Appendix: Additional Experimental Results

Optimal-Horizon Social Robot Navigation in Heterogeneous Crowds

Anonymous Authors

1 Appendix: Additional Experimental Results

1.1 Spatio-Temporal Transformer Training Results

This section presents comprehensive training results for the Spatio-Temporal Transformer used for pedestrian cooperation prediction. The Spatio-Temporal Transformer model demonstrates strong performance in classifying cooperative vs. non-cooperative pedestrian behaviors across diverse crowd scenarios. Table 1 shows the detailed model configuration and training setup. The training performance evolution is presented in Table 2. Error analysis with limitations is presented in Table 3.

Table 1: Spatio-Temporal Transformer Model Configuration

Parameter	Value	Description
Model Architecture		
d_{model}	128	Model embedding dimension
n_{head_nodes}	8	Node attention heads
n_{head_edges}	4	Edge attention heads
n_{layer}	3	Transformer layers
d_{ff}	512	Feedforward dimension
Dropout	0.1	Dropout rate
Training Setup		
Epochs	2000	Training epochs
Batch Size	4096	Batch size
Learning Rate	0.001	Learning rate
Sequence Length	8	Input sequence length

Model architecture parameters and training configuration for the Spatio-Temporal Transformer.

1.2 Environment and Training Configuration

This section presents comprehensive configuration parameters for our framework. Table 4 shows the environment setup parameters. Table 5 presents the MPC-CBF configuration parameters. Table 6 details the reward function parameters. Table 7 shows the training configuration and hyperparameters used for the RL model.

Table 2: Training Performance Evolution

Epoch	Train Acc	Test Acc	ROC AUC
940	90.62%	89.99%	97.69%
1000	91.07%	90.40%	97.92%
1200	92.29%	91.74%	98.35%
1400	93.15%	92.50%	98.73%
1560	93.71%	93.14%	98.91%

Training performance evolution showing accuracy and ROC AUC improvements over epochs.

Table 3: Error Analysis and Limitations

Error Type	Frequency	Description
False Positives (Coop)	2.8%	Misclassified as cooperative
False Negatives (Coop)	5.3%	Missed cooperative behavior
High Density Scenarios	+15% error	Crowded environments
Partial Observation	+8% error	Limited visibility
Long-term Prediction	Limited	Beyond 8 steps

Error analysis showing common failure modes and limitations of the cooperation prediction module.

Table 4: Environment Configuration Parameters

Parameter	Value	Description
Simulation Environment	CrowdNav	Extended 2D pedestrian simulation with limited sensing range
Robot Sensing Range	5m	Circular sensing range with 360° field of view
Human Population	20 pedestrians	Varying cooperation ratios (0, 5, 10 cooperative pedestrians)
Robot Kinematics	Holonomic	Maximum velocity of 1.0 m/s in both x and y directions
Safety Distance	0.5m	Minimum distance around robot and other agents
Prediction Horizon	[1, 2, 3, 4, 5, 6, 7, 8]	Dynamic selection from available horizon steps

Environment configuration parameters for the CrowdNav simulation environment.

Table 5: MPC-CBF Configuration Parameters

Parameter	Value	Description
State Cost Matrix Q	[[15, 0], [0, 15]]	Weighting matrix for state variables
Control Cost Matrix R	[1, 1]	Weighting matrix for control inputs
Terminal Cost Matrix P	[[3, 0], [0, 3]]	Weighting matrix for terminal state

MPC-CBF configuration parameters for trajectory optimization and safety constraints.

Table 6: Reward Function Parameters

Parameter	Value	Description
Terminal Reward		
Goal Reward	+10	Reward for reaching goal
Collision Penalty	-20	Penalty for collision or timeout
Potential Reward		
λ_{pot}	2	Potential field weight
Kinematic Reward		
λ_r	5	Rotation penalty weight
λ_v	2	Velocity penalty weight
Prediction Horizon Reward		
λ_h	0.1	Horizon reward weight
h_{max}	8	Maximum horizon
Visibility Social Reward		
λ_{high}	0.05	High cooperation weight
λ_{low}	0.01	Low cooperation weight
η_{high}	1.2	High cooperation exponent
η_{low}	0.8	Low cooperation exponent
Cooperation Threshold	0.5	Threshold for cooperation classification

Reward function parameters for balancing safety, efficiency, and social awareness.

Table 7: Our RL Training Configuration

Parameter	Value	Description
PPO Algorithm		
Algorithm	PPO	Proximal Policy Optimization
Learning Rate	4e-5	PPO learning rate with RMSprop optimizer
Clip Parameter	0.2	PPO clipping parameter
Value Loss Coef	0.5	Value function loss weight
Entropy Coef	0.02	Entropy regularization
Discount Factor	0.99	For reward calculation
Training Setup		
Total Steps	30M	Environment interaction steps
Parallel Processes	64	Number of parallel environments
Batch Size	64	With 64 parallel processes
Mini-batch Size	8	PPO mini-batch size
Save Interval	1M	Model checkpoint frequency
Action Space	8	Dynamic prediction horizon options (1-8)

Training configuration and hyperparameters for our RL-based dynamic horizon decision mechanism.