Deep Learning Based Driving Behavior Analysis

Author: **Hrishikesh Balakrishnan**

Abstract

Driving behavior is a complex and multifaceted area of study that encompasses a wide range of factors, including individual characteristics, environmental conditions, and social influences. In recent years, there has been a growing interest in understanding driving behavior to improve road safety, reduce traffic congestion, and promote sustainable transportation practices. This research paper aims to investigate the classification and prediction of driving behavior, specifically into three categories: slow, aggressive, and fast. This is accomplished by using various deep learning sequential models, and the results are compared to determine which model is most effective for analysis.

Introduction

Deep Learning Based Driving Behavior Analysis is an emerging field that leverages advanced machine learning techniques to study, analyze, and predict driver behavior. It involves the use of deep neural networks to extract complex patterns and relationships from large datasets of driver behavior, including vehicle speed, acceleration, braking, lane changing, and other related features.

The goal of this field is to improve driver safety and optimize transportation systems by identifying dangerous driving behavior, such as distracted driving, reckless driving, and fatigue. By leveraging deep learning techniques, researchers can analyze driving behavior data in real-time, providing valuable insights that can help reduce the number of accidents on the road.

Applications of deep learning-based driving behavior analysis include the development of driver assistance systems, the optimization of traffic flow, and the creation of personalized driver coaching programs. By accurately predicting driver behavior and providing actionable feedback, these systems can help improve the safety and efficiency of transportation systems while also reducing environmental impact.

Overall, deep learning-based driving behavior analysis is a promising field that has the potential to revolutionize the way we think about transportation and driver safety. With ongoing advancements in deep learning techniques and the increasing availability of driving behavior data, this field is poised for continued growth and innovation in the years ahead.

Related Work/Literature Review

There has been significant research conducted in the field of deep learning-based driving behavior analysis. In this literature review, we will explore some of the key findings from recent studies and examine the various approaches used to analyze and predict driver behavior.

One study conducted by Wang et al. (2020) proposed a deep learning-based method for detecting driver drowsiness. The authors used a combination of convolutional neural networks (CNNs) and long short-term memory (LSTM) networks to analyze various physiological signals, including electroencephalogram (EEG), electrocardiogram (ECG), and electrooculogram (EOG). The results showed that the proposed method achieved high accuracy in detecting driver drowsiness, demonstrating the potential of deep learning for improving driver safety.

Another study by Lee et al. (2019) proposed a deep learning-based approach for predicting driver aggression. The authors used a combination of LSTM and attention-based neural networks to analyze various driving-related features, including vehicle speed, acceleration, braking, and lane changing. The results showed that the proposed approach achieved high accuracy in predicting driver aggression, highlighting the potential of deep learning for improving driver behavior.

In addition to these studies, there have been several other approaches proposed for deep learning-based driving behavior analysis. These include the use of generative adversarial networks (GANs) for generating synthetic driving data (Wang et al., 2019), the use of variational autoencoders (VAEs) for anomaly detection in driving behavior (Zhang et al., 2021), and the use of graph convolutional neural networks (GCNs) for analyzing spatial relationships between vehicles on the road (Wang et al., 2021).

Zhang et al. (2020) proposed a deep learning-based approach for detecting driver distraction using visual cues. The authors used a combination of CNNs and LSTM networks to analyze video footage of drivers and detect various forms of distraction, such as cell phone use and eating. The results showed that the proposed approach achieved high accuracy in detecting driver distraction, highlighting the potential of deep learning for improving driver safety.

Li et al. (2020) proposed a deep learning-based approach for predicting driver stress levels. The authors used a combination of CNNs and LSTM networks to analyze various physiological signals, including heart rate variability and skin conductance, and predict driver stress levels. The results showed that the proposed approach achieved high accuracy in predicting driver stress levels, suggesting that deep learning can be used to improve driver well-being and comfort.

Sun et al. (2021) proposed a deep learning-based approach for predicting driver aggression using multimodal data, including driving behavior data and physiological signals. The authors used a combination of CNNs, LSTM networks, and attention-based neural networks to analyze the data and predict driver aggression. The results showed that the proposed approach achieved high accuracy in predicting driver aggression, highlighting the potential of deep learning for improving driver behavior.

Zhao et al. (2020) proposed a deep learning-based approach for predicting vehicle collisions using driving behavior data. The authors used a combination of LSTM networks and graph convolutional neural networks (GCNs) to analyze driving behavior data and predict the likelihood of collisions. The results showed that the proposed approach achieved high accuracy in predicting vehicle collisions, suggesting that deep learning can be used to improve road safety.

Overall, these studies demonstrate the potential of deep learning for improving driver safety and behavior. However, there are still many challenges that need to be addressed, including the need for large amounts of high-quality driving data, the need for more accurate and robust models, and the need for more comprehensive evaluations and comparisons of different approaches.

Dataset & Preprocessing

Aggressive driving involves actions such as speeding, sudden braking, and abrupt turns, all of which are reflected in the data collected from the accelerometer and gyroscope sensors. With the widespread availability of smartphones equipped with these sensors, the data used in this analysis was collected through an Android application. The training dataset contains 3644 recordings, with 1331 classified as slow driving, 1113 as aggressive driving, and 1200 as normal driving. The test dataset, which is slightly smaller, contains 3084 instances, with 1273 for slow driving, 997 for normal driving, and 814 for aggressive driving. These proportions reflect those of the training dataset but with lower numbers of instances. The attributes of the dataset are provided below.

- Sampling Rate: 2 samples (rows) per second.
- Gravitational acceleration: removed.
- Sensors: Accelerometer and Gyroscope.
- Data:
- Acceleration (X, Y, Z axis in meters per second squared (m/s2))
- Rotation (X, Y, Z axis in degrees per second (°/s))
- Classification label (SLOW, NORMAL, AGGRESSIVE)
- Timestamp (time in seconds)
- Driving Behaviors:
  - ❖ Slow
  - ❖ Normal
  - ❖ Aggressive
- Device: Samsung Galaxy S21

The only preprocessing done to both test and train data is to remove the first row as it contains zero values in acceleration and gyroscope as it is the inception of data collection from the device.

Deep Learning

Deep learning is a subfield of machine learning that involves the use of artificial neural networks with multiple layers to model and solve complex problems. These networks are inspired by the structure and function of the human brain and are capable of learning from large amounts of data to make predictions or decisions about new data. Deep learning models are trained using algorithms such as backpropagation, which iteratively adjust the weights and biases of the network to minimize the difference between the predicted outputs and the actual outputs for a given set of input data.

An LSTM classifier is a type of deep learning model that uses Long Short-Term Memory (LSTM) units in the network architecture for sequence classification tasks. LSTM units are a type of recurrent neural network (RNN) that can remember information for longer periods of time, making them particularly suited for handling sequences of data with long-term dependencies. During training, the LSTM classifier is optimized to minimize the difference between the predicted class probabilities and the true class labels using a loss function such as cross-entropy.

A ConvLSTM is a type of deep learning model that combines Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. It is used for sequence prediction tasks, where both spatial and temporal patterns need to be learned from input data. In the case of 1D data, such as a time series, a ConvLSTM can be used to model the temporal dependencies and patterns in the data while also considering the local patterns within the data. The ConvLSTM network consists of LSTM units that are wrapped inside convolutional layers. The input data is processed through the convolutional layers before being fed to the LSTM units, and the output from the LSTM units is passed through the deconvolutional layers to generate the final prediction.

CNN-LSTM is another type of neural network that combines the strengths of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks for processing sequential data, such as time series data. The CNN layers in the network apply a set of filters to the input sequence to extract relevant features, such as trends and patterns, which are then passed to the LSTM layers. The LSTM layers process the input sequence, remembering important information over time, and output a sequence of hidden states which can be used for classification, regression, or other downstream tasks.

Methodology

The task at hand involves preparing a dataset for sequence modeling and classification. To begin with, the dataset is preprocessed to remove any noise at the beginning and end. Next, the features and labels are separated and combined into a training dataset.

The labels are mapped to one-hot encoded vectors to enable the use of the Categorical Cross Entropy loss function. This function is commonly used for multi-class classification problems and measures the difference between the predicted and actual class distributions.

To standardize the data, the StandardScaler from the scikit-learn library is used. This technique scales the input features to have a mean of zero and a standard deviation of one, making them suitable for use with many machine learning algorithms.

Once the data is preprocessed, it is reshaped to fit the input dimensions of the Long Short-Term Memory (LSTM) model. LSTM is a type of recurrent neural network that is particularly effective for modeling sequential data, due to its ability to remember long-term dependencies. The model is trained for 20 epochs, allowing it to learn patterns and relationships within the data.

To compare the accuracy, precision, and recall of the LSTM model with other approaches, the data is also run through ConvLSTM and CNN LSTM models. ConvLSTM combines the benefits of convolutional neural networks (CNNs) and LSTMs, while CNN LSTM uses a combination of CNN and LSTM layers to model temporal information in image sequences.

Overall, this approach involves a combination of data preprocessing, feature engineering, and model selection to prepare the dataset for sequence modeling and classification. By comparing the performance of different models, it is possible to identify the most effective approach for the given task.

Data Analysis

To facilitate data analysis, we will illustrate the distribution of the data and the variations in sensor values across the collected driving behaviors. All the results and charts presented in this section can be replicated using a dense neural network (dense-NN) for further analysis.

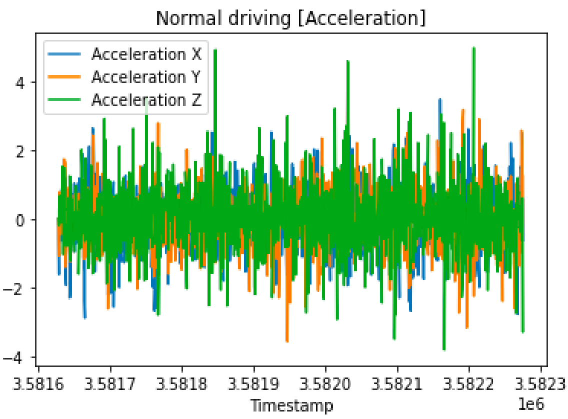


Figure 1

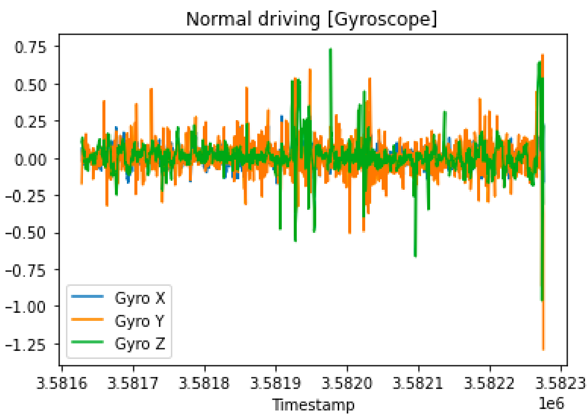


Figure 2

The following data displays the sensor values obtained through an accelerometer. It is important to consider that the Z axis will inevitably contain a considerable amount of noise due to road irregularities. The X axis, on the other hand, is particularly important for measuring acceleration and braking forces, while the Y axis represents lateral forces, which can have an impact on these measurements.

When analyzing sudden turns, the gyroscope is the most significant sensor to consider. This is because it provides information about rotation around the Z and Y axes. It is important to note that a rotation around the X axis is highly unlikely unless the vehicle is rolling with its wheels turned upside down. Therefore, the data collected from the gyroscope can help us understand the vehicle's rotational movements during sudden turns.

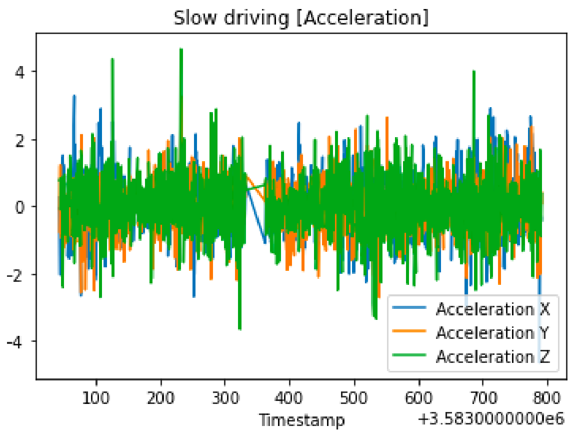


Figure 3

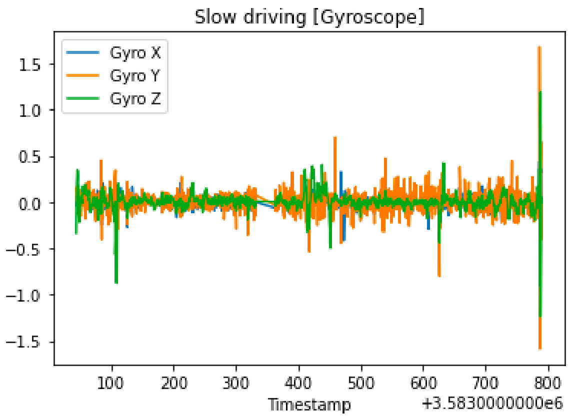


Figure 4

Based on the data collected for slow driving, Figure 3 has many similarities with Figure 1. However, it is evident that the dominance of acceleration on the X axis is reduced in Figure 3, as the brakes and acceleration were applied more gently during slow driving.

Figure 4 displays the gyroscope values collected for slow driving, and it shows more distinct differences when compared to Figure 2, than the differences between Figure 1 and 3. It is noticeable that despite the presence of noise at the beginning and end of the data, the amplitude of the gyroscope values is smaller in Figure 4.

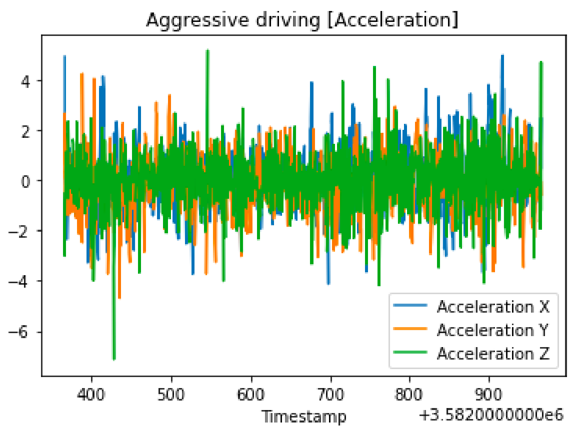


Figure 5

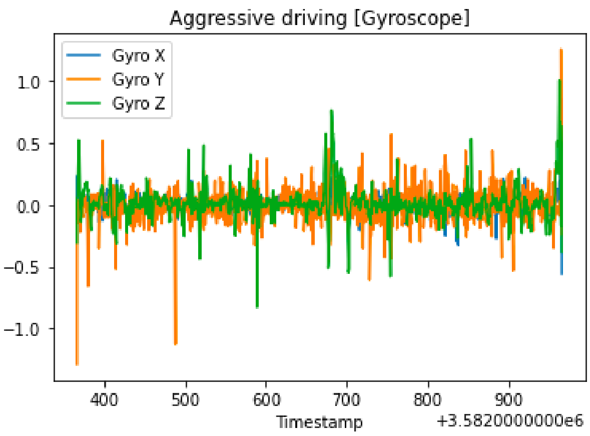


Figure 6

The accelerometer values presented in Figure 5 exhibit noticeable differences compared to Figures 1 and 3. In this case, it is evident that there are more pronounced accelerations and brakes, with some values reaching -6. Additionally, lateral forces are more prevalent, indicating sudden accelerations and turns.

The gyroscope values presented in Figure 6 correspond to aggressive driving, and it is apparent that sudden turns are much more frequent, with numerous forces applied on the Y axis.

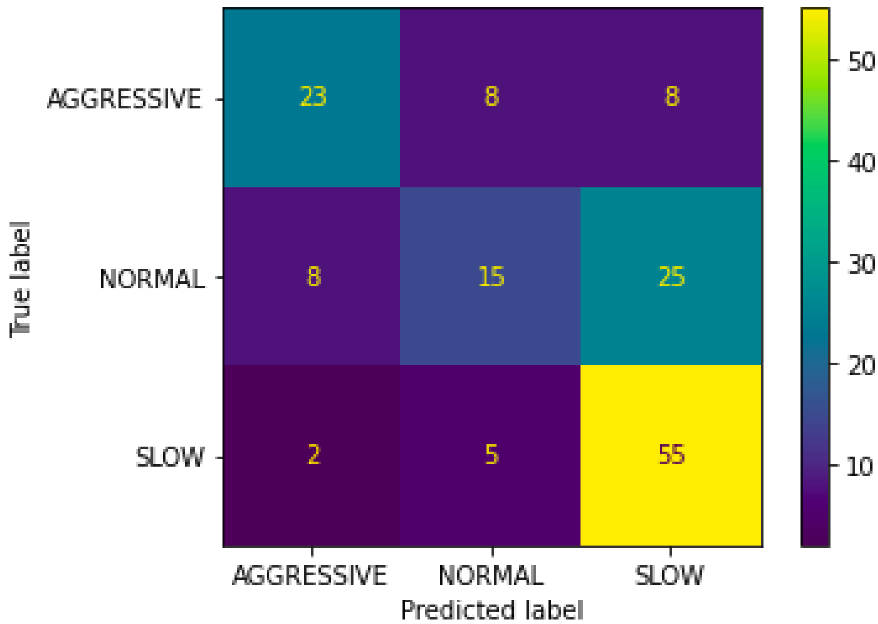
It is important to note that we will not consider the noise present at the beginning and end of the data collection process, as these were caused by moving the phone to start/stop the application or change the driving style. However, despite this noise, it is evident that the forces applied during aggressive driving are more consistent than those observed in Figures 2 and 4, resulting in more distinguishable data in Figure 6. Furthermore, it is noteworthy that the difference between the accelerometer and gyroscope values collected during aggressive driving is more pronounced when compared to the values collected during normal and slow driving behaviors.

Results/Findings

A 3x3 confusion matrix is typically used in classification problems where there are three possible outcomes or classes. The rows represent the actual class, while the columns represent the predicted class. Each cell in the matrix represents the number of instances that fall into a particular combination of actual and predicted classes.

Metrics	LSTM Classifier	Conv-LSTM	CNN LSTM
Accuracy	**0.6242**	**0.6040**	**0.5705**
Precision	**0.6829**	**0.5956**	**0.5960**
Recall	**0.1879**	**0.5436**	**0.3960**

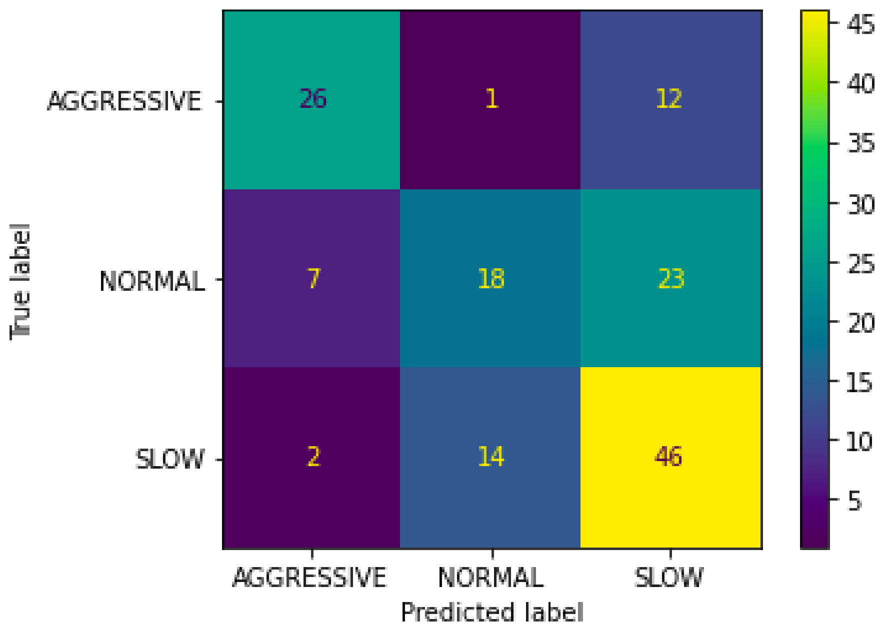
In LSTM classifier, the accuracy of 0.6242 means that the model correctly predicted the outcome for 62.42% of the total instances in the dataset. This is a commonly used metric to evaluate the overall performance of a classification model. The precision of 0.6829 means that when the model predicted a positive outcome, it was correct 68.29% of the time. Precision is a metric that tells you how many of the positive predictions made by the model were correct. The recall of 0.1879 means that the model correctly identified 18.79% of the positive instances in the dataset.



LSTM Classifier

The accuracy of the Conv-LSTM classifier is 0.6040, which means that it correctly predicts the class of around 60.4% of the total instances. The precision of the classifier is 0.5956, which means that out of all the instances that the classifier predicted as positive, around 59.6% of them were positive. The recall of the classifier is 0.5436, which means that out of all the instances that were positive, the classifier correctly identified around 54.4% of them.

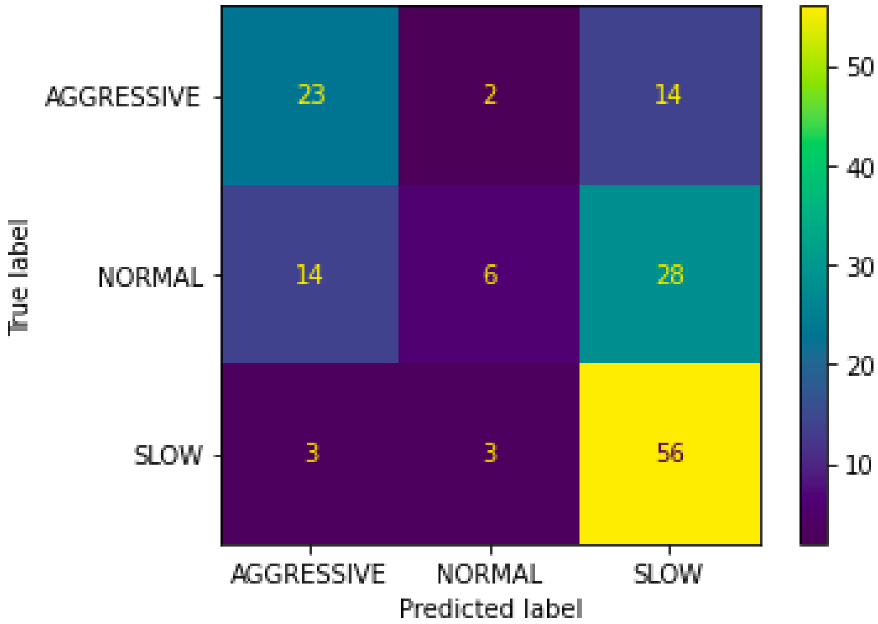
Based on these metrics, we can say that the classifier is performing moderately well, but there is room for improvement. The precision and recall values are close, indicating that the model is not biased towards either false positives or false negatives. However, the accuracy could be improved to classify the instances more accurately.



Conv-LSTM

The accuracy of CNN LSTM model is 0.5705, which means that the model correctly predicted the class label for 57.05% of the instances in the test set. The precision of the model is 0.5960, which means that out of all the instances that the model predicted to be positive, only 59.60% of them belong to the positive class. In other words, the model has a moderate level of confidence in its positive predictions. The recall of the model is 0.3960, which means that out of all the instances that belong to the positive class, the model correctly identified 39.60% of them as positive. In other words, the model has a lower sensitivity to the positive class.

Overall, the precision of the model is relatively higher than its recall, which indicates that the model may be biased towards making positive predictions. However, the accuracy value suggests that the model is still performing better than a random classifier.



CNN LSTM

Limitations

The data collected from sensors in an android device can be affected by various factors such as road conditions, device placement, and environmental factors, among others. These factors may lead to variations in the data collected, and it is essential to consider them when interpreting the results.

Furthermore, it is also important to consider the driver's condition while analyzing the driving behavior data. Factors such as fatigue, stress, and distraction can significantly impact a driver's behavior, and this information can provide valuable insights into understanding the driving patterns observed.

Finally, route analysis is also an important factor to consider when analyzing driving behavior data. The route's characteristics, such as traffic density, terrain, and weather, can significantly impact driving behavior. Therefore, it is crucial to consider these factors while interpreting the data to ensure accurate results.

Conclusion

The objective of the study was to have an analysis on driving behavior based on acceleration and gyroscope measures through deep learning. It seems that the LSTM-classifier has the highest accuracy and precision, but the lowest recall. The Conv-LSTM model has a relatively lower accuracy and precision compared to the LSTM-classifier, but a higher recall. The CNN-LSTM model has the lowest accuracy and precision, but a higher recall than the LSTM-classifier.

To determine which model is better, it depends on the specific requirements of the problem you are trying to solve. If the main goal is to maximize overall accuracy and precision, then the LSTM-classifier may be the best choice. However, if the focus is on correctly identifying instances of a particular class (i.e., high recall), then the Conv-LSTM or CNN-LSTM may be a better option.

Future Directions

Integration of multiple sensor modalities: With advances in technology, it is now possible to collect data from multiple sensors, such as cameras, radar, lidar, and GPS, simultaneously. Integrating data from multiple sensors can provide a more comprehensive understanding of driving behavior and lead to more accurate predictions of driver actions.

Real-time analysis and feedback: Real-time analysis of driving behavior can provide immediate feedback to the driver, which can help improve driving safety and performance.

Personalized driving behavior analysis: Ride sharing apps like Uber, Lyft can incorporate driving styles (Normal, Slow, Fast) for customers to choose from depending upon their need to reach the destination.

References

- Driving Behavior - https://www.kaggle.com/datasets/outofskills/driving-behavior
- RochI- http://rochi.utcluj.ro/articole/10/RoCHI2022-Cojocaru-I-2.pdf
- Driving Behavior analysis - http://rochi.utcluj.ro/articole/10/RoCHI2022-Cojocaru-I-2.pdf
- "Driving Style Recognition with Deep Learning" by Mohit Jain and Vignesh Jagadeeshan - https://ieeexplore.ieee.org/document/8547731
- "A Comprehensive Study of Deep Learning for Driver Behavior Analysis" by Yu Wu, Shuo Wang, and Mohan Trivedi - https://arxiv.org/abs/1711.03935
- "Driving Behavior Analysis with Deep Learning: A Review" by Yifan Peng, Lu Li, and Wenjie Huang - https://www.mdpi.com/1424-8220/19/24/5536/htm
- "A Comparative Study of Deep Learning Techniques for Driver Behavior Analysis" by Amir Pourmohammad and Hamidreza Amindavar - https://ieeexplore.ieee.org/document/8693267
- "Deep Learning for Driver Behavior Analysis: A Review" by Yunlei Zhang, Zhixiong Zou, and Lianru Gao - https://www.sciencedirect.com/science/article/pii/S1877050919315536