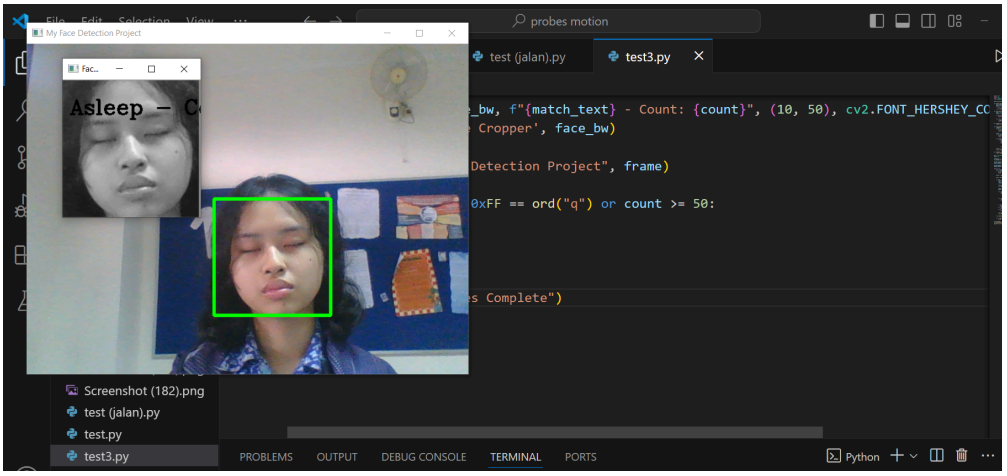


Figure 4.2 Successful Asleep

### 4.1.1.2 Error

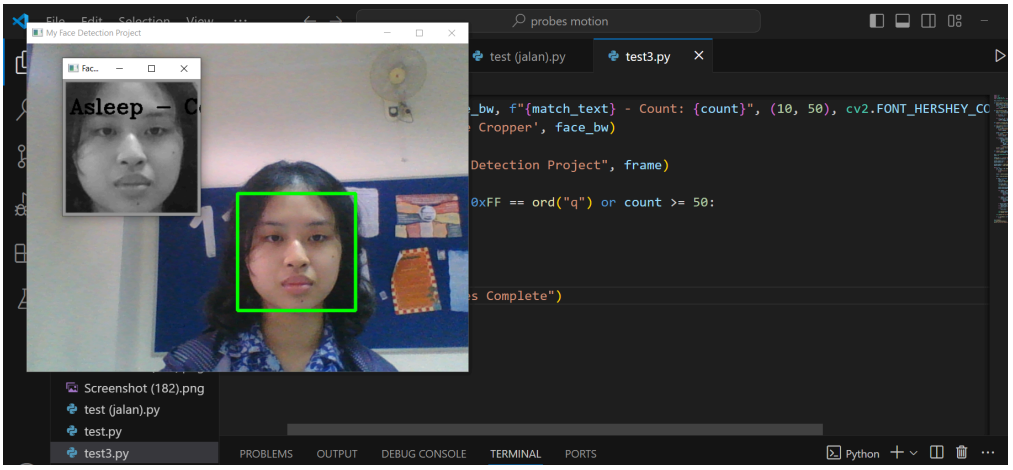


Figure 4.3 Error Detecting Sleep

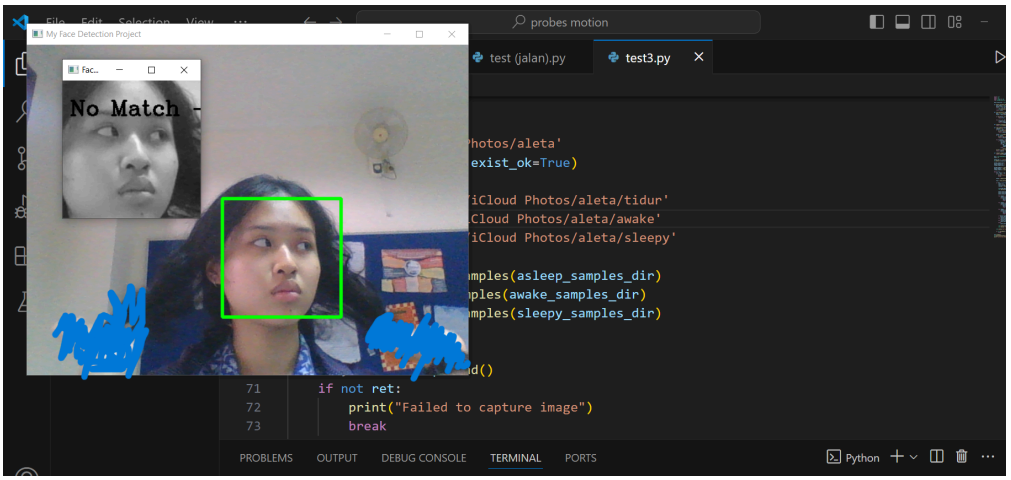


Figure 4.4 Failure Detecting Condition

### 4.1.2 Programme Confusion Matrix

#### 4.1.2.1 100 Samples

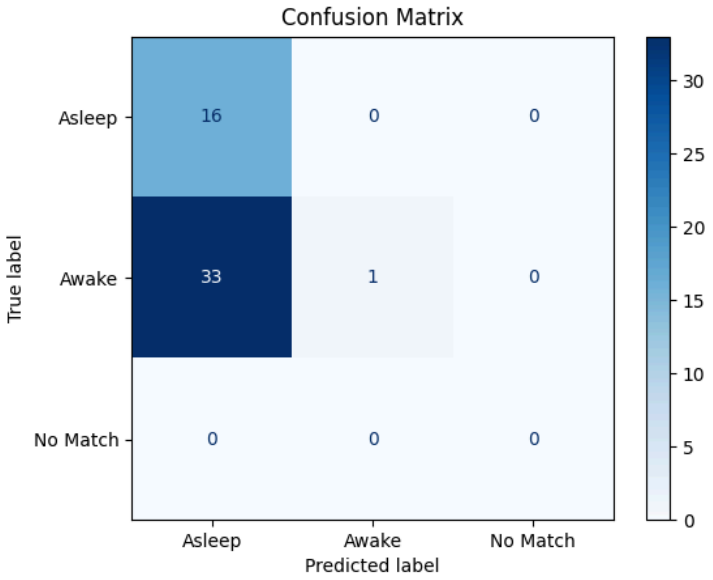


Figure 4.5 100 Sample Data Confusion Matrix

| | Precision | Recall | F1-Score | Accuracy |
|---|---|---|---|---|
| Asleep | 32.6% | 100% | 49.2% | 34% |
| Awake | 100% | 2.9% | 5.7% | |
| No Match | - | - | - | |

Table 4.1 100 Sample Data Confusion Matrix Table

#### 4.1.2.2 300 Samples

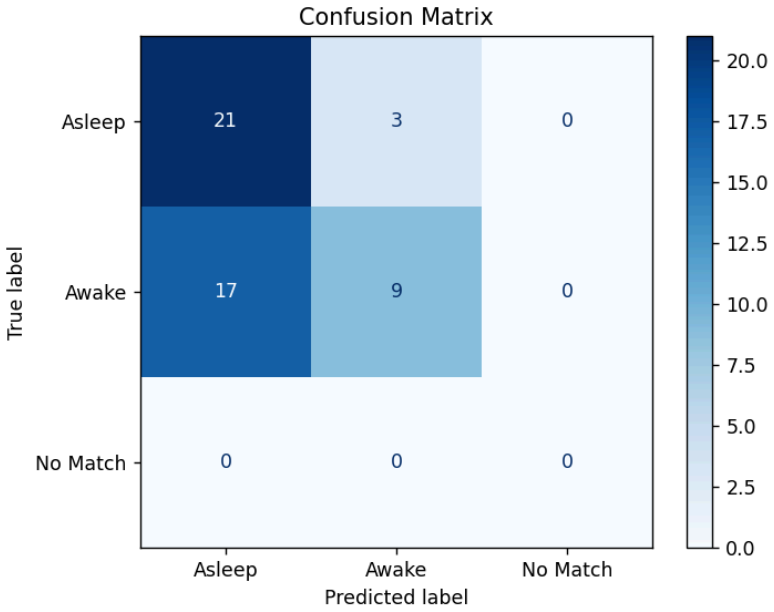


Figure 4.6 300 Sample Data Confusion Matrix

| | Precision | Recall | F1-Score | Accuracy |
|---|---|---|---|---|
| Asleep | 55% | 87.5% | 67.8% | 60% |
| Awake | 75% | 34.6% | 47.4% | |
| No Match | - | - | - | |

Table 4.2 300 Sample Data Confusion Matrix Table

### 4.1. 2.3 500 Samples

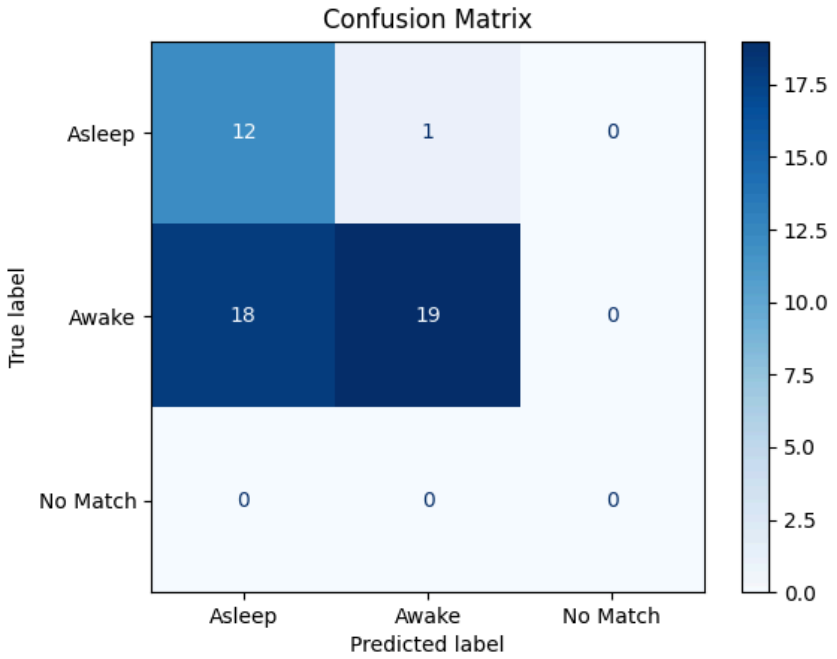


Figure 4.7 500 Sample Data Confusion Matrix

| | Precision | Recall | F1-Score | Accuracy |
|---|---|---|---|---|
| Asleep | 40% | 92.3% | 55.8% | 62% |
| Awake | 95% | 51.4% | 66.7% | |
| No Match | - | - | - | |

Table 4.3 500 Sample Data Confusion Matrix Table

## 4.2 Analysis

### 4.2.1 100 Samples

The confusion matrix from 100 samples shows a low accuracy of 34%, indicating poor overall performance. The model demonstrates perfect precision (100%) for the "awake" class, but this is likely due to very few predictions. The precision for the "asleep" class is low at 32.6%, indicating a high rate of false positives. The recall for "asleep" is perfect at 100%, showing the model identifies all true "asleep" instances. However, the recall for "awake" is extremely low at 2.9%, suggesting the model misses almost all actual "awake" cases. The F1-scores reflect these imbalances, with both classes showing poor overall performance. These results indicate a strong bias towards predicting "asleep" states, leading to significant misclassification errors.

### 4.2.2 300 Samples

The confusion matrix from 300 samples shows an accuracy of 60%, indicating a moderate overall performance and an improvement over the 100-sample test. The model demonstrates good precision (75%) for the "awake" class, meaning that most predictions for this class were correct. The precision for "asleep" is lower at 55%, indicating some false positives where the model predicts "asleep" for non-asleep cases. The recall for "asleep" is high at 87.5%, showing the model effectively identifies most true "asleep" instances. The recall for "awake" is lower at 34.6%, suggesting the model still misses many actual "awake" cases. The F1-scores reflect these imbalances, with the "asleep" class showing better overall performance.

### 4.2.3 500 Samples

The confusion matrix from 500 samples shows an accuracy of 62%, indicating a moderate overall performance. The model demonstrates high precision (95%) for the "awake" class, meaning that when it predicts "awake," it's usually correct. However, the precision for the "asleep" class is lower at 40%, indicating a tendency to overpredict this state. The recall for "asleep" is high at 92.3%, showing the model effectively identifies most true "asleep" cases. In contrast, the recall for "awake" is 51.4%, suggesting the model misses about half of the actual "awake" cases. The F1-scores reflect these imbalances, with the "awake" class showing better overall performance compared to previous 100 and 300 sample results.

### 4.2.4 Error Analysis

There are several needs for improvements to enhance the model's ability to accurately distinguish between sleep states:

- Factors such as lighting, distance, and movement may affect the ability of machine learning to detect sleep states through the subject's face.
- A higher number of sample counts is needed to increase the precision of the prediction.

# CHAPTER V
# CONCLUSION AND FUTURE WORK

## 5.1 Conclusion

Based on our hypothesis, it is true that the programme will function in detecting sleep, however the inaccuracy rate might be high. Comparing the data from 100, 300, and 500 samples, it can be concluded that a higher number of samples will increase the accuracy of detecting sleep behaviours. This is supported by the data of the model's overall accuracy, which improved from 34% with 100 samples to 62% with 500 samples. In addition, the analysis suggests that external factors such as lighting conditions, subject distance, and other supporting factors may significantly impact the model's ability to accurately detect sleep states through facial analysis. These findings highlight both the strengths and weaknesses of the current approach, pointing out areas that need improvement and further study.

## 5.2 Future Work

For future research some improvements can be made to optimize the research. These improvements include:

1. Use a laptop/device with a higher RAM capacity, which is at least 16GB
2. Expand the training for the machine learning
3. Increase the number of samples
4. Compress image files to optimize processing time

## REFERENCES


Rosebrock, A. (2018, July 19). *OpenCV tutorial: A guide to learn OpenCV.* PyImageSearch. https://pyimagesearch.com/2018/07/19/opencv-tutorial-a-guide-to-learn-opencv/

Zeynep, T. (2022, March 21). *The science of sleep: What makes us sleepy?* Vox. https://www.vox.com/even-better/23752433/sleepy-tired-fatigue-differences

Oxford Learning. (2023, March 17). *Should teachers let students sleep in class?* https://www.oxfordlearning.com/should-teachers-let-students-sleep-in-class/

Shore, L. (n.d.). *The importance of sleep in the classroom*. Education World. https://www.educationworld.com/a_curr/shore/shore005.shtml

Start School Later. (n.d.). *Wake-up calls: Fast facts about school start times.* https://www.startschoollater.net/wake-up-calls-fast-facts.html

Brue, A. (2014, March 21). *Study: Homework stress can keep kids from sleeping.* CNN. https://edition.cnn.com/2014/03/21/health/homework-stress/

Centers for Disease Control and Prevention. (n.d.). *Sleep in middle and high school students.* https://www.cdc.gov/healthyschools/sleep.htm

American Physical Therapy Association. (2018, January 29). *CDC: Most middle and high school students don't get enough sleep.* https://www.apta.org/news/2018/01/29/cdc-most-middle-and-high-school-students-dont-get-enough-sleep

B-Sync. (2022). *2022 student sleep statistics: An endemic problem.* https://b-sync.life/blogs/science/2022-student-sleep-statistics-an-endemic-problem

West Lion's Roar. (2021, February 25). *The effects homework can have on teens' sleeping habits.* https://www.westlionsroar.com/features/2021/02/25/the-effects-homework-can-have-on-teens-sleeping-habits/

Canapari, C. (2015, August 31). *Homework vs. sleep: A cause of stress in teens?* Dr. Craig Canapari. https://drcraigcanapari.com/homework-vs-sleep-a-cause-of-stress-in-teens/

OpenCV. (n.d.). *About OpenCV*. https://opencv.org/about/

TutorialsPoint. (n.d.). *OpenCV - Overview*. TutorialsPoint. https://www.tutorialspoint.com/opencv/opencv_overview.htm

Python Software Foundation. (n.d.). *Python: The programming language.* https://www.python.org/doc/essays/blurb/

Awan, M. I., & Anwar, M. (2021). *A study of the impact of online learning on students' performance.* International Journal of Research in Education and Science, 7(1), 395-368. https://journal.ijresm.com/index.php/ijresm/article/view/395/368

Lutz, M. (2001). *Programming python*. " O'Reilly Media, Inc.". https://search.worldcat.org/title/1049734214

Python, W. (2021). Python. *Python releases for windows*, *24*. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=1f2ee3831eebfc97bfafd514ca2abb7e2c5c86bb

Packt Publishing. (2013). *Deep learning: Building machine learning systems with Python.* http://archive.keyllo.com/L-%E7%BC%96%E7%A8%8B/DeepLearning-Building%20Machine%20Learning%20Systems%20with%20Python-Packt%20Publishing%20(2013).pdf

Mahesh, B. (2020). Machine learning algorithms-a review. International Journal of Science and Research (IJSR).[Internet], 9(1), 381-386. https://www.researchgate.net/profile/Batta-Mahesh/publication/344717762_Machine_Learning_Algorithms_-A_Review/links/5f8b2365299bf1b53e2d243a/Machine-Learning-Algorithms-A-Review.pdf?eid=5082902844932096t

Maulana, F. (2021). Machine Learning Object Detection Tanaman Obat Secara Real-Time Menggunakan Metode Yolo (You Only Look Once). https://epub.imandiri.id/repository/docs/journal/Jurnal_361601016.pdf

Bi, Q., Goodman, K. E., Kaminsky, J., & Lessler, J. (2019). What is machine learning? A primer for the epidemiologist. *American journal of epidemiology*, *188*(12), 2222-2239. https://sph.umsha.ac.ir/uploads/18/2023/May/22/kwz189.pdf

Aoyagi, Y., Park, S., Cho, S., & Shephard, R. J. (2018). Objectively measured habitual physical activity and sleep-related phenomena in 1,645 people aged 1–91 years: The Nakanojo Community Study. *Preventive Medicine Reports*, 11, 180-186. https://doi.org/10.1016/j.pmedr.2018.06.013

Kusunoki, M. (1985). Body movements during sleep as an indicator of comfort. *Nippon Eiseigaku Zasshi (Japanese Journal of Hygiene), 39*(6), 886-893. https://www.jstage.jst.go.jp/article/jjh1946/39/6/39_6_886/_pdf/-char/ja

Toyong, P. J. A. (2020). Sleeping habits, classroom behaviour and academic performance of senior high school students. *International Journal of Multidisciplinary Applied Business and Education Research*, *1*(1), 54–63. https://www.neliti.com/publications/581441/sleeping-habits-classroom-behaviour-and-academic-performance-of-senior-high-scho

Narkhede, S. (2021, June 15). Understanding Confusion Matrix - towards Data science. *Medium*. https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62

GeeksforGeeks. (2024, July 8). *Understanding the confusion matrix in machine learning*. GeeksforGeeks. https://www.geeksforgeeks.org/confusion-matrix-machine-learning/