Capstone Project – Autonomous Driving

Part 1 – Vehicle Detection (Deep Learning + Bounding Box)

Step 1 – Import Required Libraries

All required libraries were imported:

- **Pandas, NumPy** for data handling
- **Matplotlib/Seaborn** for visualization
- **OpenCV** for image preprocessing
- **TensorFlow/Keras** for deep learning model building

Step 2 – Load Dataset (Images + Labels)

- Loaded the dataset consisting of `Images/` and the annotation file `labels.csv`.
- Verified that `labels.csv` contained image IDs, class labels, and bounding box coordinates.

Step 3 – Preprocess Images

- Resized all images to **224×224 pixels**.
- Converted color channels from BGR to RGB.
- Normalized pixel values to the range [0,1].
- Verified preprocessing using sample outputs.

Step 4 – Prepare Bounding Box Data

- Extracted bounding box values (x, y, width, height).
- Normalized bounding box coordinates relative to image dimensions.
- This normalization ensured stable training and scale-independent predictions.

Step 5 – Build a CNN Model

- Constructed a **multi-output CNN** using Keras Functional API.
- **Two heads in the model:**
  - **Classification head (Softmax):** predicts vehicle type.
  - **Regression head (Sigmoid):** predicts bounding box coordinates.
- Compiled the model with:
  - `categorical_crossentropy` for classification loss
  - `mean_squared_error (mse)` for bounding box regression

Step 6 – Train & Validate the CNN Model

- Split dataset: **80% training, 20% testing**.
- Trained for multiple epochs.
- Observed training/validation accuracy and bounding box loss trends.

Step 7 – Visualize Predictions

- Plotted sample test images with bounding boxes:
  - **Green box = Ground Truth**
  - **Red box = Model Prediction**
- Visual inspection highlighted gaps in bounding box accuracy.

Step 8 – Fine-tune the Model

- Continued training for additional epochs.
- Classification accuracy improved with more epochs.
- Bounding box regression remained challenging due to dataset limitations.

Step 9 – Evaluate on Test Data

- Evaluated the trained model using:
  - **Classification Report:** precision, recall, and F1-score for each class.
  - **Average IoU (Intersection over Union):** for bounding box accuracy.

Final Results:

- Classification Report successfully generated.
- ✅ **Average IoU on test samples: 0.007850177420282108**

✅ **Part 1 Completed – Vehicle Detection Model (Steps 1 to 9)**

Autonomous Driving Accident Analysis –

Part 1 – Data Preparation

- The accident dataset was imported from CSV.
- Data cleaning steps included removing duplicates, handling missing values, and standardizing column names.
- After preprocessing, the dataset was ready for exploratory analysis.
- Initial inspection gave an overview of the number of records and highlighted that accident reporting increased in recent years.

Part 2 – Exploratory Data Analysis (EDA)

2.1 Events Over Time

- **Year-wise trend:** Tesla accidents increased steadily, especially after 2018.
- **Month-wise trend:** Seasonal variations were observed in reporting.
- **Country-wise distribution:** The USA reported the highest number of accidents, followed by China, Germany, and Canada.

2.2 Victim Analysis

- **Deaths per accident:** Most crashes involved a single death, but some events had multiple fatalities.
- **Driver fatalities:** A significant share of accidents recorded the Tesla driver among the victims.
- **Model involvement:** Model S and Model 3 appeared most frequently in the dataset.

2.3 Collision Analysis

- **Other vehicles:** Many accidents involved collisions with other vehicles.
- **Cyclists/Pedestrians:** Some accidents included vulnerable road users such as cyclists and pedestrians.

Part 3 – Model & Autopilot Analysis

3.1 Event Distribution Across Tesla Models

- Accident counts varied across Tesla models.
- Model S and Model 3 showed the highest accident counts.

3.2 Verified Tesla Autopilot Deaths Distribution

- Verified Autopilot-linked deaths were counted separately.
- These numbers were smaller compared to the overall accident counts.

3.3 Verified vs All Reported Deaths (NHTSA)

- Comparison between verified Autopilot deaths and all deaths reported to NHTSA showed a large difference.
- Only a subset of overall reported deaths are officially verified as Autopilot-related.

Part 4 – Visualisation & Insights

Visualisations

- Bar chart of top countries by accident count.
- Trend of total deaths per year.
- Bar charts for Tesla models and Autopilot-related deaths.

Key Insights

- Accident counts rose sharply after 2013, especially post-2018.
- The USA accounted for the majority of accident reports.
- Most crashes involved one fatality, though some were more severe.
- Tesla drivers were frequently among the fatalities.
- Model S and Model 3 appeared most in accident records.
- Most accidents involved other vehicles; some included cyclists/pedestrians.
- Verified Autopilot deaths were much fewer than all deaths reported to NHTSA.

Conclusion

The project covered data preparation, exploratory analysis, model and Autopilot analysis, and final insights. The findings show trends in Tesla accidents by year, geography, victim type, collision type, and Autopilot involvement.