Fusing ConvNeXt and Histogram of Oriented Gradients (CNHOG) for Modular and Efficient Driver Distraction Detection

Lakhani, Samay (School: Jericho High School)

In 2021, the National Highway Traffic and Safety Administration recorded 43,000 fatalities from automotive crashes, an 11% increase from 2020. Automotive accidents (94%) are caused by human driver error and are preventable. To reduce fatalities, a race to design a highly accurate and efficient distraction detection system is underway. Current distraction detection models rely on combined handcrafted and deep learning models but use outdated feature extractors such as AlexNet and utilize compute-heavy deep learning ensembles. Herein, I created a modernized hybrid architecture named CNHOG (ConvNeXt Histogram of Oriented Gradients). Fusing target localization from handcrafted features with adaptive generalizability from modern deep learning architectures enables accuracy and scalability to differentiate types of distracted driving like texting, calling, adjusting the radio or makeup, drinking, and reaching into the backseat. Using the discovered hyperparameters from a sweep, two variants (CNHOG-Tiny and CNHOG-Large) of the model were trained and compared to previous literature with the goal of deployment in motor vehicles. CNHOG-Tiny achieved 98.8% test accuracy and CNHOG-Large performed at 99.14% test accuracy, outperforming the previous state-of-the-art video model (97.5%) and image model (97.2%), indicating new state-of-the-art accuracies for the Driver Monitoring Dataset. CNHOG-Large model is 65.6% less erroneous than the previous state-of-the-art video model. CNHOG is scalable to mobile devices and is 12x more efficient than previous works (4.5 GFLOPs). Future regulations will mandate driver monitoring systems in all new cars by 2026. With modularity, state-of-the-art accuracy, privacy, and high-efficiency, CNHOG could revolutionize driver safety.