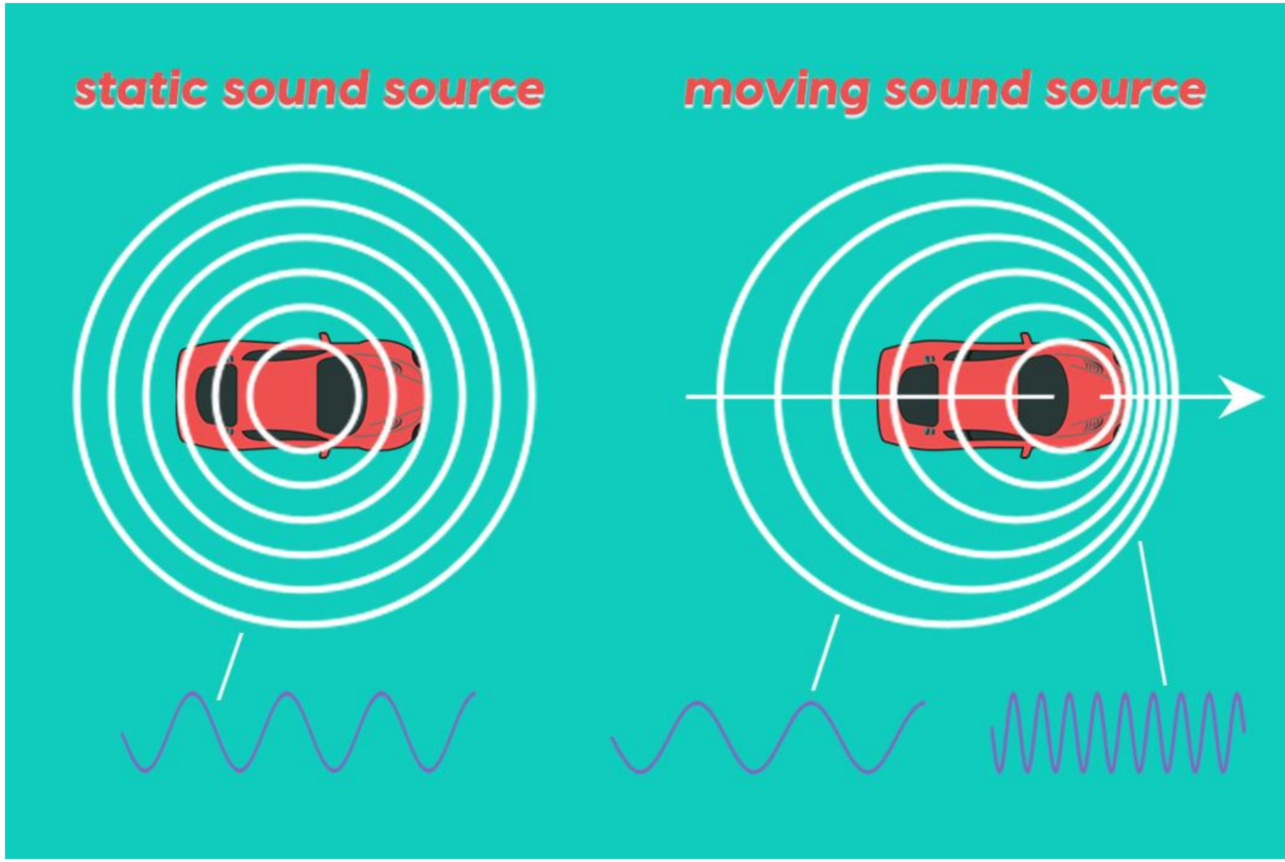
MTE 546: PROJECT

DOPPLER EFFECT BASED SENSOR FUSION FOR SPEED ESTIMATION



Group 9: Cameron Ritchie, Connor Johansson, Christopher Nikolik, Shivraj Singh Vishnoi, Shourrya Guha.

Department Name: Mechanical and Mechatronics Engineering Department

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

INTRODUCTION

Project Introduction and motivation

Introduction

Traditional speed measurement relies on radar, LIDAR, or GPS, but what if we could estimate speed using just acoustic sensors? This experiment explores a low-cost passive method to measure a car’s speed using the Doppler effect

Assumptions:

- The car produces a consistent, identifiable, high frequency sound.
- Phones are time-synchronized for accurate delay measurement.
- Environmental conditions remain stable during the test.
- No other sounds consistently at the same frequency

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

Motivation

Low-Cost Emergency Vehicle Speed Detection

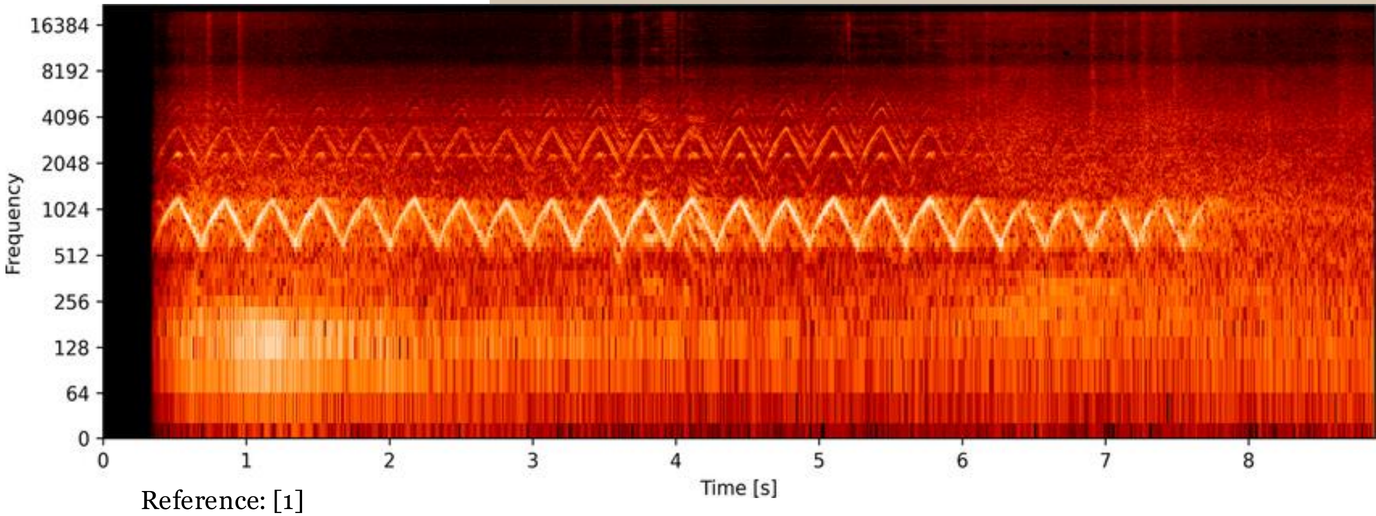
- Helps traffic management systems detect approaching ambulances and prioritize traffic signals.

Potentially Cheaper and More Effective Alternative to Speed Cameras

- Works in areas where vision is obstructed, or cameras may be unusable (tunnels, dense urban zones).
- Low-cost alternative to LIDAR/radar-based speed traps.

Other Applications

- Could be used for mobile robots in industrial/manufacturing situations where obstructions/weak signals are present



Reference: [1]

BACKGROUND

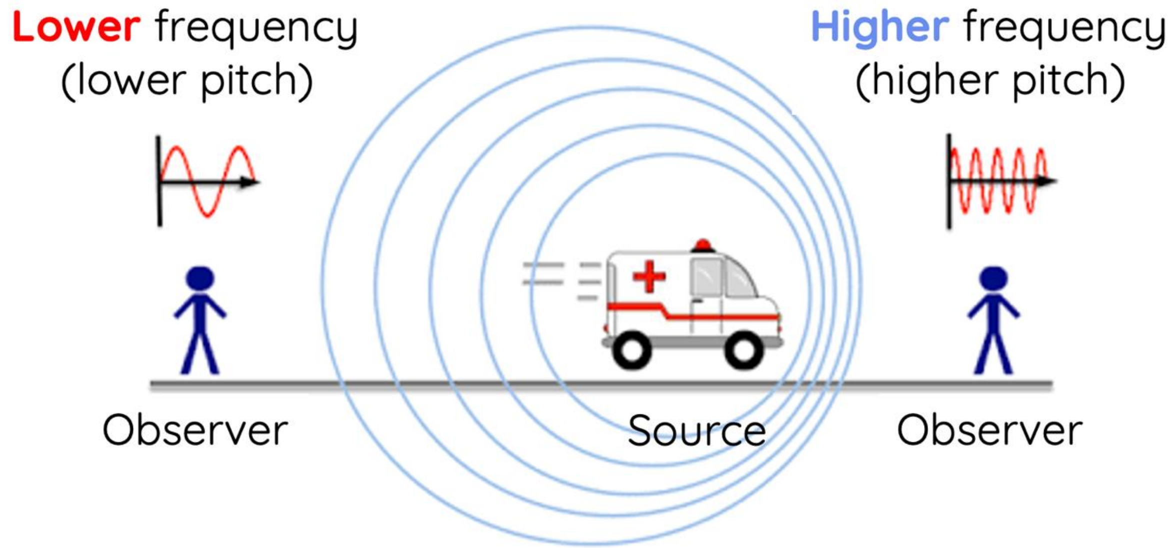
Doppler Effect and sensors used

Doppler Effect

The Doppler effect can be described as the effect produced by a moving source of waves in which there is an apparent upward shift in frequency for observers towards whom the source is approaching and an apparent downward shift in frequency for observers from whom the source is receding [2].

Formula: $f_o = \frac{v+v_0}{v+v_s \cos\theta} f_s$

- f_0 Observer frequency of sound
- f_s Actual frequency of sound waves
- v_o Observer velocity
- v Speed of sound waves
- v_s Speed of source
- θ angle between line of sight of source and observer



Sensors Used

Smartphones

- Record high frequency speaker tone as the car moves
- Used to analyze the **Doppler shift** in sound frequency
- Positioned along a **100m** straight path at equal intervals
- Capture time delay between sound arrival at different locations

Ground Truth

Car's Speedometer

- Provide real-time speed readings
- Used as a baseline for accuracy evaluation

Experimental Setup

1. Audio receivers are placed at points 1 through 5
2. The car begins speeding up at ‘Start’ and gets to the required speed at ‘Pickup speed’.
3. The car travels from point 1 to 5 at 10, 20, 30, 40 and 50 kph constant speeds.
4. A speaker playing a 2000 Hz tone is held outside the car window for the receivers.

Drawbacks

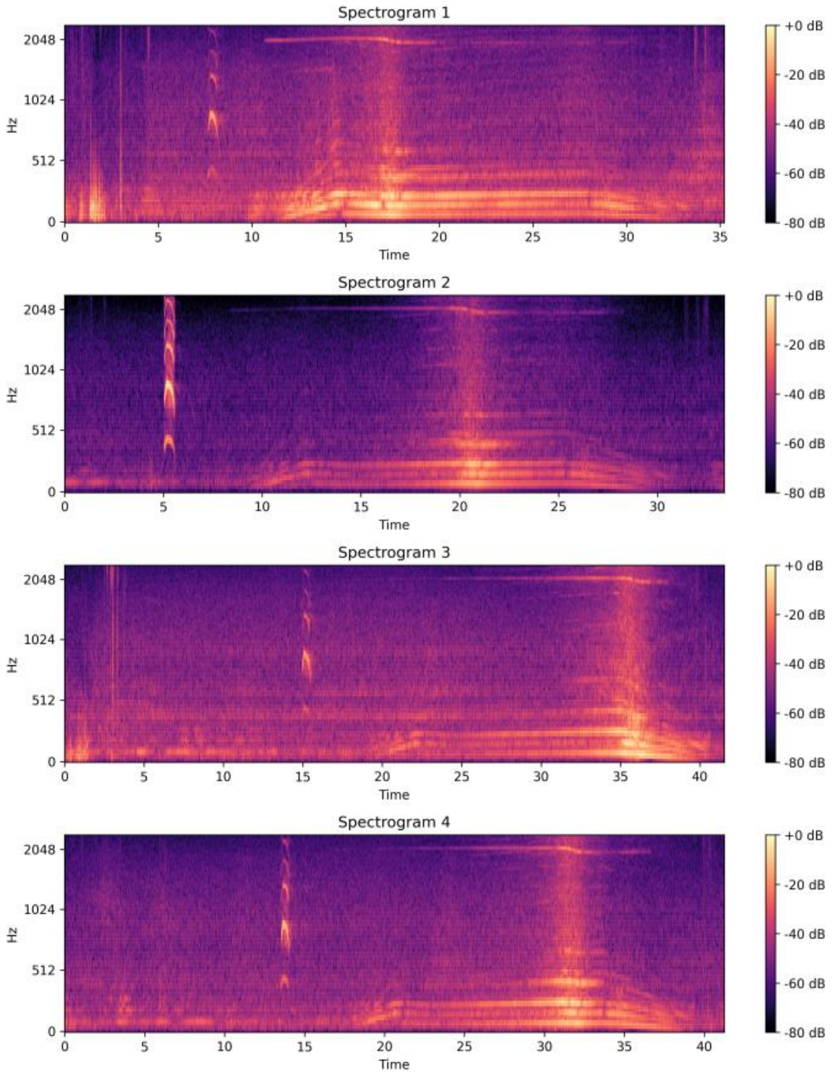
At speeds higher than 40kpmh, the car did not reach the required speed till point 1. Hence there might be a slight offset in speed that needs to be accounted for.

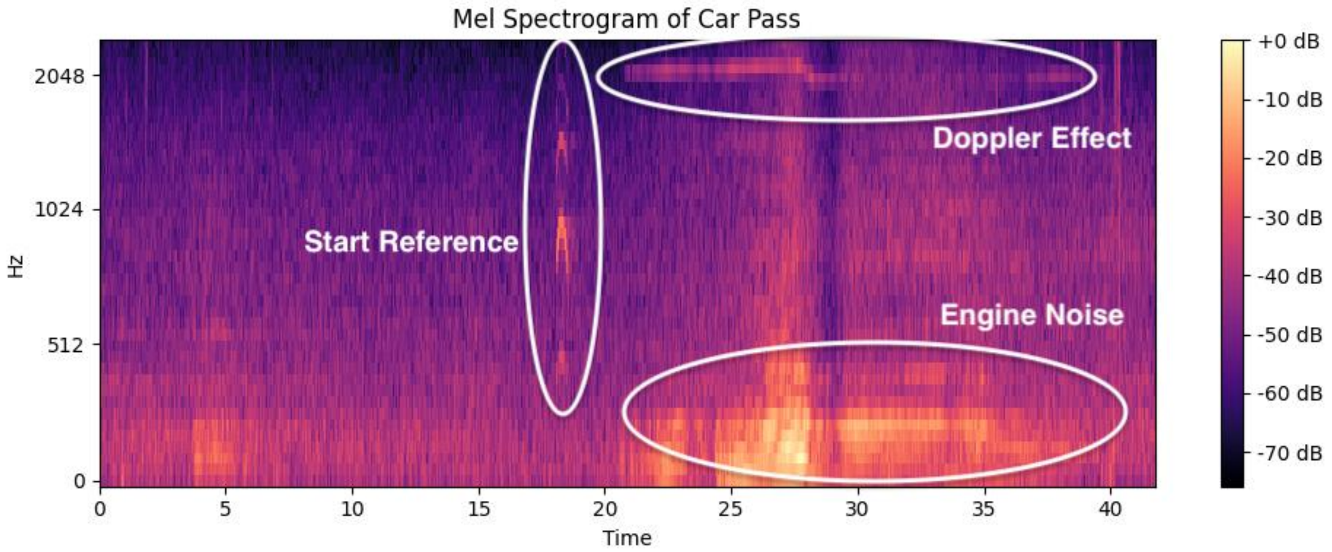
UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

SENSOR FUSION

Unscented Kalman Filter and Convolutional Neural Network

Spectrogram Results

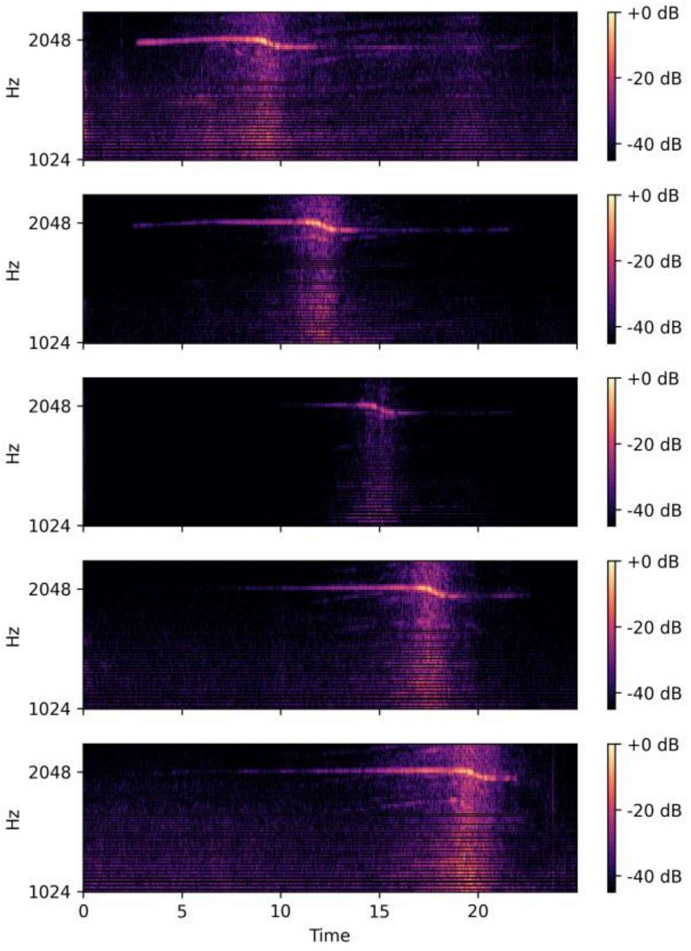




A spectrogram is a visual representation of the spectrum of frequencies of a signal as it varies with time[3]. Emergency vehicle signals have a clear identifiable frequency on the spectrogram[1].

It shows a color intensity which indicates the amplitude (or decibel level) of the signal at each frequency. Brighter or warmer colors signify higher intensities, while darker or cooler colors represent lower intensities.

Vehicle Progression Spectrogram Results

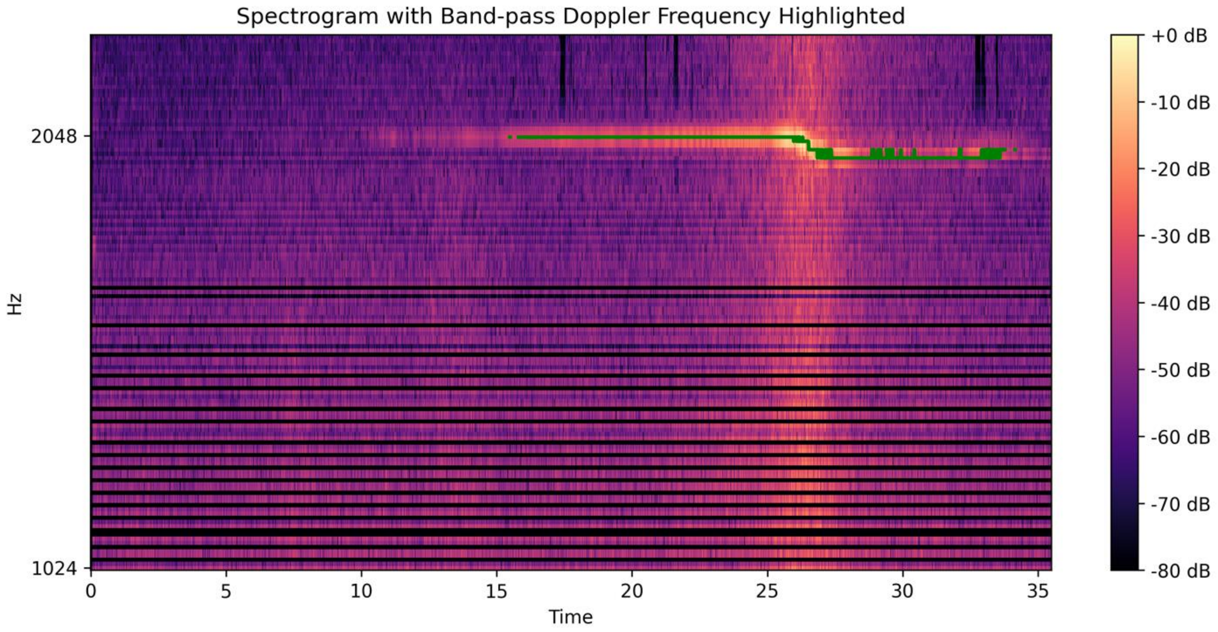


- **Time-Shifted Peak Intensity** – The peak signal intensity gradually shifts from left to right across the spectrograms, indicating the movement of the vehicle over time.
- **Consistent Frequency Signature** – The dominant frequency components remain relatively stable, corresponding to the 2000Hz sound source
- **Sensor-to-Sensor Propagation** – Each subplot (representing a different sensor) captures the same event at slightly different times, demonstrating the sequential detection of the moving vehicle.

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

Spectrogram Processing

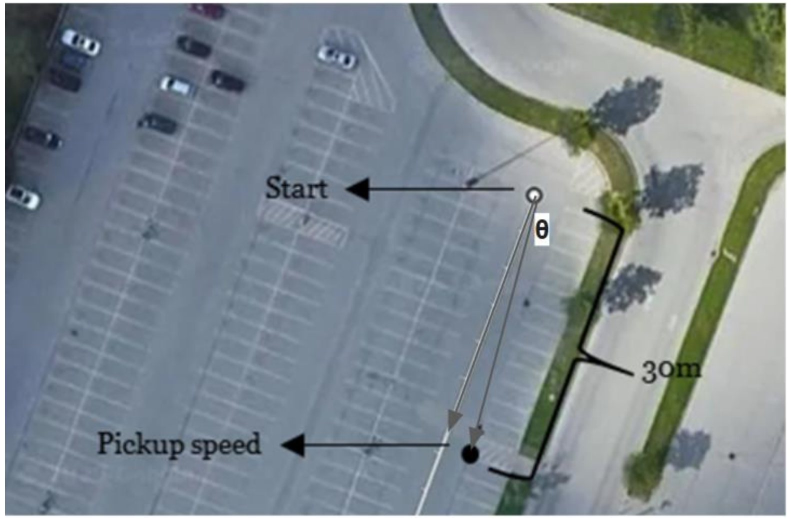
- **Spectrogram Filtering** – Applied a band-pass filter (1940–2100 Hz) to isolate the vehicle's 2000 Hz signal and accept values only within a certain decibel range.
- **Overlay Highlight** – Added an overlay to the high frequency audio signal.
- **Data Logging** – Logged frequency values and corresponding timestamps in a CSV file.
- **UKF Input** – Formatted data for Unscented Kalman Filter processing.



UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

Sensor Fusion - UKF

$$x_k = \begin{bmatrix} x_k \\ y_k \\ v_{x_k} \\ v_{y_k} \\ a_{x_k} \\ a_{y_k} \end{bmatrix}$$



$$\begin{bmatrix} h(x_k)_1 \\ h(x_k)_2 \\ h(x_k)_3 \\ h(x_k)_4 \\ h(x_k)_5 \end{bmatrix} = \begin{bmatrix} f_s \cdot \dfrac{v}{v - \left(\begin{bmatrix} v_{x_1} & v_{y_1} \end{bmatrix} \cdot \dfrac{\begin{bmatrix} x'_1 \\ y'_1 \end{bmatrix}}{\sqrt{x'^2_1 + y'^2_1}} \right)} \\ f_s \cdot \dfrac{v}{v - \left(\begin{bmatrix} v_{x_2} & v_{y_2} \end{bmatrix} \cdot \dfrac{\begin{bmatrix} x'_2 \\ y'_2 \end{bmatrix}}{\sqrt{x'^2_2 + y'^2_2}} \right)} \\ f_s \cdot \dfrac{v}{v - \left(\begin{bmatrix} v_{x_3} & v_{y_3} \end{bmatrix} \cdot \dfrac{\begin{bmatrix} x'_3 \\ y'_3 \end{bmatrix}}{\sqrt{x'^2_3 + y'^2_3}} \right)} \\ f_s \cdot \dfrac{v}{v - \left(\begin{bmatrix} v_{x_4} & v_{y_4} \end{bmatrix} \cdot \dfrac{\begin{bmatrix} x'_4 \\ y'_4 \end{bmatrix}}{\sqrt{x'^2_4 + y'^2_4}} \right)} \\ f_s \cdot \dfrac{v}{v - \left(\begin{bmatrix} v_{x_5} & v_{y_5} \end{bmatrix} \cdot \dfrac{\begin{bmatrix} x'_5 \\ y'_5 \end{bmatrix}}{\sqrt{x'^2_5 + y'^2_5}} \right)} \end{bmatrix}$$

$$f(x) = \begin{bmatrix} 1 & 0 & \Delta t & 0 & \frac{1}{2}\Delta t^2 & 0 \\ 0 & 1 & 0 & \Delta t & 0 & \frac{1}{2}\Delta t^2 \\ 0 & 0 & 1 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 1 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$z_k = \begin{bmatrix} f_{o,1} \\ f_{o,2} \\ f_{o,3} \\ f_{o,4} \\ f_{o,5} \end{bmatrix}$$

Tuning Process for UKF:

Process Noise Covariance (Q)
If Q is too large -> UKF may overreact to measurements
If Q is too small -> UKF may ignore system uncertainties, leading to poor tracking

μ: Sigma Points
For a 'n' dimensional space '2n+1' Sigma points are needed. Sigma points are not unique

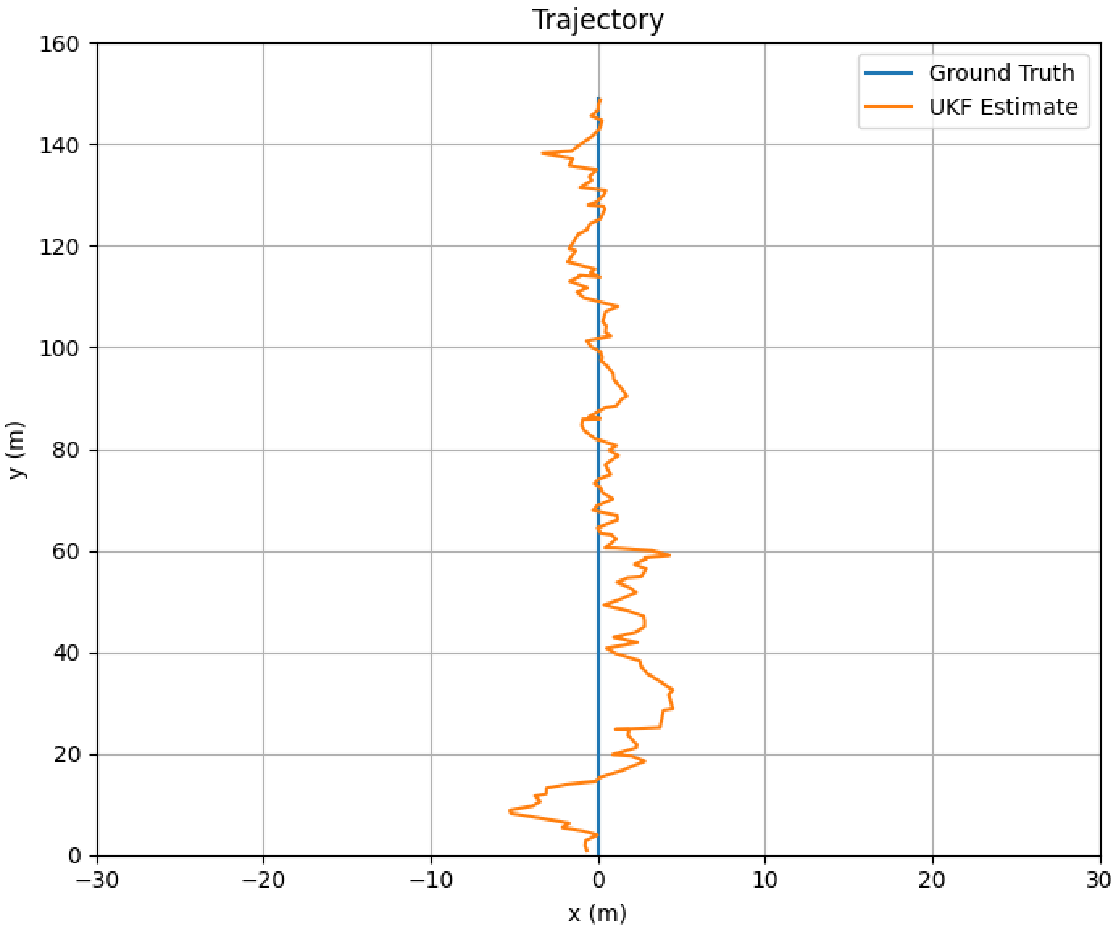
α: Scaling Parameter
Typically 0.001 for high dimensional systems.
Values range between 0 to 1 ($\alpha \epsilon (0,1]$).
Tuning: Smaller Values -> Reduce nonlinearity but may cause underestimation
Large Values -> Increases accuracy but may cause instability

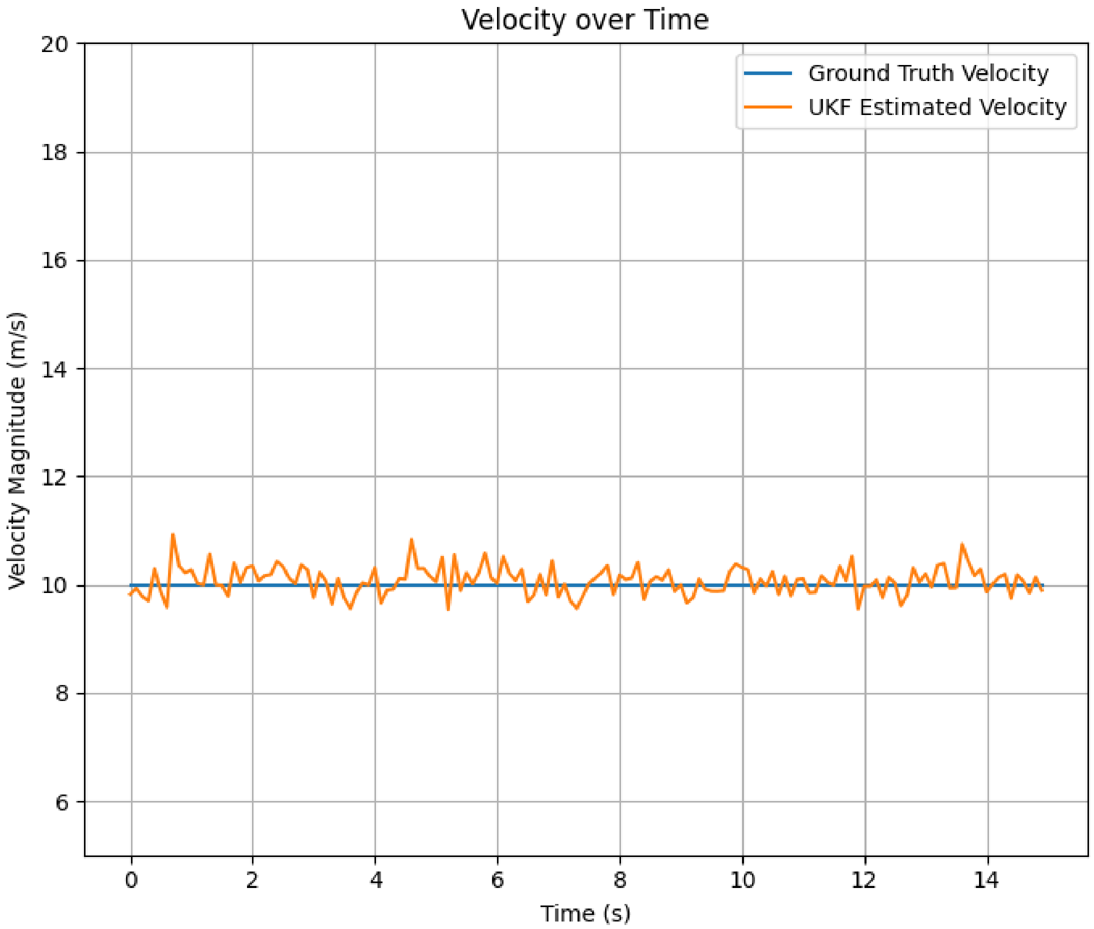
Measurement Noise Covariance (R)
If R is too large -> Filter relies on prediction rather than measurements
If R is too small -> Filter relies on noisy measurements.

β: Prior knowledge about the distribution
Set to 2 for a Gaussian distribution
Tuning: Start with 2 and adjust based on performance

κ: Secondary Scaling
$\kappa \geq 0$ for a Gaussian distribution
Tuning: Lower Values -> Reduces impact of higher order terms
Large Values -> Broader spread, useful for highly nonlinear system

UKF - Results

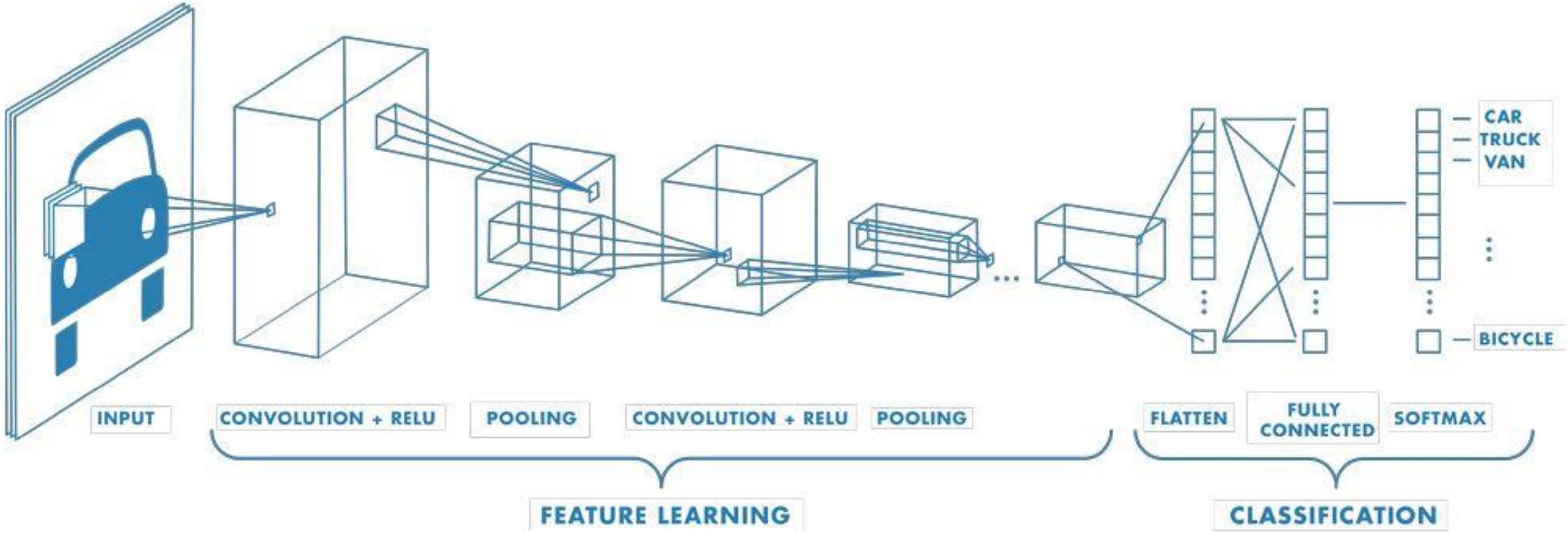




UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

Convolutional Neural Network

- Like an ANN for 2D data
- Filters act as specialized neurons looking for patterns
  - Like a 2D matrix of weights that is moved over the entire dataset. Each filter tries to identify important features
- Same learning methods used
- Common layers:
  - Convolution: putting images through filters
  - ReLU
  - Pooling: down-sampling to improve speed and emphasize features
  - Dense: fully connected layer (same as in ANNs)
  - Flatten: turns multi-dimensional data into 1D. Prepares data for decision



Reference [4]

CNN Pipeline

1. Process Data and Convert to Spectrogram
   - Padding with -45dB data to ensure all data is equal in length
   - Filter out undesirable frequencies (1024 < desired < 2400 Hz)
2. Prepare Data for Training
   - Reshape data into required sizes
   - Normalize data
     - All signal power below -45 dB is clipped to -45 dB. Then normalized so -45 dB = 0 and 0dB = 1
     - Important to ensure proper scaling during learning
   - Split data randomly into training and validation sets
     - 80/20 training/test

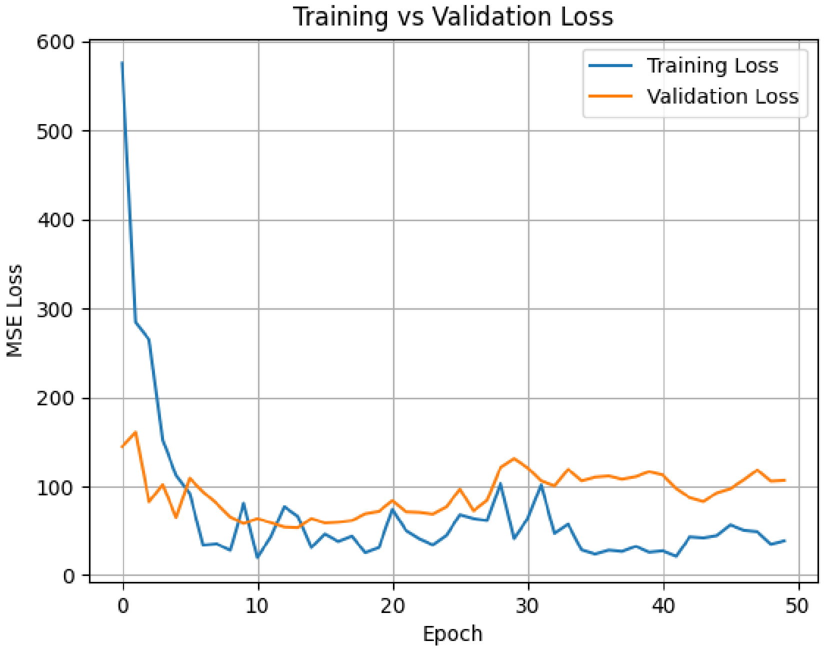
CNN Pipeline Continued: Architecture

3. Build CNN
   - Basic Layers:
     1. Convolution of *8* filters
     2. Pooling
     3. Convolution of *16* filters
     4. Pooling
     5. Flatten
     6. Dropout – decreases likelihood of overfitting by randomly turning off neurons (prevents reliance on single neuron)
     7. Dense (final speed prediction) – No classification in our case

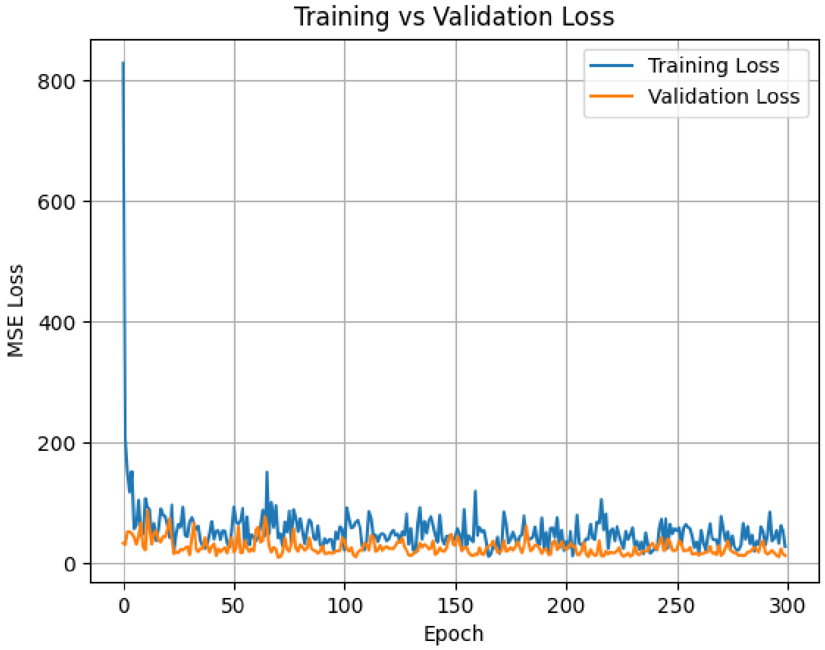
Training

- Optimizer: Adam with $\eta = 0.001$
- Loss: mean squared error
- Metric: mean absolute error
- Epochs: 75 - 300
- Batch Size: 4
- Training time: 3 - 10 mins

Overfitting (too many filters)



Longer Training Time



Select Training Results

Epochs	Depth of Convolution and Pooling	Number of Filters	Notes	Results (Mean Average Error)
50	2 layers	Layer 1: 16 Layer 2: 32	Overfit	9.34 km/h
75	2 layers	Layer 1: 8 Layer 2: 32	Removed Sensor 1 (Closest to Start Pt)	3.43 km/h
150	3 layers	Layer 1: 8 Layer 2: 16 Layer 3: 32	N/A	4.70 km/h
300	**2 layers**	**Layer 1: 8** **Layer 2: 16**	**Removed Sensor 1 (Closest to Start Pt)**	**2.87 km/h**

CONCLUSIONS

Results, Validation, Approach and References

Expected Results & Conclusions

- **UKF for Sensor Fusion**
  - In need of further tuning
  - Drawbacks: fairly involved to set up, requiring lots of preprocessing and prior knowns
- **CNN Velocity Prediction**
  - Achieved 2.87 km/h average error
  - Further improvements to the CNN hope to lessen the error further

Approach

Strengths:

- **Low-Cost & Easy Setup** – Uses only smartphones, no extra hardware needed.
- **Data Fusion Improves Accuracy** – Combines Doppler & time delay with UKF.
- **Verifiable Results** – Results can be easily verified by comparing with speedometer readings.

Weaknesses:

- **Noise Sensitivity** – Background sounds can interfere with recordings. Need accurate filtering
- **Sound & Weather Dependency** – Speed of sound varies with environmental conditions.
- **Irregular Speed** – Vehicle did not reach target speed completely before the first data collection point.

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

References

[1] M. Asif, M. Usaid, M. Rashid, T. Rajab, S. Hussain, and S. Wasi, “Large-scale audio dataset for emergency vehicle sirens and road noises,” *Scientific Data*, vol. 9, no. 1, Oct. 2022, doi: https://doi.org/10.1038/s41597-022-01727-2.

[2] The Physics Classroom, “The Doppler Effect,” *The Physics Classroom*, 2019. https://www.physicsclassroom.com/class/waves/Lesson-3/The-Doppler-Effect

[3] “What is a Spectrogram?,” *Pacific Northwest Seismic Network*. https://www.pnsn.org/spectrograms/what-is-a-spectrogram

[4] MathWorks, "Convolutional Neural Network," 28 March 2025. [Online]. Available: https://www.mathworks.com/discovery/convolutional-neural-network.html

QUESTIONS ?