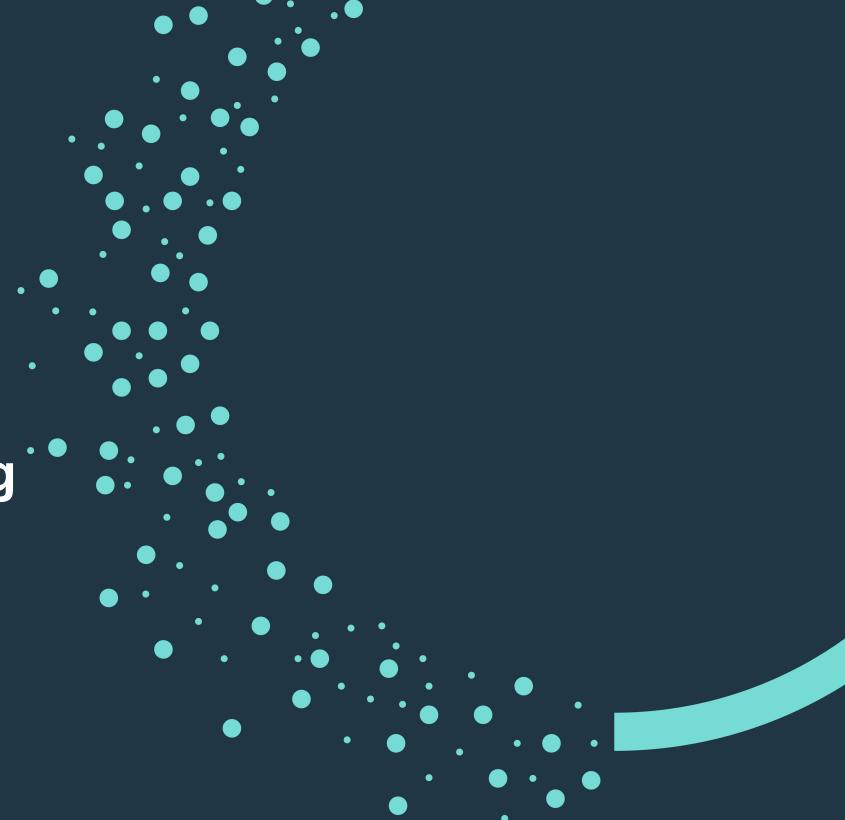




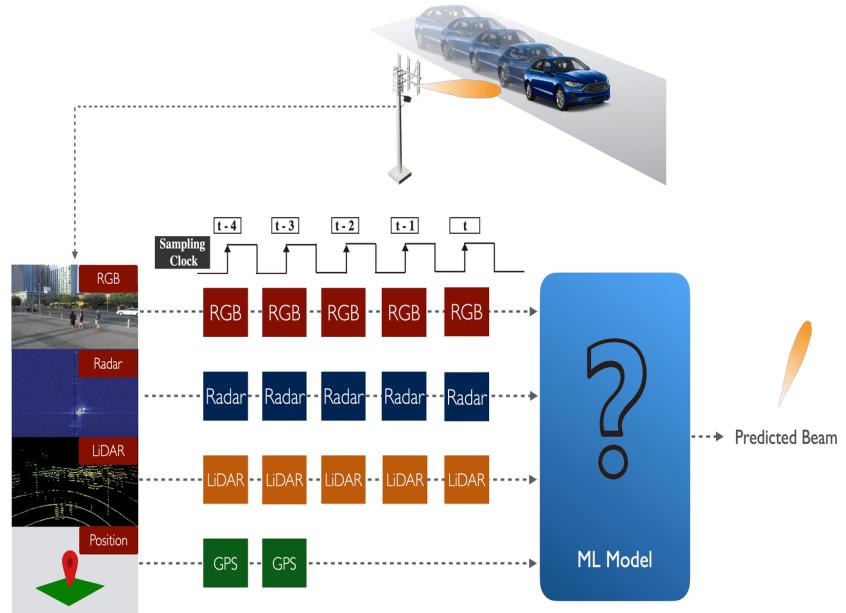
Multi-Modal Transformer Deep Learning for Sensing assisted Communications

Team TII



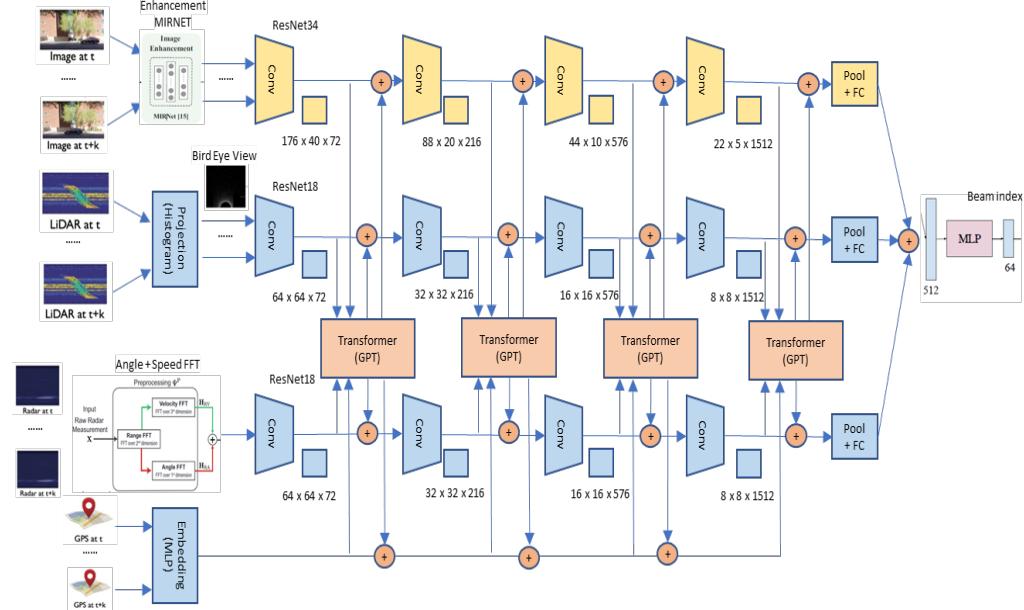
Sensing for Communications

- **Communications beyond 5G**
 - High frequency and **narrow beams**
 - Boost capacity, increase SINR, reduce energy
 - **Mobility and beam management**
 - Propagation loss, high speed, high reliability
- **Sensing-assisted** beam prediction
 - Beam prediction from **multi-modal** sensors
 - Sequential camera, LiDAR, radar, GPS data
 - Predict beam with maximum received power
 - GPS: high latency, energy, interruption issues
 - Sensors: environmental blockage, reflection

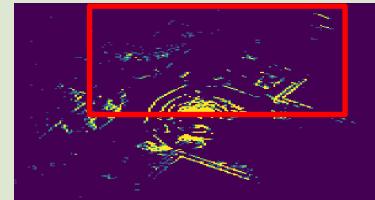
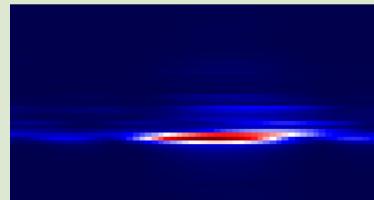
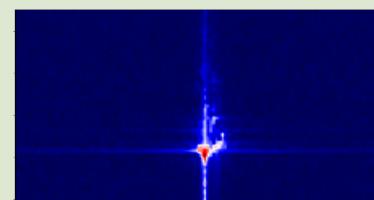
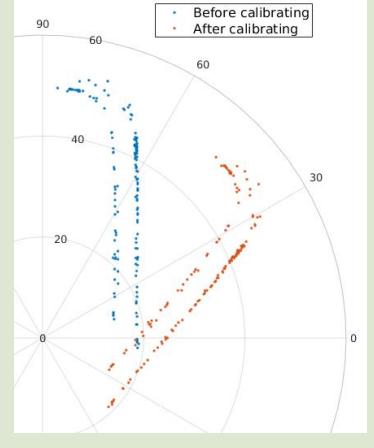


Multi-Modal Transformer Deep Learning

- **Feature extraction** with CNN
 - Abstraction of multi-modalities
 - Transfer from pretrain model
 - Projection to align dimensions
- **Feature fusion** with transformer
 - Fusion on abstraction spaces
 - Generalized to applications
 - Flexible multi-modality inputs
- **Generalized** multi-modal learning



Multi-Modal Sensor Data Processing

Image	Point-cloud	Radar	GPS
Enhance brightness Semantic segment	Project Bird Eye View Custom Field of View Filtering background	2D Fourier transform Range-angle maps Range-velocity maps	Relative coordinates Min-max normalize Calibrate view angles
 	 	 	

Comparisons

- Multiple **sequences**
 - Improve accuracy with more timestamps preprocessed data
- Multiple **modalities**
 - Improve accuracy with more modalities sensor data
- **Transformer** model
 - Fusion multi-modality data and outperform CNN, LSTM

Comparing data sequence performance

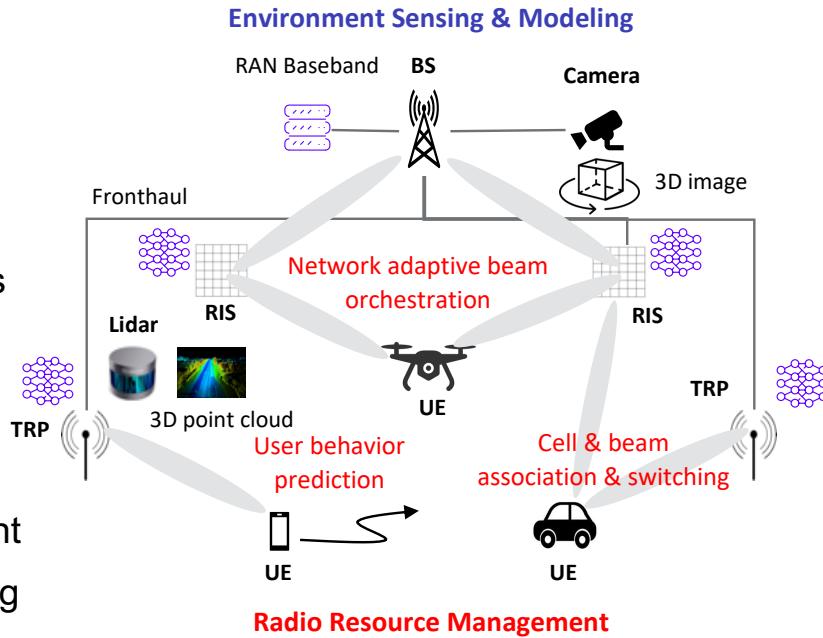
31	32	33	34	Overall	Data
0.5331	0.7173	0.7910	0.8209	0.6671	Preprocessed
0.5056	0.7160	0.7800	0.7720	0.6411	5 timesteps
0.5044	0.6988	0.7281	0.7527	0.6234	3 timesteps
0.4718	0.6222	0.6933	0.7328	0.5891	1 timestep

Compare data modality and model performance

Top-1	Top-2	Top-3	DBA	Modalities	Model
47%	53%	54%	0.62	camera, lidar, radar	Transformer
44%	48%	48%	0.57	camera, lidar	Transformer
49%	50%	52%	0.55	radar	CNN
5%	8%	9%	0.21	GPS	LSTM

Benefits

- Contribution
 - **Transformer** deep learning for beam prediction
 - **Preprocess** sequential multi-modal sensor data
 - **Generalize** to various scenario and applications
- Advantage
 - **Tailorable** model size, data sequences, modalities
 - **Robust** in extreme environments: fog, rain, cloud
 - **Diverse** devices and sensors in wireless network
- Applications
 - Beam, power, resource, load, mobility management
 - Sensing, localization, trajectory prediction, planning
 - Collaborative control vehicle, robotics, road traffic





Thank you

