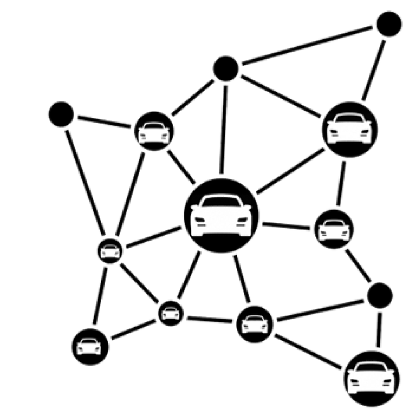
Controlling Steering Angle for Cooperative Self-driving Vehicles utilizing CNN and LSTM-based Deep Networks

Rodolfo Valiente*, Mahdi Zaman*, Sedat Ozer**, Yaser P. Fallah*

*Center for Research in Electric Autonomous Transport (CREAT), **University of Central Florida



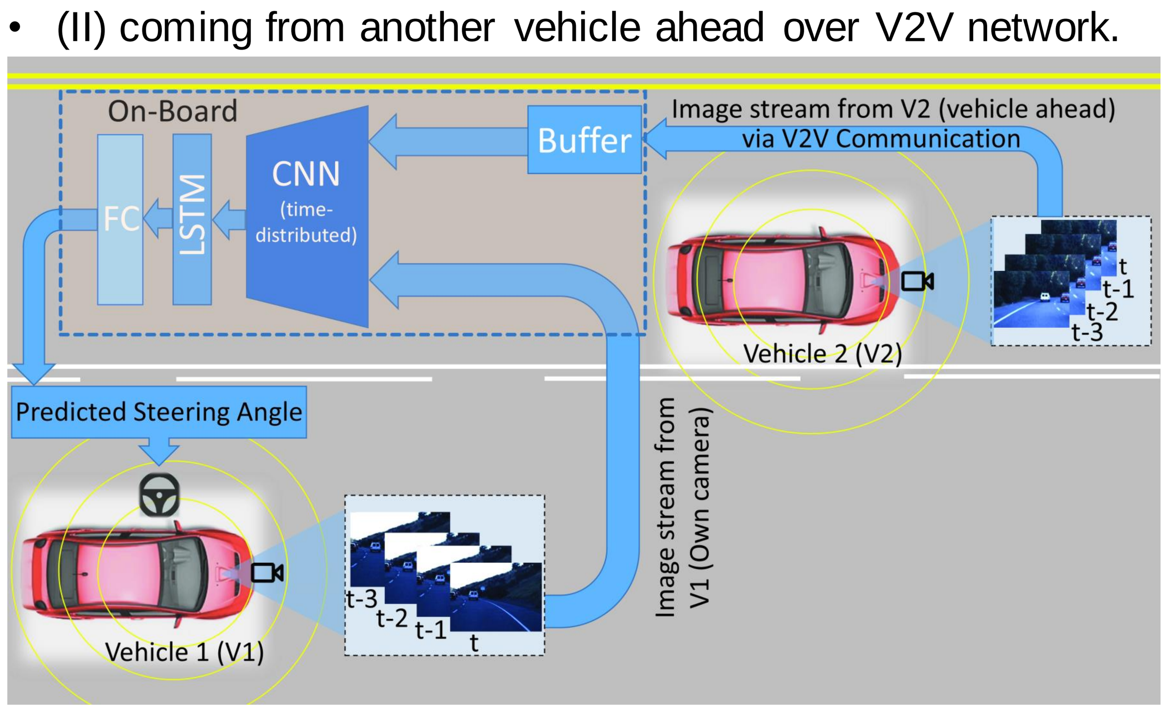
1. Overview

- A fundamental challenge in **Autonomous Vehicles** (AV) is **adjusting the steering angle** at different road conditions.
- We present and study a **new end-to-end deep architecture** to predict the steering angle by using **Long-Short-Term-Memory** (LSTM) for co-operative systems.
- We utilize **multiple sets of images shared** between two AV to improve the accuracy of controlling the steering angle by considering the temporal dependencies between the image frames. Our model uses both present and future images (shared by vehicle ahead via **Vehicle-to-Vehicle (V2V) communication**) as input.
- We demonstrate that using series of images in cooperative systems can significantly improve the accuracy of prediction in terms of RMS and MAE error.

2. Proposed System

Our deep-learning based approach utilizes two sets of images:

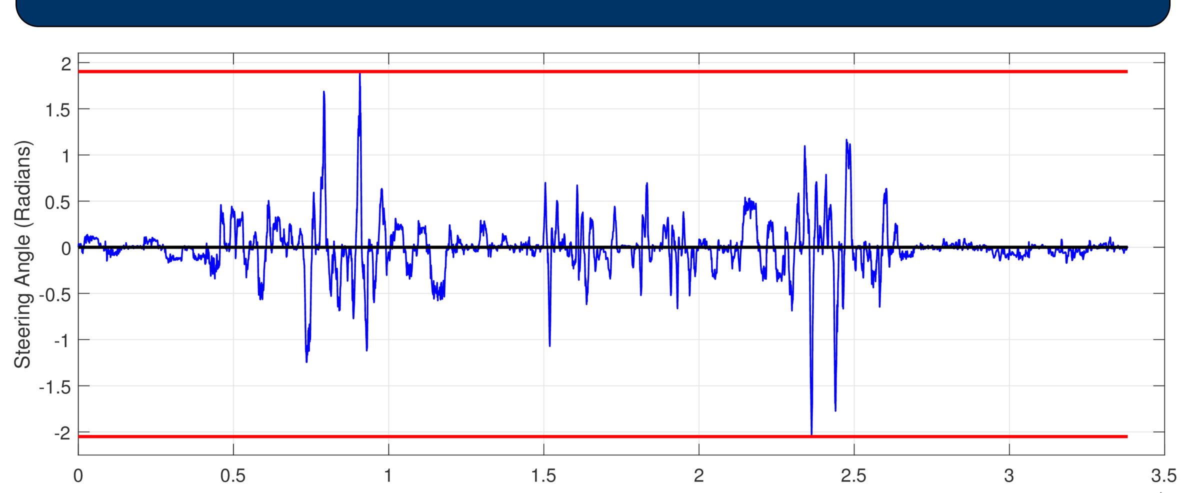
- (I) coming from the onboard sensors e.g. camera,
- (II) coming from another vehicle ahead over V2V network.



The overview of our proposed vehicle-assisted end-to-end system is given above:

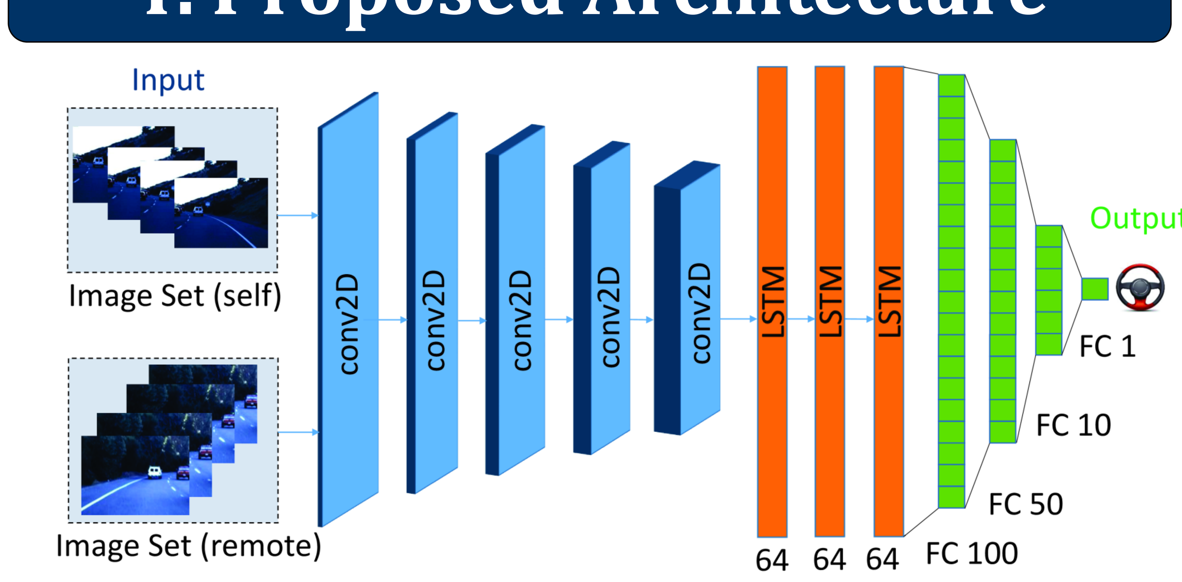
(I) Vehicle 2 (V2) sends his information to Vehicle 1 (V1) over V2V communication. (II) V1 combines that information along with its own information to control the steering angle. (III) The prediction is made through our CNN+LSTM+FC network.

3. Dataset



Visualization of the self-driving car dataset by Udacity: number of images (frames) vs. steering angle of each image. Udacity dataset forms the ground truth for our experiments.

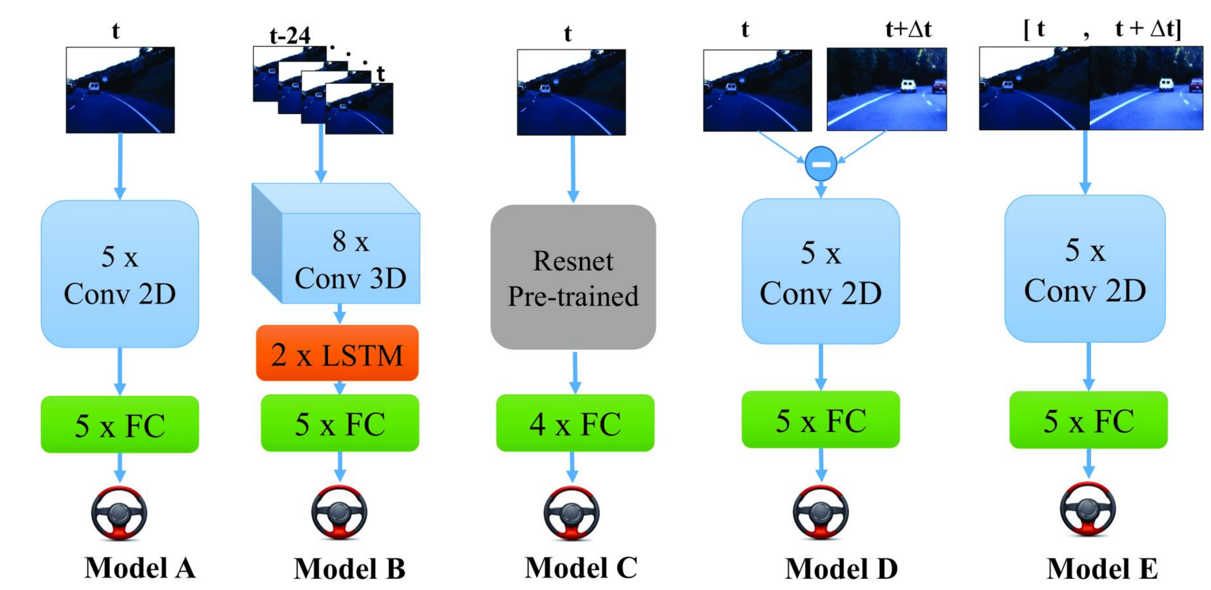
4. Proposed Architecture



- Our CNN + LSTM + FC Image sharing model.
- **Image set (self)**: Images obtained from on-board camera of the Host Vehicle (V1). **Image set (remote)**: Images obtained from vehicle ahead (V2)
- For training, we used: $epochs = 14$, $minibatch\ size = 64$, Adam Optimizer with Learning Rate of 10^{-2}.

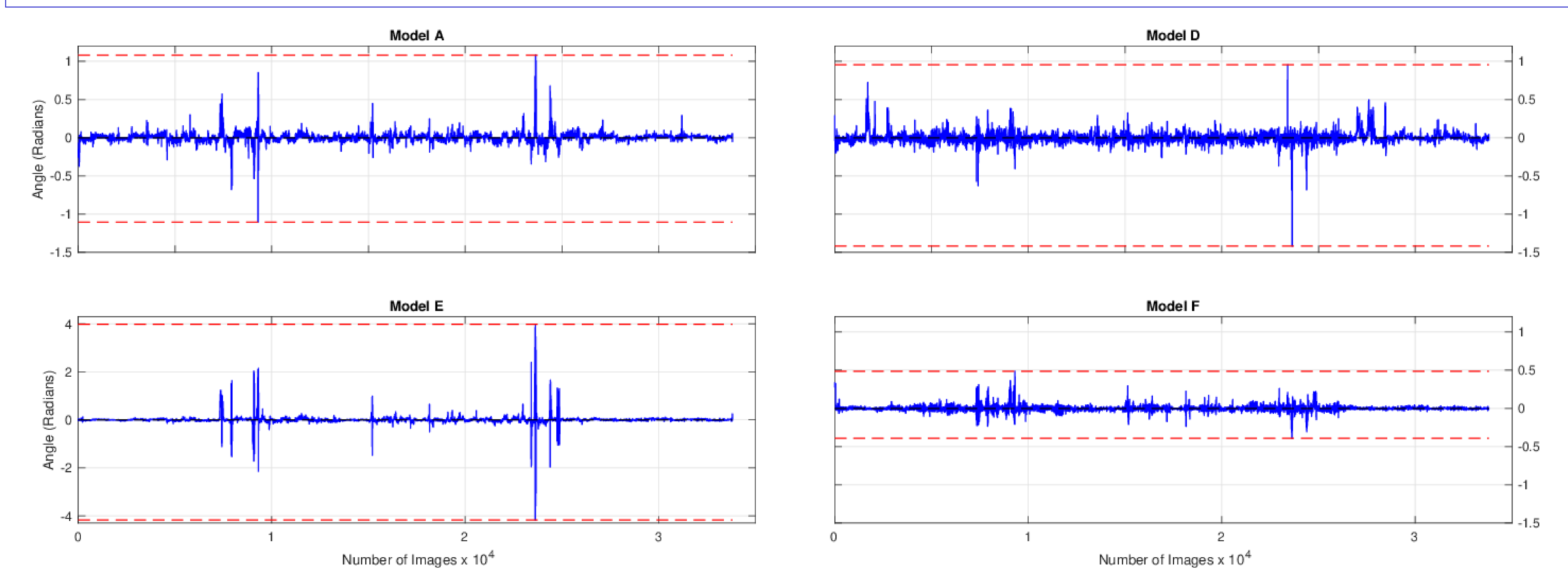
5. Analysis and Results

- Benchmark Models chosen for performance comparison:
  - Model A: Bojarski et al, End to End Learning for Self-Driving Cars. 2016.
  - Model B, C: Shuyang Du, Haoli Guo, and Andrew Simpson. Self-Driving Car Steering Angle Prediction Based on Image Recognition. Technical report, 2017.
  - Model D, E: Dhruv Choudhary and Gaurav Bansal. Convolutional Architectures Self-Driving Cars. Technical report, 2017.

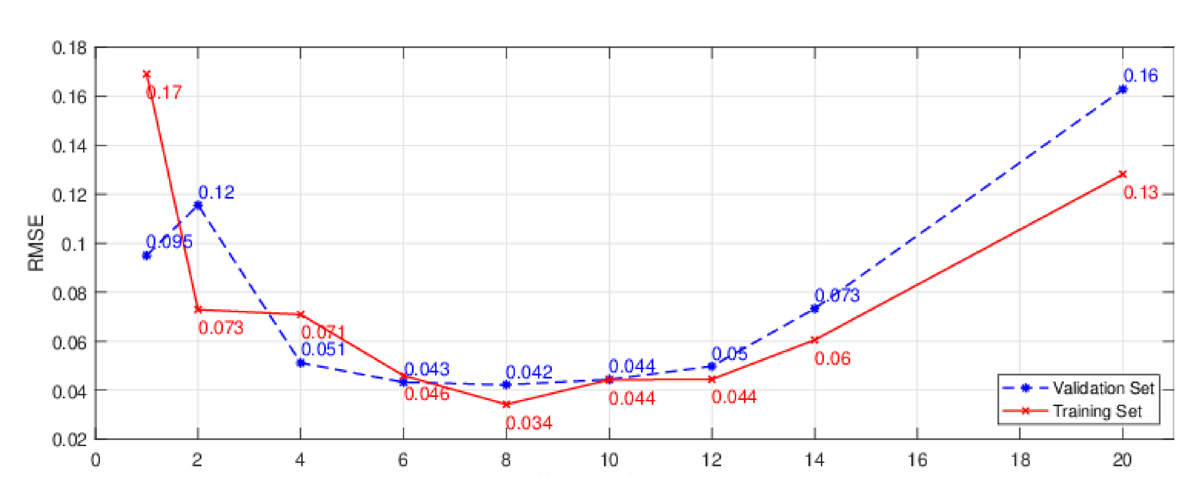


- **Model F and G are our proposed architectures**. Model F uses $x = 8$ images per set, while Model G uses $x = 10$.
- **RMSE and MAE comparison**: our model F (with $x = 8$ and two image sets 30 frames apart) obtains the minimum error values for both training and validation.

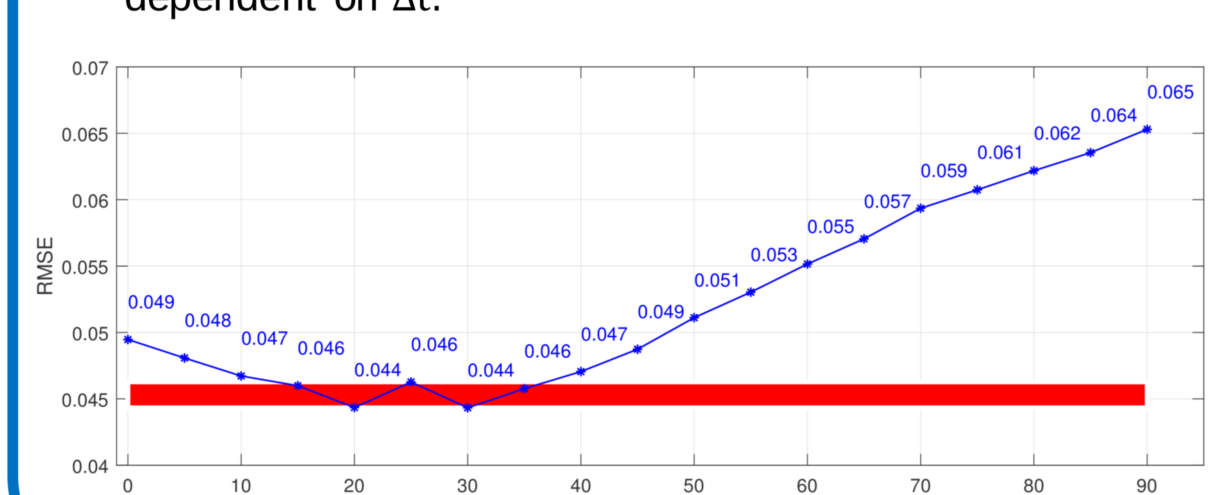
	Model:	A	B	C	D	E	F	G
RMSE	Training	0.099	0.113	0.077	0.061	0.177	**0.034**	0.044
	Validation	0.098	0.112	0.077	0.083	0.149	**0.042**	0.044
MAE	Training	0.067	x	x	0.038	0.046	**0.022**	0.031
	Validation	0.062	x	x	0.041	0.039	**0.033**	0.036

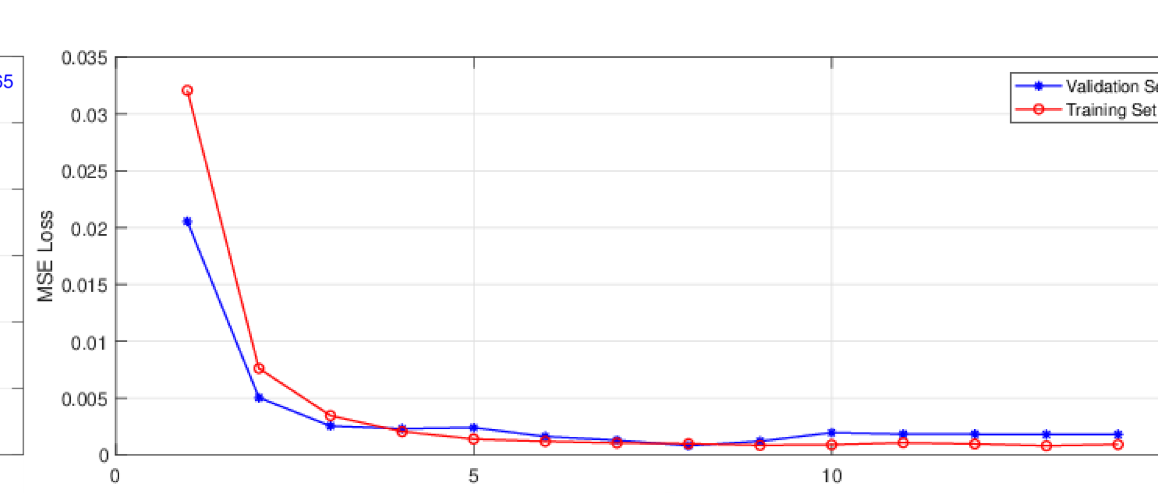


- **Top Figure: Individual error values (in radian)** made at each frame over the whole dataset is plotted for Model A, D, E and F. The upper and lower red lines highlight the maximum and minimum errors.
- **Right and Bottom figures**: we tested with tuning three hyperparameters that affect the system performance;
  - We trained our model at various x values and the minimum $RMSE$ is obtained at $x = 8$, which also provides minimum validation $RMSE$ (0.042).



- Changing Δt corresponds to changing the distance between two vehicles. Within the red area, the $RMSE$ remains lowest (0.044~0.046) but it is non-linearly dependent on Δt.
- On our experiments, $\Delta t = 30$ (i.e Vehicles 1.5 seconds apart). With $x = 8$ and $\Delta t = 30$, the training and validation losses stabilize at 14[th] epoch.





6. Concluding Remarks

- We present a new approach **by sharing images** between **cooperative self-driving vehicles** to improve the **control accuracy of steering angle**.
- Our results demonstrate that using shared images increases prediction accuracy compared to other existing models.
- Our model is sensor-agnostic in regards that it uses the shared images rather than manual feature decomposition (e.g road or lane marking detection).
- We are in the process of collecting real data obtained from actual cars communicating over V2V and will perform more detailed analysis on that larger new data.

This work was done as a part of CAP5415 Computer Vision class in Fall 2018 at UCF. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the GPU used for this research.

Questions and/or Comments: rvalienter90@knights.ucf.edu, http://www.eecs.ucf.edu/NSL/