behavior model. Changes to these characteristics are used to monitor driver fatigue. Again the accuracy is high (98%) on the proposed datasets, but no discussion about hardware is made.

In this paper, we decided to utilize neural networks to detect drowsiness since image processing techniques are very heavy to implement on embedded processors. Also, the use of biomedical techniques can result in an accurate system in some situations, but not affordable.

As a hardware-software codesign matter, the selected method must run on suitable hardware. In recent years RiscV has gained popularity due to its open-source nature. For deep learning algorithms, however, the majority of research indicates that the designers paired a base CPU with an accelerator to decrease the latency of the implemented neural network. In [18], the authors proposed a RiscV-based hardware accelerator designed for the Yolo object detection system. The designed system executed Yolo in 400ms. In [19], the authors compared two hardware accelerators, NVDLA and Gemmini, and they proved that NVIDIA's NVDLA accelerator outperforms Gemmini by 3.77x running ResNet-50 on an equivalent configuration using the same system setup. However, in [20], the author mentioned that a high-end FPGA is required to test the NVIDIA's NVDLA accelerator. As a result, they proposed to integrate NVDLA into a real RiscV SoC on the Amazon cloud FPGA using a tool named FireSim. As a result, although adding an accelerator to a base CPU is logical, but the added area

requirements is not supplied. Subsequently, in this paper, frames of the videos from reference [22] are extracted to design an accurate system. The frames of these videos are extracted at the rate of 1 frame per second, then the frames are labeled in four distinct classes, including normal, distraction, yawn, and sleep. After the initial design with the extracted dataset, several images were added to the dataset to increase the system's accuracy. In this paper, we consider a scenario where the camera is fixed on the car's mirror. Fig. 1 (a) to (d), illustrate a sample of distinct classes, including normal, distraction, sleep, and yawn, respectively.

For these classes, intending to improve the system accuracy and sensitivity, extracted frames are augmented by adding Gaussian noise to the images, changing the brightness of images, translation, and rotation, as can be seen in Fig. 1 (e) to (h). It must be noted that the system quality and received images might be affected due to driving cars in tunnels, bumpy roads, and the like. Consequently, we consider Gaussian noise with a wide range of $\sigma$, including 0.01, 0.02, and 0.04. The reason for considering noises is due to the various quality of the camera and changing the quality of input images. Furthermore, while cars are derived at different times, the light varies considerably. For this reason, to improve the system sensitivity, the effect of the light is considered. It must be noted that, due to the mobility of cars, moving in various directions, and the hills and valleys on the roads, the driver's location, and position extracted by the inputs of the camera witness conspicuous changes.