VISION-GUIDED FORECASTING - VISUAL CONTEXT FOR MULTI-HORIZON TIME SERIES FORECASTING

Eitan Kosman

Technion - Israel Institute of Technology

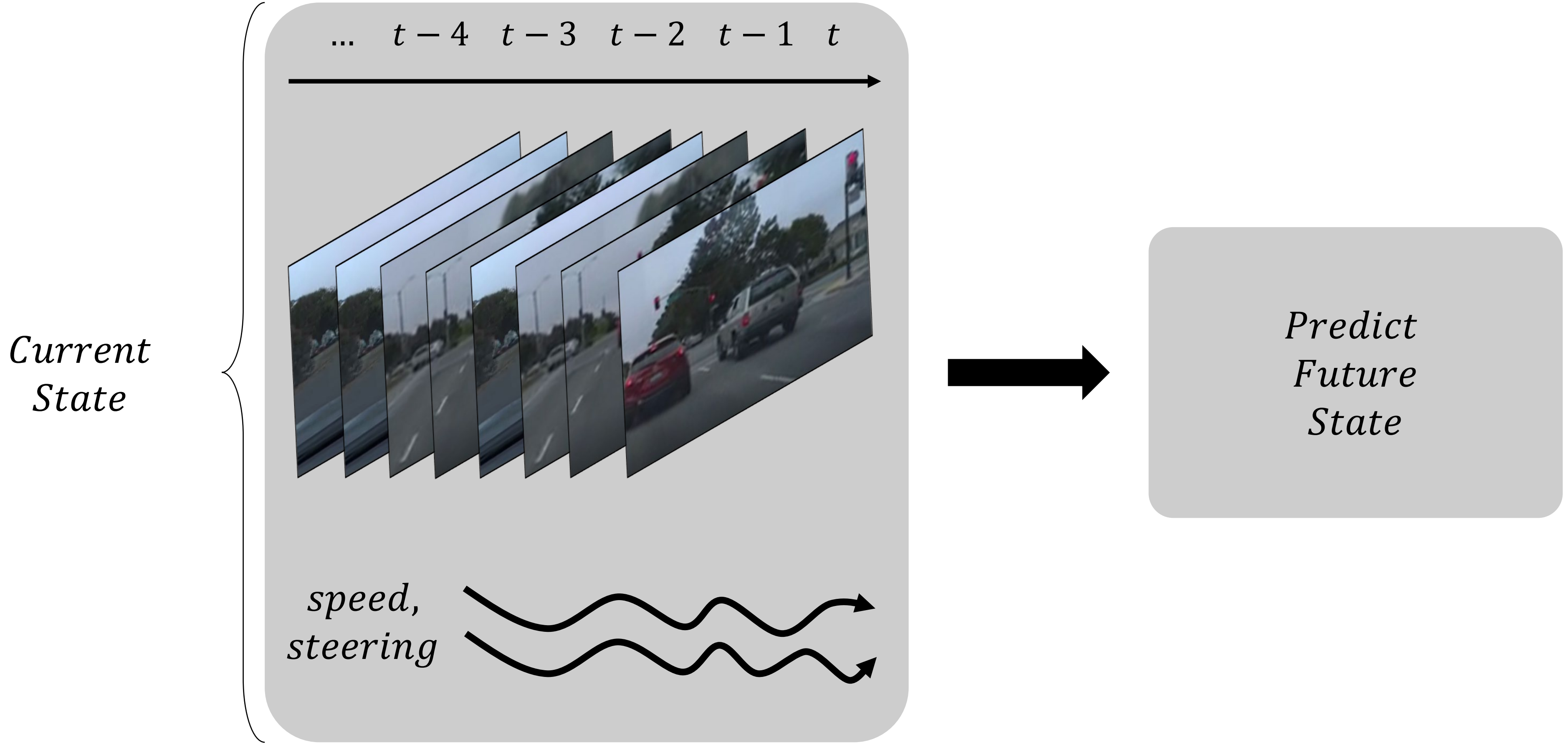
Dotan Di Castro

Bosch Center for Artificial Intelligence

2021 ICCV OCTOBER 11-17 VIRTUAL

BOSCH

Background and Motivation



Background and Motivation

Why forecasting?

In real-world, driving decisions are made by short time planning because the driver's attention is focused on the front visual view.

BOSCH

Background and Motivation

Why forecasting?

Predict anomalous situations, such as

1) Abrupt braking
2) Dangerous maneuvers

BOSCH

Background and Motivation

$\ldots \quad t-4 \quad t-3 \quad t-2 \quad t-1 \quad t$

Current State

speed, steering

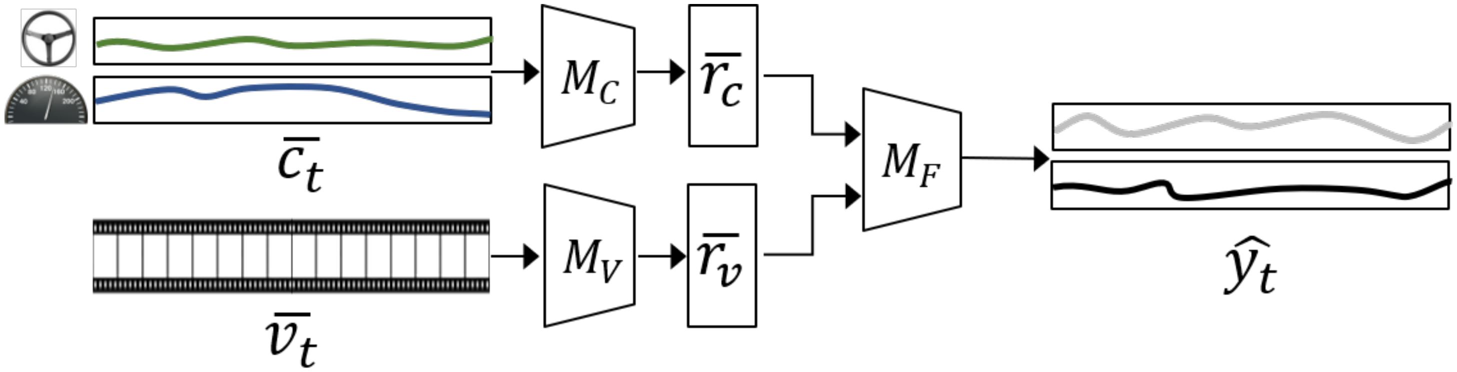
$t+\Delta \quad t+\Delta+1 \quad t+\Delta+2 \quad t+\Delta+3 \ \ldots.$

Predict Future States

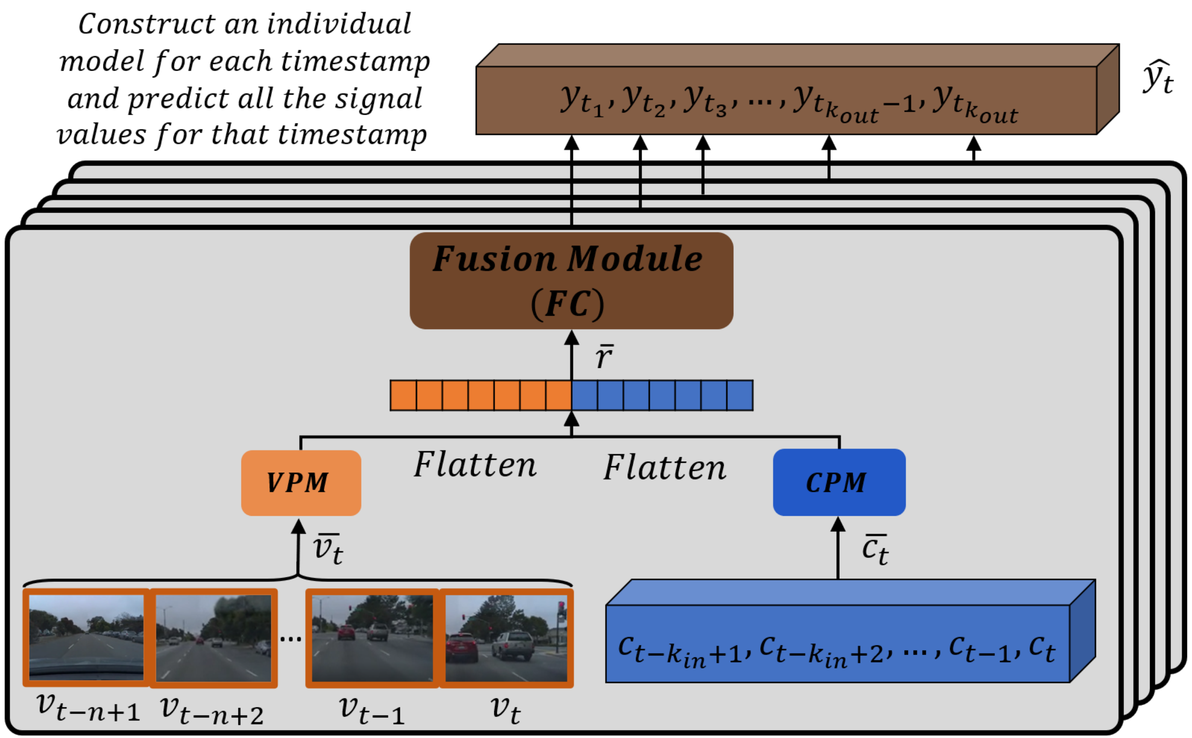
The Method

We rely on a model consisting of 3 modules:

1) Video Processing Module (VPM; denoted by M_V)
2) A Controller Area Network (CAN-Bus) Signals Processing Module (CPM; denoted by M_C)
3) A Fusion Module for predicting the output sequences from both feature representations (denoted by M_F). We develop and compare 3 different fusion modules in the following.

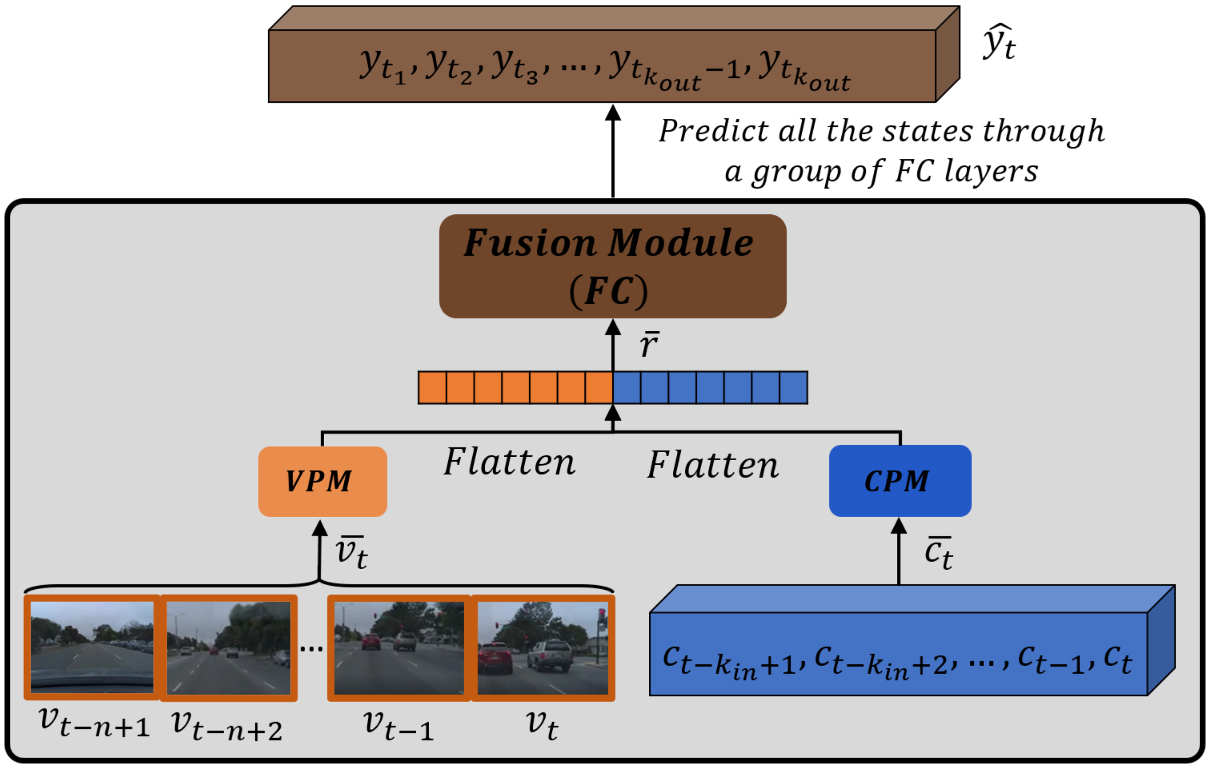


The Method - MH-IND-FC

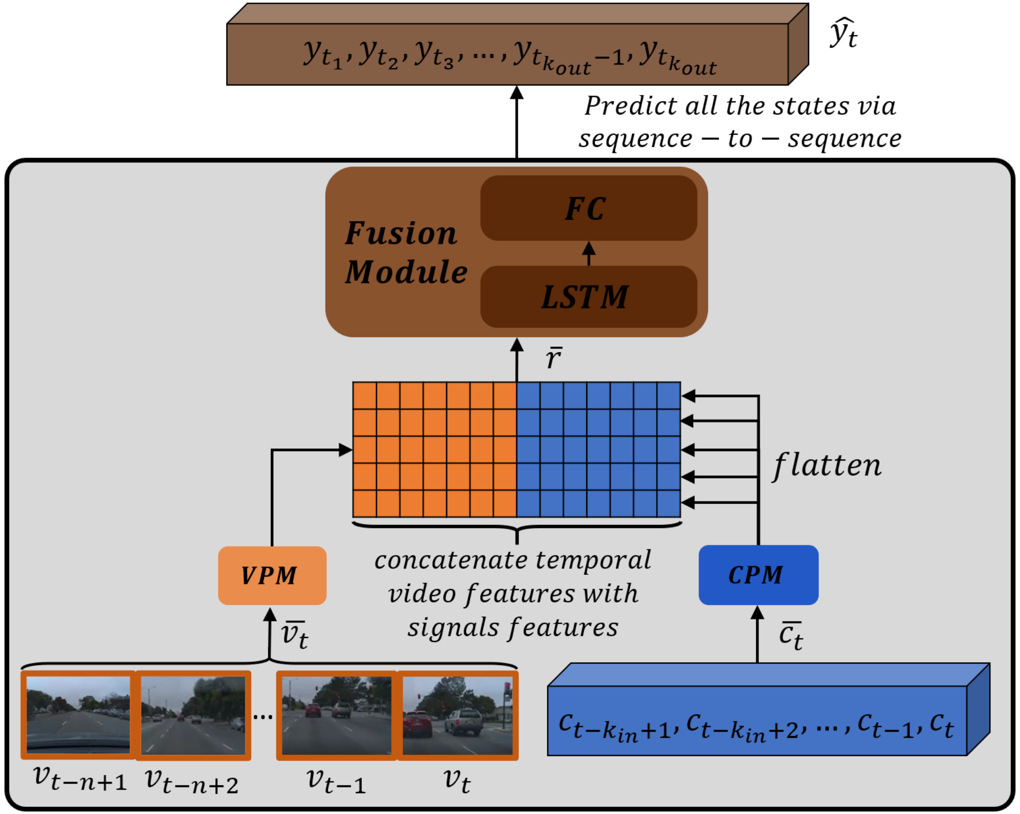


BOSCH

The Method - MH-SIM-FC



The Method - MH-SIM-LSTM



EXPERIMENTS

Experiments and Results – Evaluation Protocol

We follow the conventional evaluation protocol used in the literature which is Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

Additionally, in order to compensate for the bias of low steering angles, we propose evaluating the prediction performance on per-range basis. Given a dataset of n samples where y_i is the target steering angle of the i^{th} sample, and $\hat{y}_i$ is the predicted steering angle, the $MAE@\alpha$ is calculated by:

$$\mathrm{MAE}@\alpha = \frac{\sum_{i=1}^{n} \mathbb{1}_{|y_i| \geq \alpha} |\hat{y}_i - y_i|}{\sum_{i=1}^{n} \mathbb{1}_{|y_i| \geq \alpha}}.$$

BOSCH

Experiments and Results - Datasets

We used 2 datasets for our evaluation:

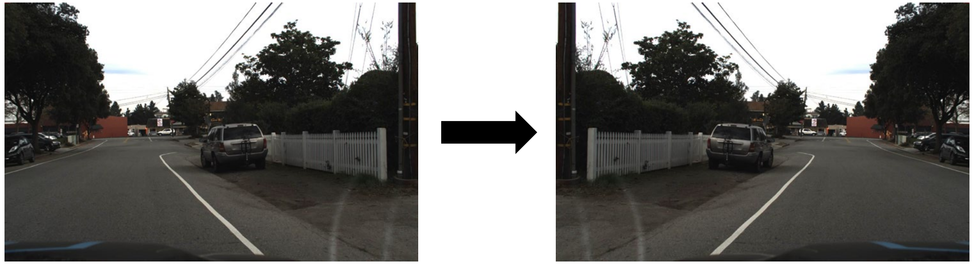
1) **Udacity Driving Dataset** which is an open-source collection of video frames along with the corresponding steering-angles, braking and throttle pressure data.
2) **Comma2k19** which captures over 33 hours of driving data. It includes frames captured by a road-facing camera, along with phone GPS, thermometers, 9-axis IMU and CAN-bus data.

More detailed information about these datasets is included in our paper.

Experiments and Results – Preprocessing & Augmentations

Data enrichment is achieved by:

1) Horizontal flips and multiplication of the corresponding steering values by -1 with probability of 0.5. This helps dealing with the skew of the steering angle values.
2) We avoid training samples with low-speed values since steering angles in these situations are not informative and hurt the training.



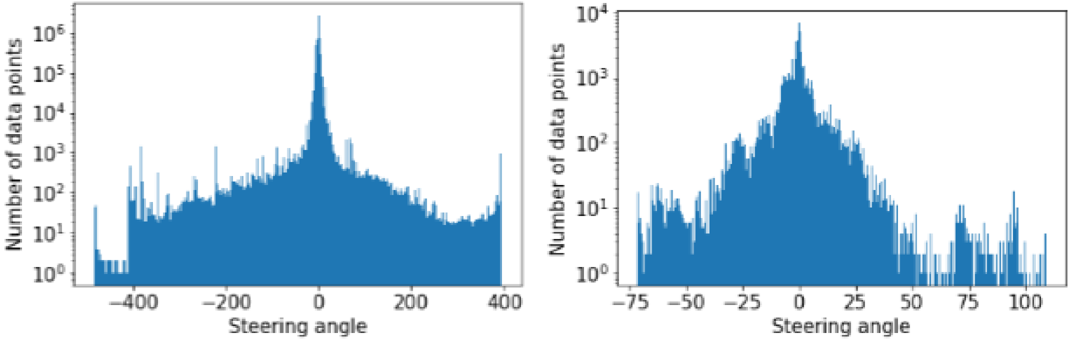


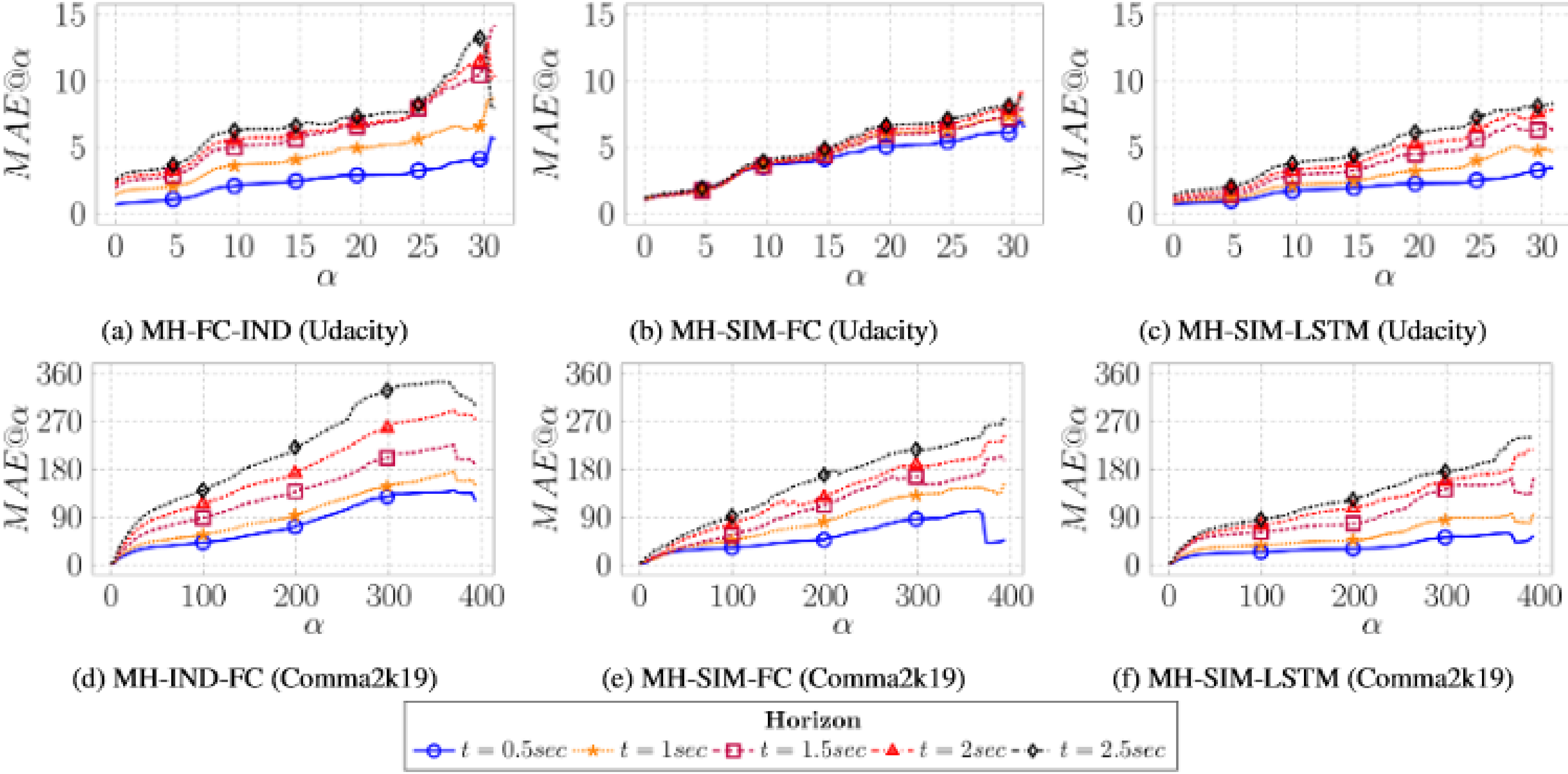
Figure 2: Histogram of the steering angles in log scale, represented as the double long-tailed distribution; Left - comma.ai. Right - Udacity.

RESULTS

ARCHITECTURES COMPARISON

		Steering				Speed			
		comma.ai		Udacity		comma.ai		Udacity	
Method	Horizon	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
MH-IND-FC	0.5	1.186	4.838	1.472	1.295	0.183	0.337	2.404	2.673
	1	1.742	6.874	1.591	2.119	0.179	0.315	2.531	2.739
	1.5	2.228	9.18	1.468	1.959	0.259	0.439	2.751	3.053
	2	2.556	10.853	1.408	1.897	0.354	0.589	2.868	3.297
	2.5	2.814	12.004	1.736	2.371	0.556	0.982	2.802	3.141
MH-SIM-FC	0.5	1.154	3.984	1.345	1.897	0.123	0.249	1.641	1.899
	1	1.78	7.003	1.368	1.952	0.182	0.34	1.713	2.019
	1.5	2.255	9.308	1.555	2.203	0.261	0.476	1.834	2.152
	2	2.552	10.568	1.592	2.233	0.352	0.617	1.929	2.294
	2.5	2.757	11.473	1.761	2.456	0.446	0.769	2.102	2.41
MH-SIM-LSTM	0.5	**1.08**	**3.444**	**0.677**	**1.394**	**0.112**	**0.212**	**0.133**	**0.313**
	1	**1.578**	**6.474**	**0.772**	**1.706**	**0.129**	**0.335**	**0.168**	**0.322**
	1.5	**1.926**	**8.378**	**0.781**	**1.733**	**0.227**	**0.415**	**0.185**	**0.357**
	2	**2.353**	**9.657**	**0.885**	**1.988**	**0.301**	**0.586**	**0.208**	**0.395**
	2.5	**2.586**	**10.442**	**1.084**	**2.235**	**0.388**	**0.539**	**0.247**	**0.462**

Table 1: Test results for prediction both steering angle and speed for the multi-horizon forecasting architectures. Values are in degrees. Lower is better. In bold are the best errors for prediction for each column and horizon where is apparent the MH-SIM-LSTM is the best architecture in our case.



(a) MH-FC-IND (Udacity)
(b) MH-SIM-FC (Udacity)
(c) MH-SIM-LSTM (Udacity)
(d) MH-IND-FC (Comma2k19)
(e) MH-SIM-FC (Comma2k19)
(f) MH-SIM-LSTM (Comma2k19)

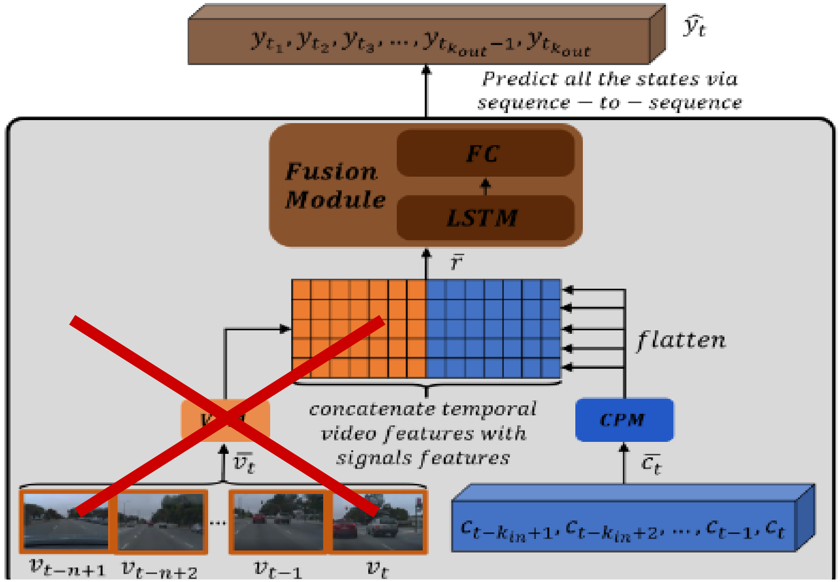
Figure 9: MAE@α for the 3 architectures for various timestamps on the two datasets. Lower values are better.

RESULTS

CONTRIBUTION OF VISION TO FORECASTING

Results - Contribution of Vision to Forecasting

- We experimented with multi-horizon forecasting without vision features in order to examine its effect on accuracy.
- Specifically, we choose the architecture that delivered the best results, i.e. MH-SIM-LSTM and omitted the VPM stem.



BOSCH

		comma.ai		Udacity	
Method	Horizon	MAE	RMSE	MAE	RMSE
W/ Vision	0.5	**1.08**	**3.444**	**0.677**	**1.394**
	1	**1.578**	**6.474**	**0.772**	**1.706**
	1.5	**1.926**	**8.378**	**0.781**	**1.733**
	2	**2.353**	**9.657**	**0.885**	**1.988**
	2.5	**2.586**	**10.442**	**1.084**	**2.235**
W/O Vision	0.5	1.806	4.689	1.458	1.458
	1	2.092	6.718	2.289	2.289
	1.5	2.506	8.858	2.528	2.528
	2	2.805	10.462	2.64	2.64
	2.5	3.019	11.562	2.919	2.919

Table 2: Accuracy comparison of the architecture *MH-SIM-LSTM* with the *VPM* stem and without using the *VPM* stem. Values are in degrees. Lower is better. In bold are the best errors for prediction for each column and horizon where is apparent the MH-SIM-LSTM with vision is the best.

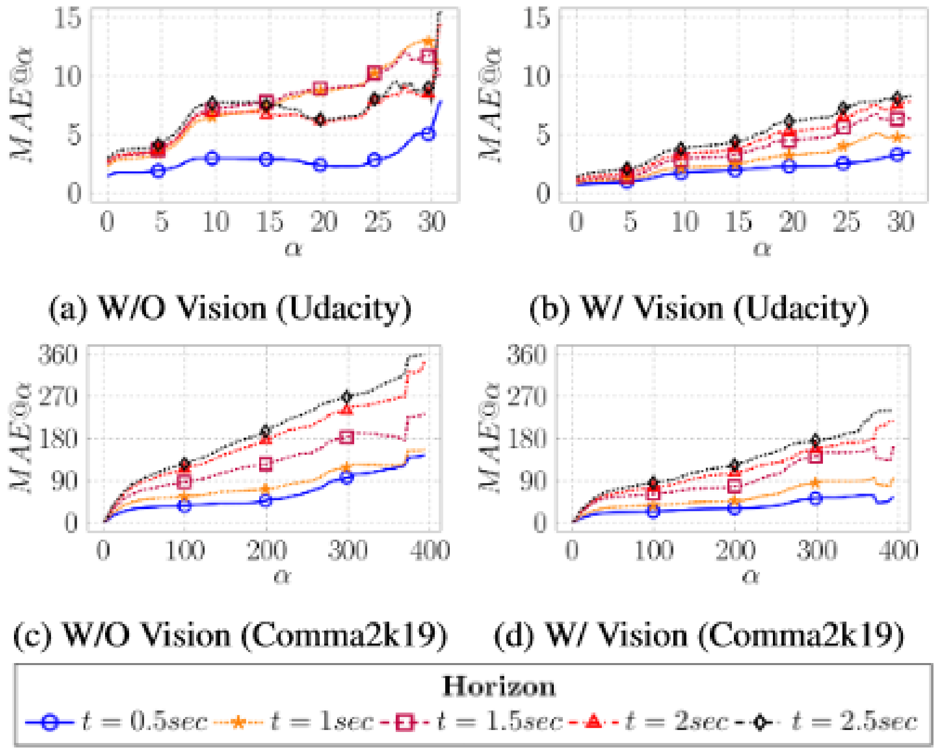


Figure 10: MAE@α for the MH-SIM-LSTM architecture, with and without using the vision stem. Lower is better. A model fed with vision achieves an error that is 56.6% and 66.9% of the error achieved by a model that doesn't use those features, on Udacity and Comma2k19 respectively.

THANK YOU