Abstract

Operating a motor vehicle when sleepy is known as drowsy driving, and it can influence any individual who gets in the driver's seat. Sleepy driving altogether expands the danger of mishaps, prompting an alarming number of wounds and deaths consistently.

Sleepy driving fundamentally expands the danger of fender benders. Microsleeps are the point at which an individual naps off for only a couple of moments, and when they happen while driving, it's simple for the vehicle to run off the street or crash into another vehicle. The harm from these accidents increments when they happen at high velocities.

We intend to build a system that helps alert a drowsy driver to make the roads more safe by detecting and analysing eye patterns of the driver. By leveraging the power of computer vision and machine learning a quick affordable and safe mechanism can be integrated to make driving more safe.

This system can be integrated into hybrid automated cars that will serve as the future of automobiles.

Introduction

Drowsy Driving

Drowsy driving ( drained driving, sluggish driving, or exhausted driving) is the activity of an engine vehicle while being intellectually disabled by an absence of rest. Lack of sleep is a significant reason for engine vehicle mishaps, and it can weaken the human brain however much intoxication can.

Drowsy driving is risky regardless of whether an individual really nods off. Research shows that lack of sleep prompts mental disability that is like drunkenness with 24 hours of lack of sleep generally likening to a blood liquor content (BAC) of 0.10%.

This hindrance makes an individual less mindful of their environmental elements and all the more quickly drawn offtrack. It eases back their response time, making it harder to stay away from risks in the street. Deficient rest is additionally attached to demolished navigation, which can prompt danger taking in the driver's seat.

At the point when an individual doesn't get a sufficient measure of rest, their capacity to work is impacted. Their coordination is disabled, has a longer response time, weakens judgment, and memory is debilitated.



Numerous variables can assume a part in tired driving:

* Lack of sleep: Lack of rest is a main source of exorbitant daytime tiredness, which can actuate microsleeps or other risky driving conduct. Grown-ups ought to get somewhere in the range of seven and nine hours of rest every evening, except countless grown-ups regularly neglect to get this suggested measure of rest.
* Rest problems: Many rest issues, like obstructive rest apnea, cause an individual's rest to be limited, intruded, and less helpful. Many rest problems go undiscovered and, when left untreated, can cause daytime tiredness.
* Liquor: Drinking liquor can incite sluggishness while additionally influencing response time and decision-production in manners that increment the dangers of car collisions.
* Prescriptions: Numerous drugs cause drowsiness. Tranquilizers, including professionally prescribed medications, over-the-counter prescriptions, and dietary enhancements, that are taken around evening time might cause waiting tiredness the following morning. Tiredness is likewise a symptom of meds utilized for some different conditions.
* Season of day: Auto collisions from lazy driving happen most as often as possible between 12 PM and six a.m. or on the other hand in the mid-evening, which are twice when languor tops.
* Sluggish driving can influence any individual who jumps in the driver's seat, yet certain individuals are at higher danger of vehicle crashes identified with tired driving, including:
* Individuals who drive professionally, for example, long stretch drivers or transport drivers.
* Individuals who work extended periods, sporadic movements, or night shifts.
* Individuals with genuine rest issues including a sleeping disorder or other rest problems.
* Teens who have less driving experience and high paces of rest deficiency.

Objectives

* Affordable and easy to implement system that alerts a drowsy driver by creating a beep sound if the driver’s eyes are closed for more than a few seconds.
* This system leverages the power of computer vision to classify if the person is drowsy or not. Classic ML approach will be used to first train a image model (MobileNet)
* The model will be trained on images of various individual subjects which can be found publicly in MRL repository
* Transfer learning will be used to make a system more accurate and train the model from a pre existing checkpoint rather than starting from scratch

Literature Survey

Computer Vision

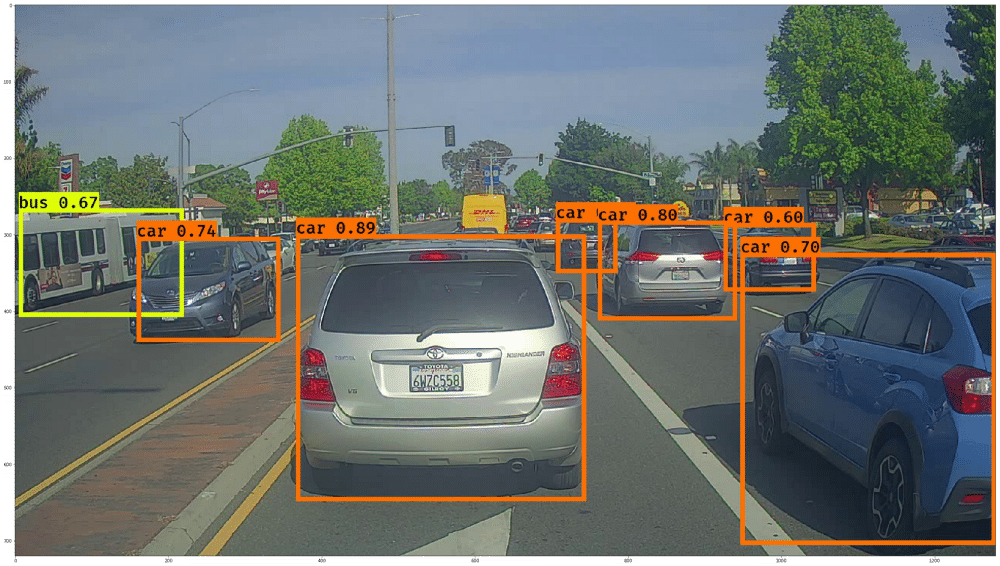
Computer vision (CV) is the subcategory of Artificial Intelligence (AI) that spotlights on building and utilizing advanced frameworks to process, examine and decipher visual information. The objective of Computer vision is to empower processing gadgets to effectively distinguish an article or individual in a computerized picture and make a fitting move.

Computer vision utilizes convolutional neural networks(CNNs) to process visual information at the pixel level and profound learning repetitive neural networks (RNNs) to see how one pixel identifies with another.

Pictures on Computers are regularly put away as large matrices of pixels. Every pixel is characterized as shading, put away as a blend of 3 added substance essential tones: RGB(Red Green Blue). These are consolidated in changing powers to address various shadings. Shadings are put away inside pixels.

How about we consider a straightforward calculation to follow a dazzling orange football on a football field. For that, we'll take the RGB worth of the centermost pixel. With that worth saved, we can give a Computer program a picture, and request that it observe the pixel with the nearest shading match. The calculation will actually take a look at every pixel at a time, calculating the distinction from the objective tone. Having checked out each pixel, the best match is a logical pixel from the orange ball. We can run this calculation for each casing of the video and track the ball over the long haul. However, if one of the groups is wearing an orange pullover, the calculation may get confused. Thus, this methodology doesn't work for highlights bigger than a solitary pixel, like edges of articles, which are composed of numerous pixels.

To distinguish these elements in pictures, Computer vision calculations need to think about little areas of pixels, called patches. For instance, a calculation that tracks down vertical edges in a scene, to assist a robot with exploring a field of deterrents. For this activity, a numerical documentation is utilized, which is known as a part or channel. It contains the qualities for a pixel-wise augmentation, the amount of which is saved into the middle pixel.



## Convolutional Neural Networks (CNNs)

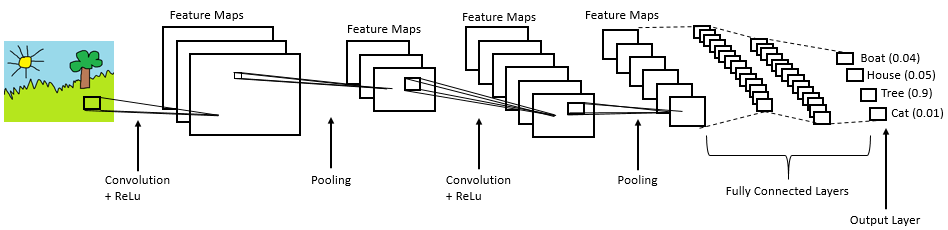
An artificial neuron is the basic structure of a neural network. It takes a progression of data sources, and increases each by a predefined weight, then, at that point, aggregates those qualities by multiplying them. The information parameters are comparable to kernel values, yet not at all like a predefined kernel value, neural networks can become familiar with their own features that can perceive intriguing highlights with regards to pictures.

CNNs use banks of these neurons to handle picture information, each yielding another picture, basically processed by various learned parts. These yields are then handled by resulting layers of neurons, considering rehashed convolutions.

The first convolutional layer may observe things like edges, the following layer may convolve on those edge elements to perceive straightforward shapes, containing edges, similar to corners. The following layer may convolve on those corner includes, and contain neurons that can perceive straightforward items, similar to mouths and eyebrows. This continues rehashing, developing in intricacy, until there's a layer that does a convolution that perceives all elements: eyes, ears, mouths, nose, and so forth, and says "it's a face".

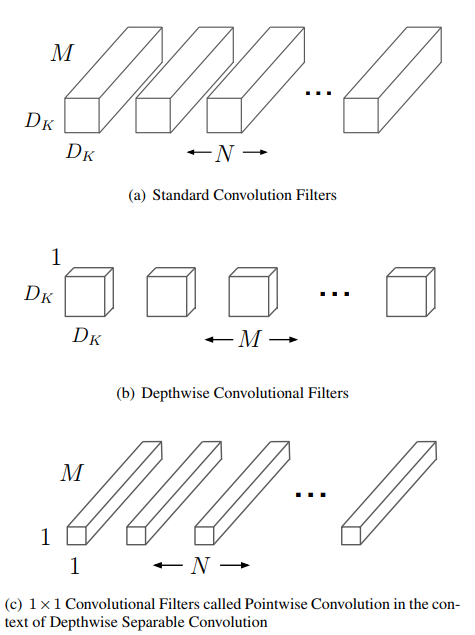
Whenever we've detached a face, we can apply more specific Computer vision calculations to pinpoint facial milestones, similar to the tip of the nose and the edges of the mouth. This can be utilized to decide whether the eyes are open, the situation of the eyebrows, which is simple once you have the milestones. The general distance among eyes and eyebrows can measure up to uncover shock or pleasure.

Also, all of this data can be deciphered by emotion recognition algorithms, enabling Computers to derive fundamental 'mind-sets'. This can assist Computers with becoming human, i.e., mindful of their environmental elements



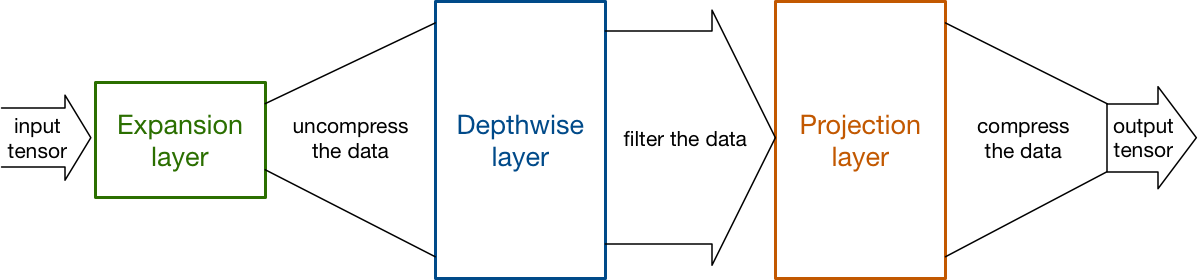
## Mobile net

Mobile net is a streamlined neural network architecture that is used in embedded vision based applications. Due to its depth wise separable convolutions, its lightweight deep neural network is suitable for mobile applications and embedded applications that have limited computing power.



It utilizes depthwise separable convolutions which fundamentally implies it plays out a solitary convolution on each color channel rather than joining every one of the three colors and smoothing it. This has the effect of filtering the input channels.

It is exceptionally low maintenance and performs very well with high speed. There are many kinds of pre-prepared models with the size of the network in memory. The speed and power utilization of the network is corresponding to the quantity of MACs (Multiply-Accumulates) which is a proportion of the quantity of combined Multiplication and Addition activities.



## Problem Statement

As discussed earlier, drowsiness is a dangerous state for vehicle drivers to drive. This is the origin of the problem statement. We intend to put forward a system that can detect the state of drowsiness in vehicle drivers and activate an alert mechanism if drowsiness is on unsafe levels. By doing so, the roads will be safer to drive and a lot of accidents can be avoid

## Solution

We leverage the power of Machine learning and computer vision to detect the eye state of the driver. The biggest symptom of drowsiness is heavy eyes and periodic microsleep. Ie. sleeping for a few seconds. This can be detected by a trained image classifier that first detects eyes in the picture and then detects the state of the eyes. Ie. if it's opened or closed. If eyes are closed for more than specified time, a bepper can play, alerting the driver. The advantages of such a system will be tremendous.

Eg

* Fewer accidents
* Better quality driving

## Methodology

We will be using transfer learning as the main paradigm for this project. The basic structure of the project is:

1. Data pre processing
   1. First we convert the image to numpy arrays
   2. These arrays are labelled into closed and open
   3. Then we split the array into labels and features
2. Load Mobile net model for training,
3. Alter the structure of the model,
4. Add a flatten and dense layer with one node
   1. One node will signify 0 if the model has predicted closed eyes or 1 if the model has predicted open eyes.
5. Train the model
   1. Model is trained for 5 epochs, yields an accuracy of 98-99%
   2. Model is trained on GPU, using loss function as binary\_crossentropy and adam optimizer
6. Test the model
   1. Phase 1 Testing is done on seen images ie the images that the model are trained on
   2. Phase 2 Testing is done on completely new image that the model has not been trained on
7. Load CV2’s haar facial detection for getting cropped eyes image from any picture
8. CV2 also provides ways to render live video from webcam
9. The live footage can be broken down, frame by frame. These frames can be feeded to the trained model to detect closed eyes
10. If close eyes detected for 5 frames, a beep noise will get generated

## How to run the code?

The code is written in python with a jupyter notebook as IDE. Jupyter notebooks are text editors that are integrated with a python interpreter. This makes it easy for documentation and running code simultaneously.

To run the code, we should first install the dependencies. We will be using a virtual environment for the dependencies as it is better to run this code in isolated mode. Anaconda is a famous virtual environment tool in python that we will use to install and modify these dependencies.

To replicate the environment we use,

| conda create env |
| --- |

This command reads the environment file and installs all the required dependencies.

Next step is to run the jupyter notebook, to do this, we run the following command,

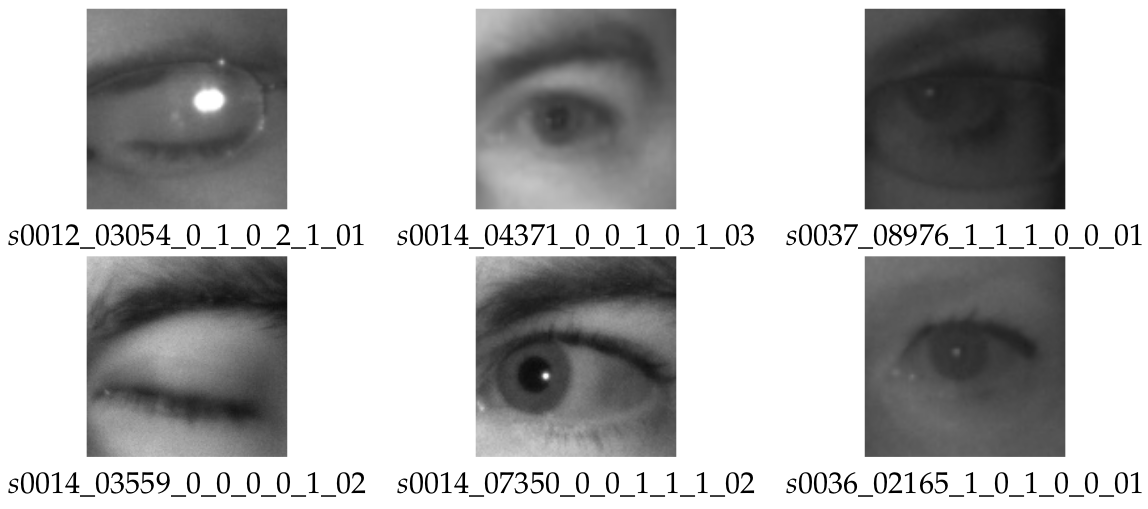
| jupyter notebook |
| --- |

This command will run the notebook and open the IDE in the browser. We are now ready to run the code.

The actual code is divided into 2 jupyter notebooks. The first notebook is called Seg.ipynb. The purpose of this book is to segregate the dataset into “open” eyes and “closed” eyes.

## The Dataset

The MRL eye dataset is a subset of a bigger dataset. It consists of eye pictures of 37 different people photographed in different conditions. This dataset is publically available and one of the most used dataset in the field of CV.



These are a few examples of the dataset. Each image is annotated as mentioned below,

* subject ID; in the dataset, we collected the data of 37 different persons (33 men and 4 women)
* image ID; the dataset consists of 84,898 images
* gender [0 - man, 1 - woman]; the dataset contains the information about gender for each image (man, woman)
* glasses [0 - no, 1 - yes]; the information if the eye image contains glasses is also provided for each image (with and without the glasses)
* eye state [0 - closed, 1 - open]; this property contains the information about two eye states (open, close)
* reflections [0 - none, 1 - small, 2 - big]; we annotated three reflection states based on the size of reflections (none, small, and big reflections)
* lighting conditions [0 - bad, 1 - good]; each image has two states (bad, good) based on the amount of light during capturing the videos
* sensor ID [01 - RealSense, 02 - IDS, 03 - Aptina]; at this moment, the dataset contains the images captured by three different sensors (Intel RealSense RS 300 sensor with 640 x 480 resolution, IDS Imaging sensor with 1280 x 1024 resolution, and Aptina sensor with 752 x 480 resolution)

## Target Deliverables

The goal of this project is to make driving more safe and reduce the number of accidents that take place in day to day life. Since a large number of vehicle accidents happen due to drowsy driving, this system will help in keeping that number low.

The scope of this project is limitless. Various other systems can be implemented to make a solid error-free and safe driving assistance application that will make driving even more pleasant and advanced. The model can be trained on better data making it adaptable to various lighting conditions and angles.

## Requirement Analysis

HARDWARE REQUIREMENTS

Processor : Any Processor above 2.3 GHz Base clock ideally + Graphics Card for training

RAM : 16GB

Hard Disk : 10 GB

Input device : Standard Keyboard and Mouse

Output device : HDMI Monitor

SOFTWARE REQUIREMENTS

Operating system : Windows

Text Editor : Jupyter notebook

Compiler : Python 3.6

## Result

(Paste your result screen shot here, from both jupyter notebook and webcam)

We have effectively constructed and shown the fundamental working of the framework. Further the framework can be upgraded via preparing the model on more information or by utilizing an all the more remarkable neural network architecture.

## Conclusion

In conclusion we would like to bring about the work that has gone in the project which can be used for further research and development. We have learnt a great deal about the concepts of machine learning and computer vision. We wish to continue learning more about this topic and make better systems that make a difference to society.

The project has also taught us how to work in a team and solve tasking problems together. This worked are interpersonal communication skills and made us stronger and more refined to the real world industry.