



Improved efficiency based on learned saccade and continuous scene reconstruction from foveated sampling

Jiayang Liu¹, Yiming Bu¹, Daniel Ts'o², Qinru Qiu¹

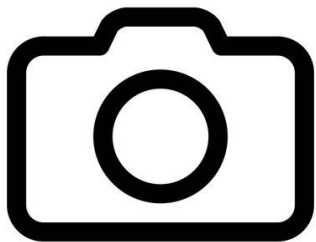
Syracuse University¹, SUNY Upstate Medical University²



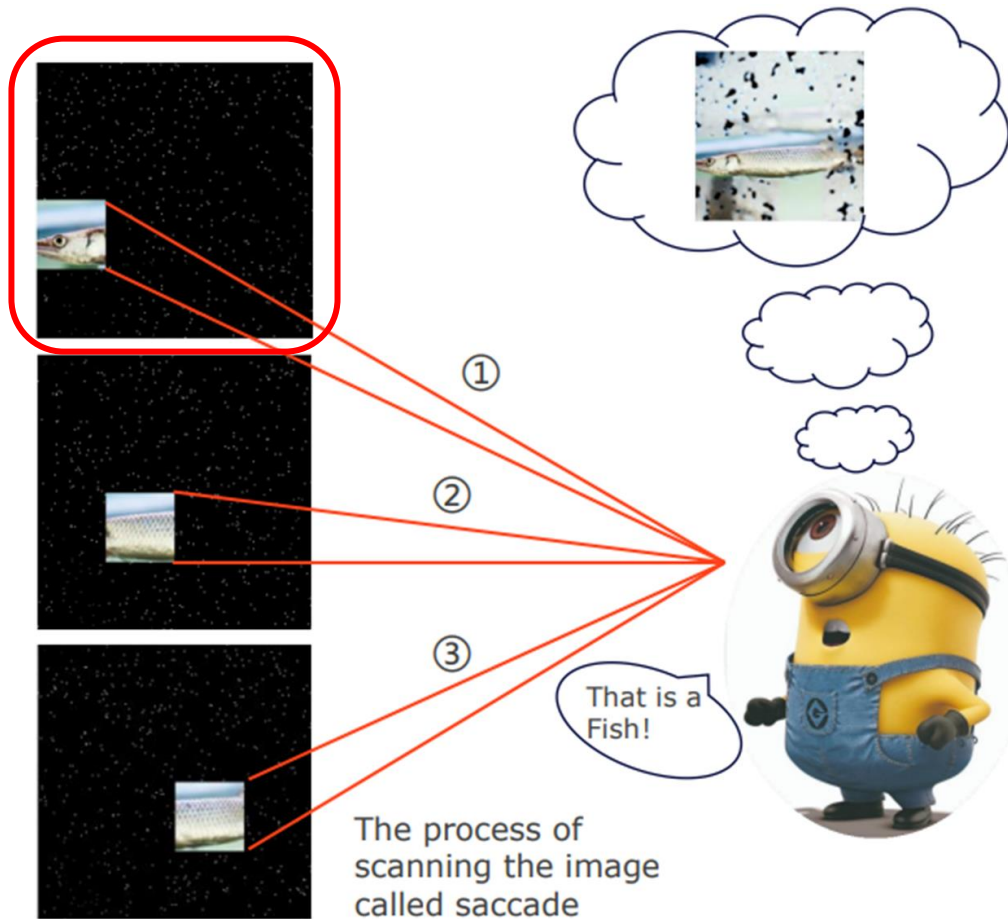
Motivation

High-quality images offer:

- Detailed object/scene presentation.
- Improved accuracy for detection and object
- Increased computational workload.
- Lower camera frame rates
- Raised data communication demands.



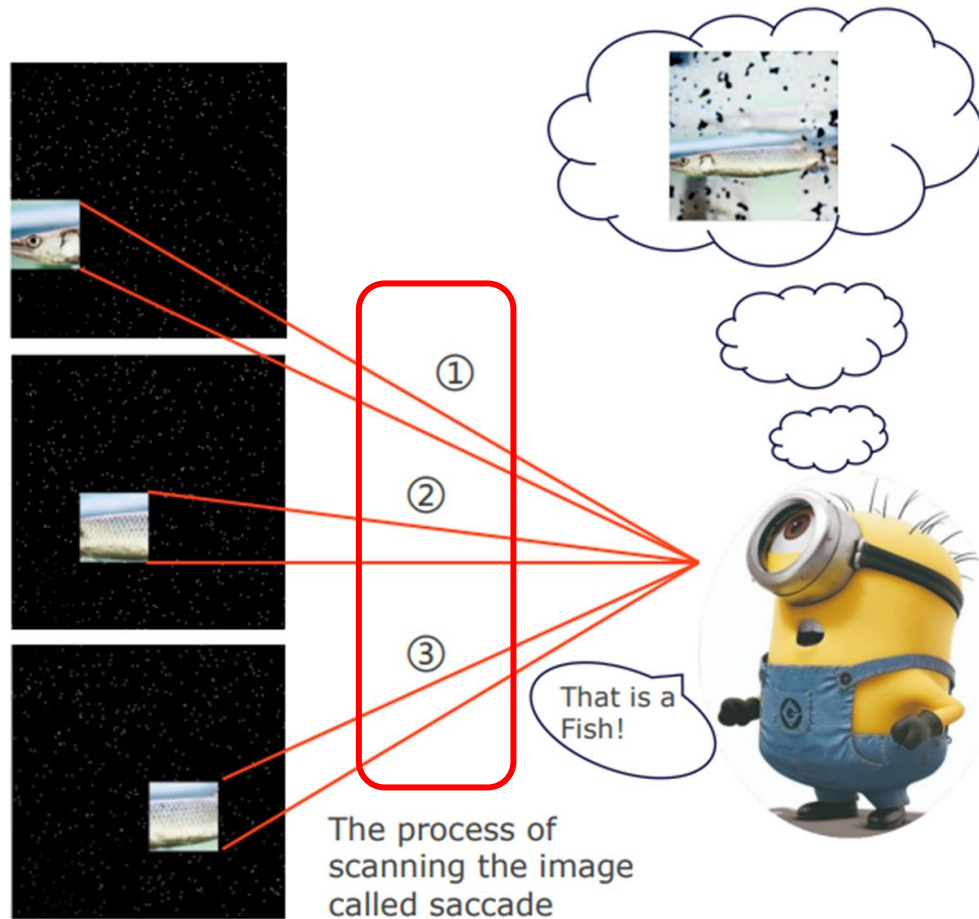
Foveal-Peripheral View



Foveal-peripheral vision:

- The human eye has a visual field of about 140 degrees.
- The central 1 degree, known as the fovea, observes high-resolution details.
- The rest of the field is sampled at lower resolution towards the periphery.

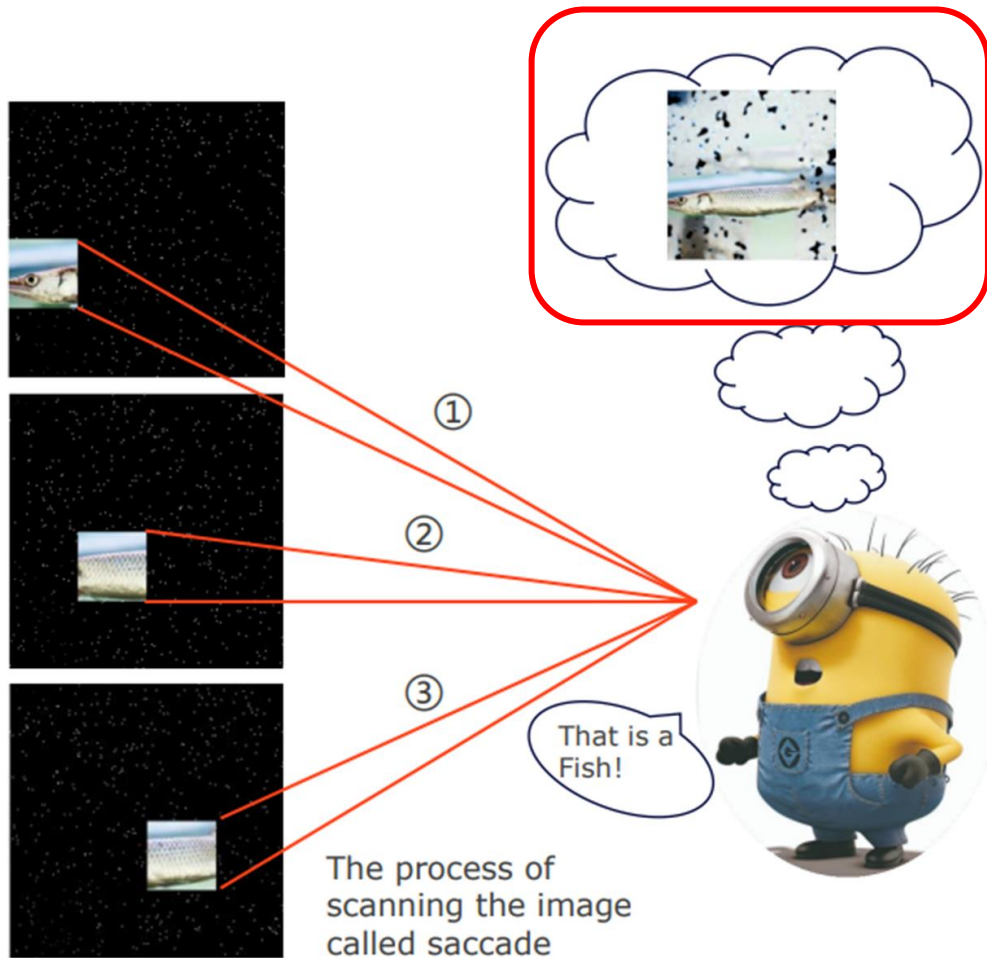
Saccade & Planning



Saccade & Planning:

- The process of human eye rapidly shifts its attention from one fixation to another called saccade.
- Saccade planning helps the brain process visual stimuli more efficiently.

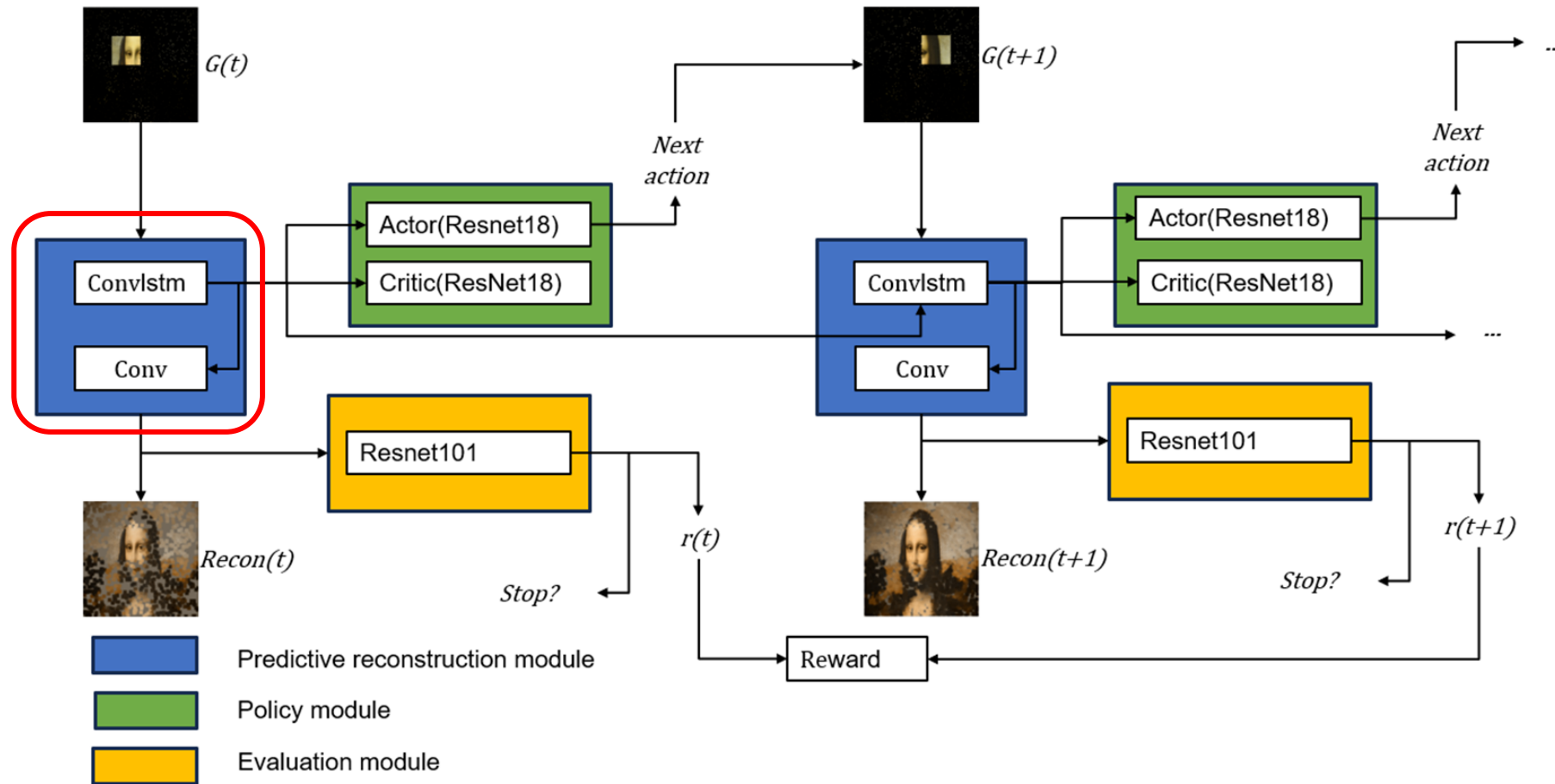
Filling-in Phenomomonon



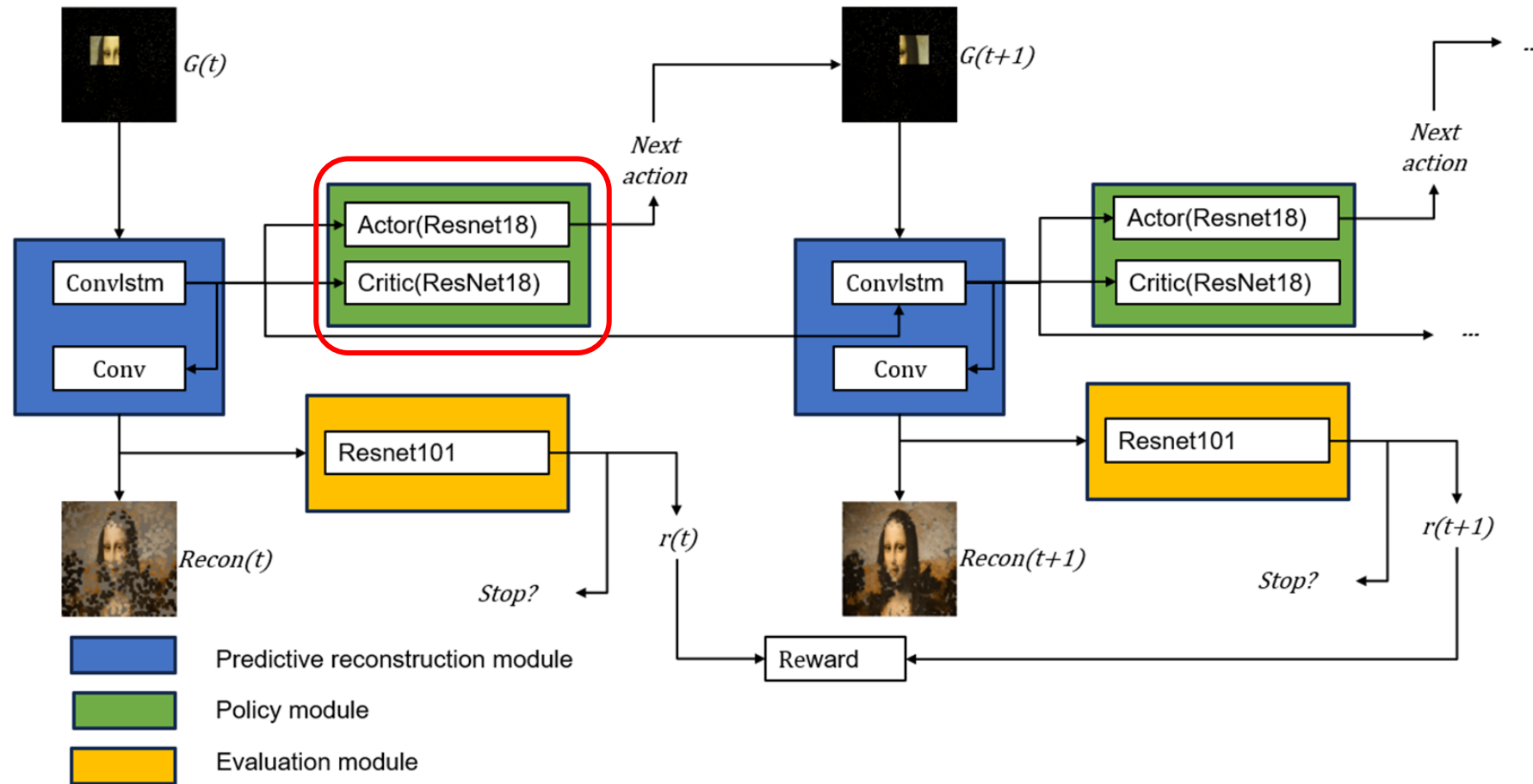
The brain's filling-in capability allows us:

- To complete missing information across the physiological blind spot.
- To yield visual perception that is continuous and seamless.

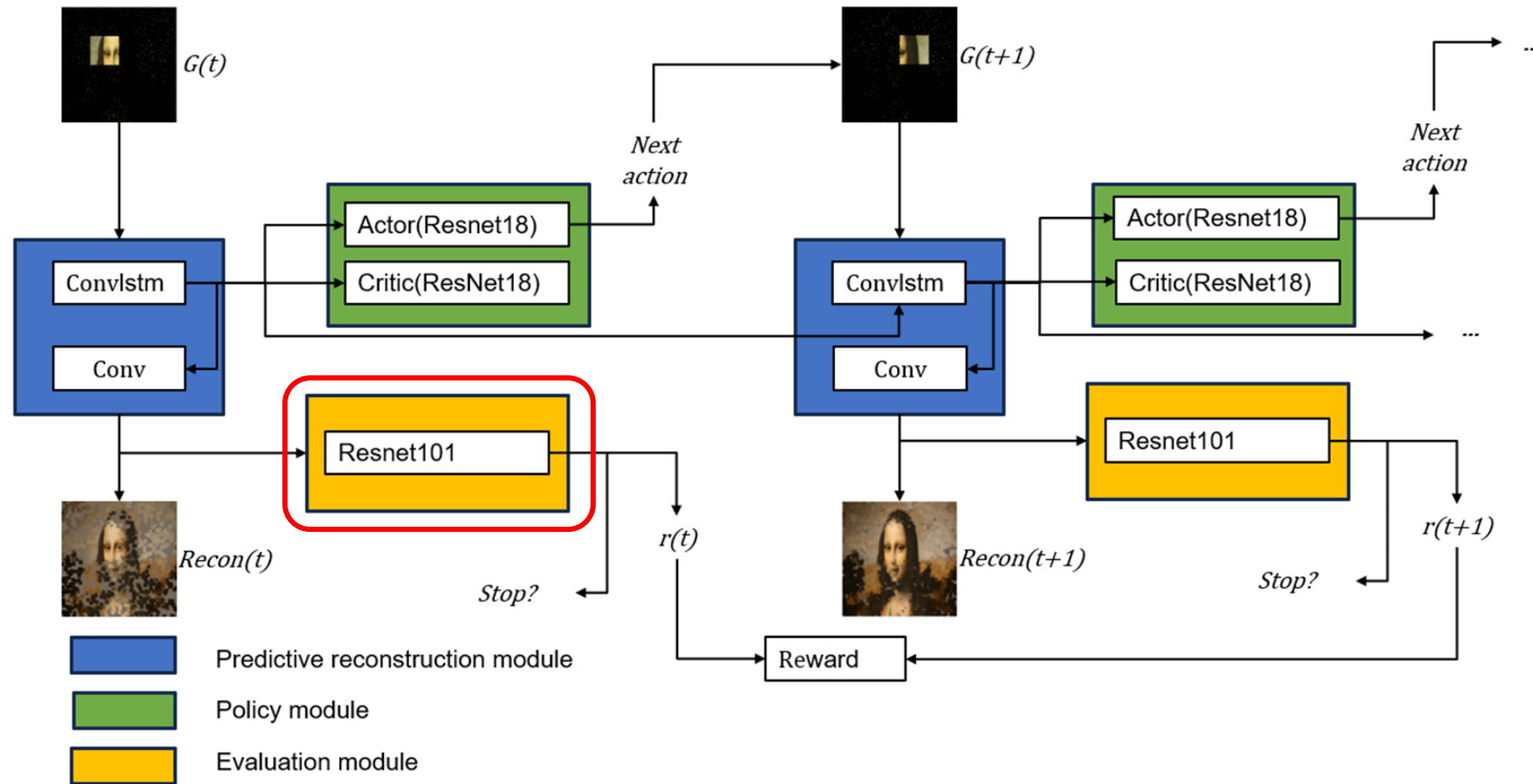
Proposed Architecture



Proposed Architecture

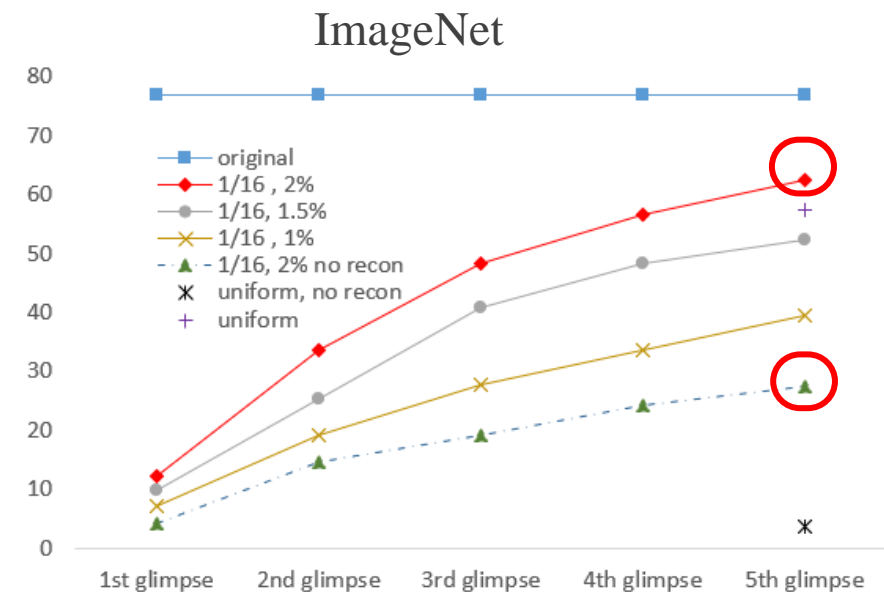
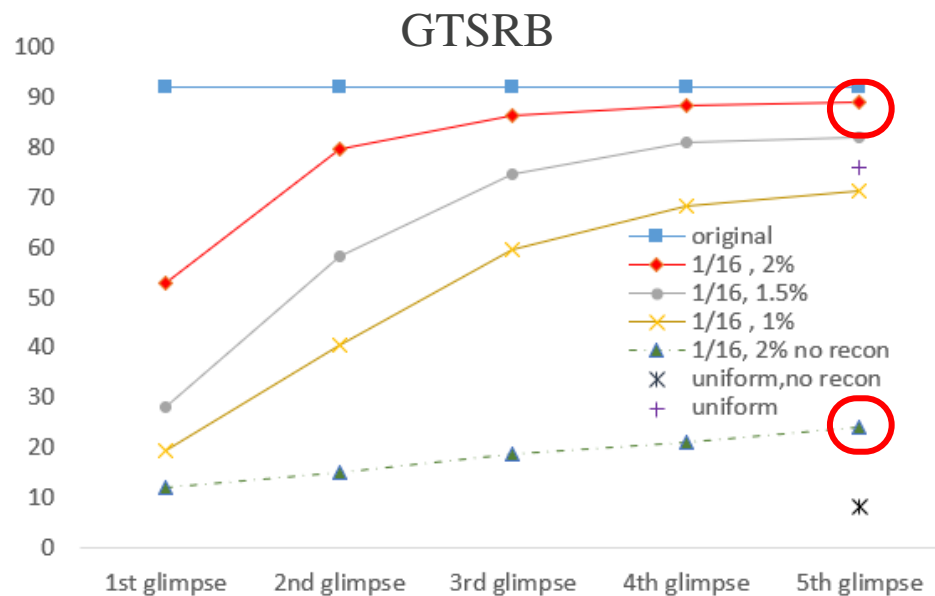


Proposed Architecture



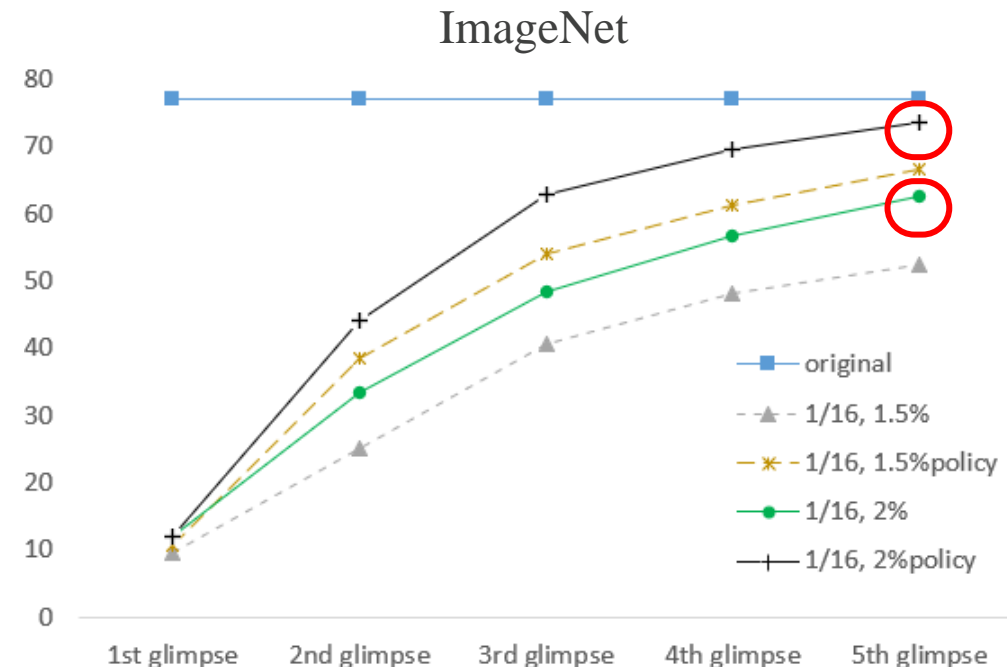
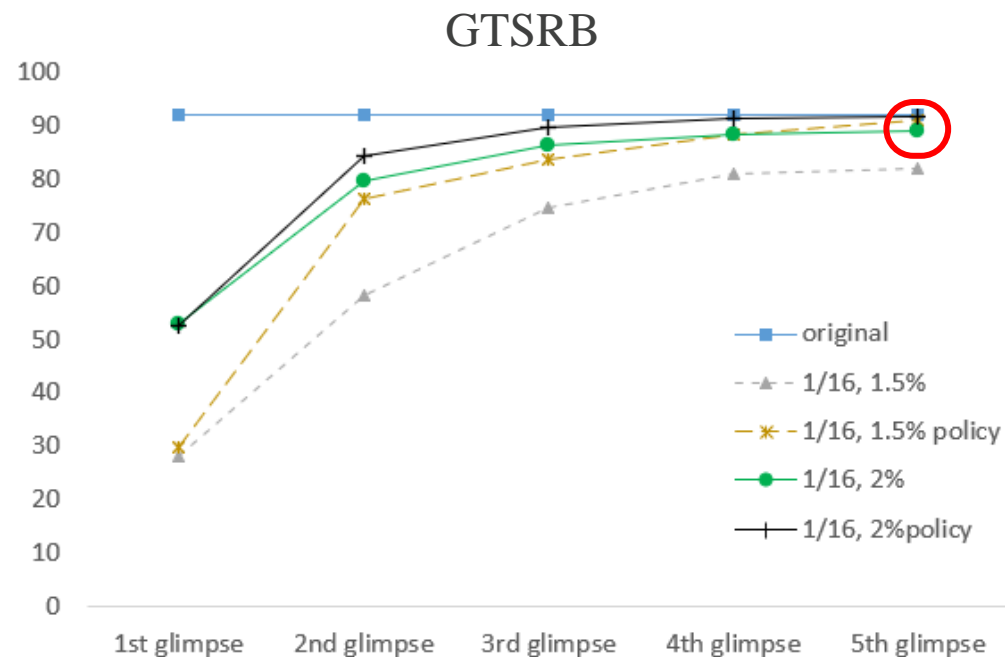
Predictive Reconstruction Performance

- Random saccades are used in this experiment.
- With predictive reconstruction model, the classification accuracy increase from 27.4% to 62.6% after 5 glimpses for ImageNet dataset. GTSRB dataset shows similar result.



Saccade Controller Performance

- Compared with random saccades, learned saccades improve the classification accuracy by 11.1% for ImageNet and 1.9% for GTSRB when the peripheral sampling proportion is 2%.



Comparison with Related Works

- In bio-FCG, the next saccade fixation is scheduled based on the classification of the previous frame, which undermines its reliability.
- Saccader requires more glimpses and more pixels to reach comparable result.
- Both Bio-FCG and Saccader have re-trained or fine-tuned their classifier using the foveal-peripheral view, while our framework did not.

Model	Test accuracy[%]	Pixel usage rate[%]	Average glimpses
Bio-FCG	65.17	27	2
Saccader	70.0	70	7
Ours 7×7	70.9	28	7
Ours 4×4	73.7	41	5
Ours 7×7 early stop	66.4	22	5.4
Ours 4×4 early stop	69.4	30	3.5

Data Efficient Policy Training

- Training the controller using data from only half of the classes in the training set performs almost equally well as training the model using the entire training set.
- Our saccade strategy is highly transferable.

training classes	acc on all classes[%]	acc on training classes[%]	SSIM
Random	62.6	-	0.85
SSIM	66.3	-	0.90
100	70.3	77.2	0.87
200	71.5	75.3	0.87
300	72.4	74.7	0.87
400	73.1	74.2	0.88
500	73.6	73.8	0.88
1000	73.7	-	0.88

Thanks for listening