Monte-Carlo Siamese Policy on Actor for Satellite Image Super Resolution

CVPR-EarthVision 2020

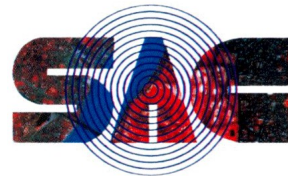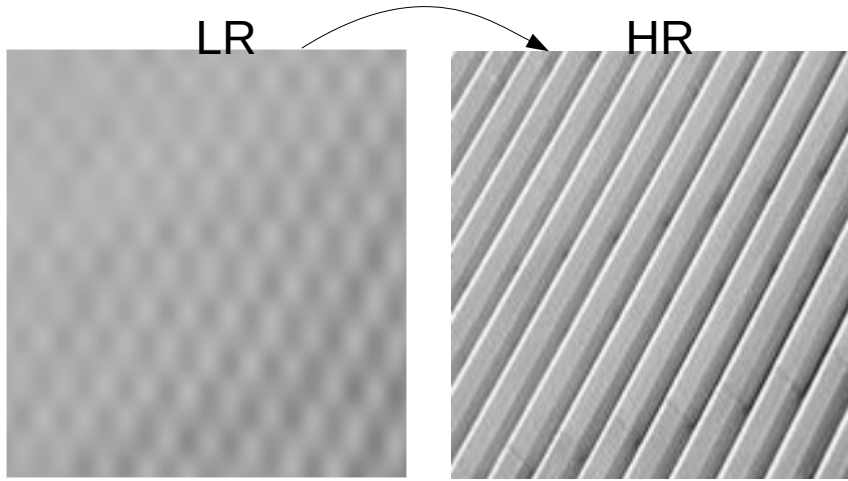**Litu Rout**   Saumyaa Shah   S Manthira Moorthi   Debajyoti Dhar

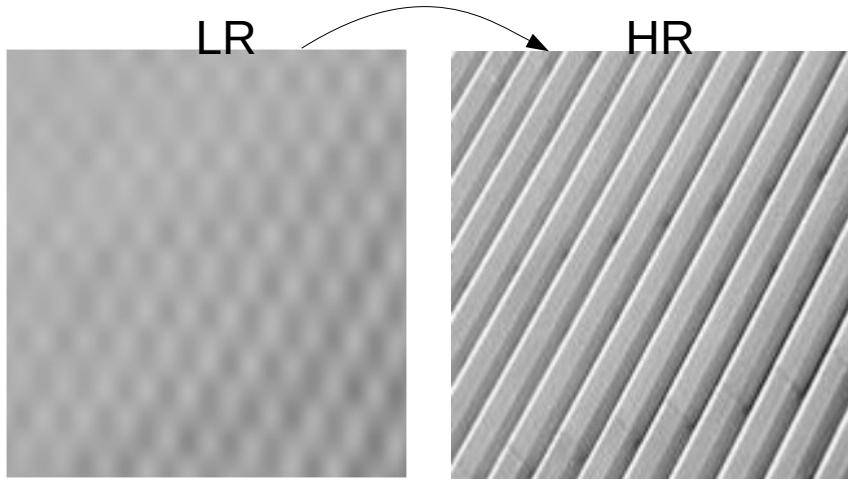Space Applications Centre

Indian Space Research Organisation

Single Image Super Resolution

LR → HR

Popular Approaches
- Supervised Learning

- Adversarial Learning

Single Image Super Resolution

LR → HR

Popular Approaches

- Supervised Learning

- Adversarial Learning

What about RL?

- Significant progress in complex environments.

Single Image Super Resolution

LR   HR

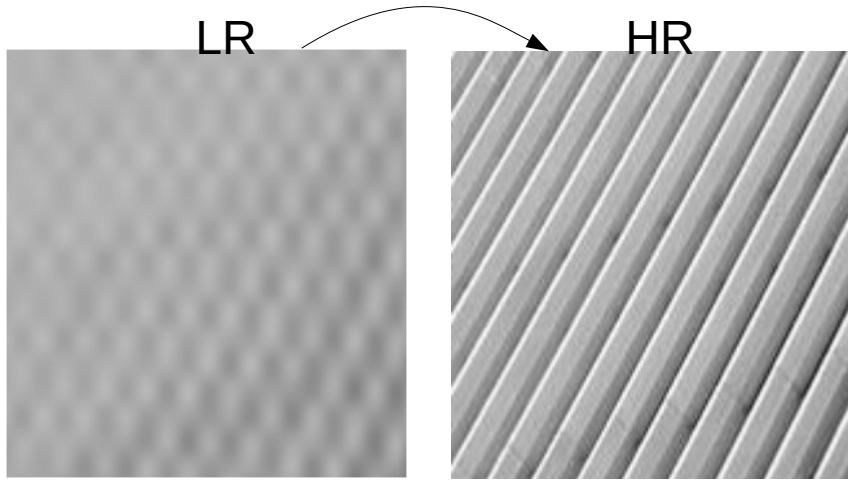## Popular Approaches

- Supervised Learning

- Adversarial Learning

What about RL?

- Significant progress in complex environments.

Single Image Super Resolution

LR

HR

Can we use RL?
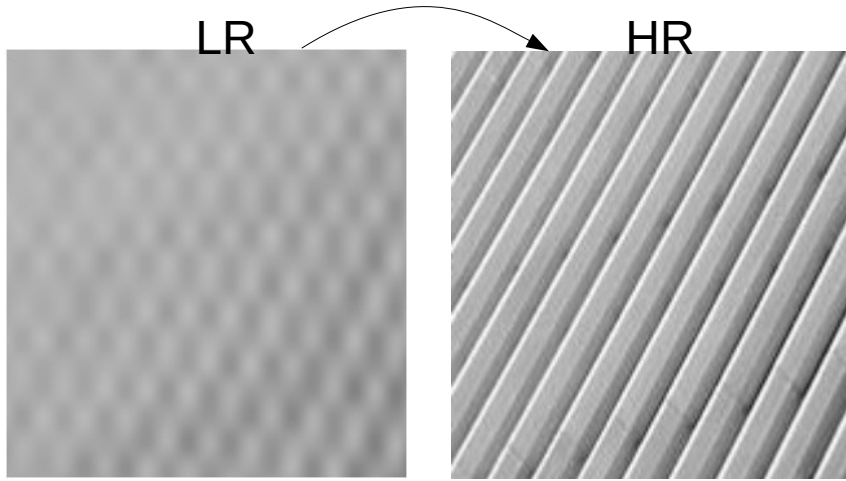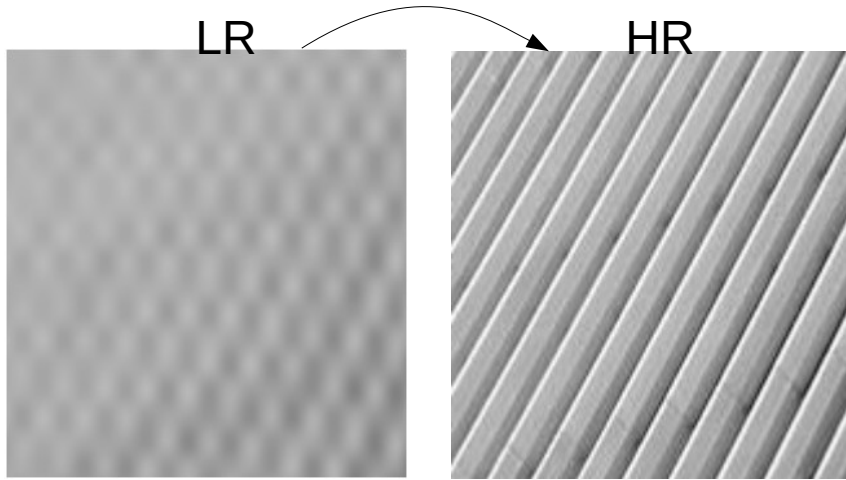
Popular Approaches

- Supervised Learning

- Adversarial Learning

What about RL?

- Significant progress in complex environments.

Single Image Super Resolution

LR HR

Popular Approaches

- Supervised Learning

- Adversarial Learning

What about RL?

Can we use RL?

- Significant progress in complex environments.

- Yes, but there is a cache.

- A straightforward implementation of RL is not adequate.

- Action variables are not fully known in super resolution environment.

Single Image Super Resolution

LR → HR

Popular Approaches

- Supervised Learning

- Adversarial Learning

What about RL?

- Significant progress in complex environments.
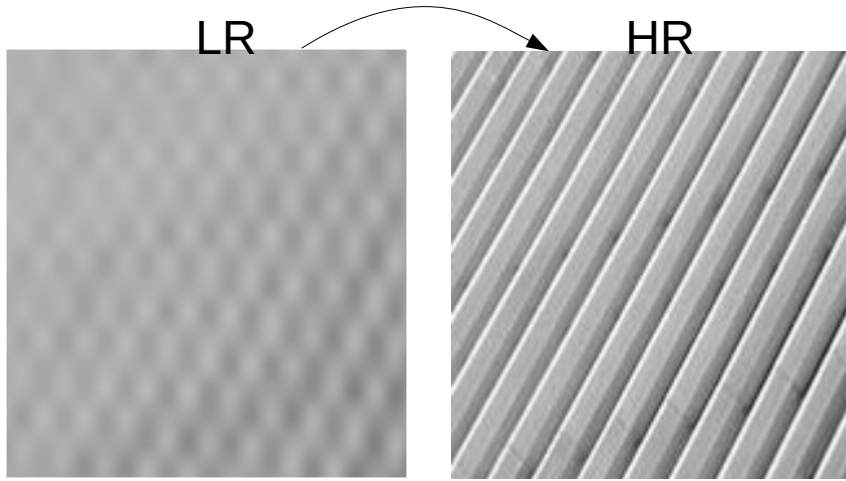
Can we use RL?

- Yes, but there is a catch.

- A straightforward implementation of RL is not adequate.

- Action variables are not fully known in super resolution environment.

Single Image Super Resolution

LR   →   HR

Popular Approaches

- Supervised Learning

- Adversarial Learning

What about RL?

- Significant progress in complex environments.
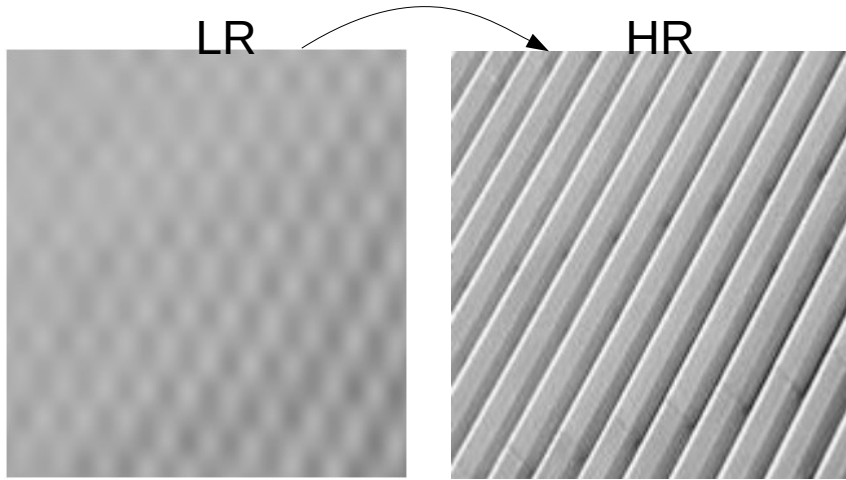
Can we use RL?

- Yes, but there is a catch.

- A straightforward implementation of RL is not adequate.

- Action variables are not fully known in super resolution environment.

Single Image Super Resolution

LR → HR

Popular Approaches

- Supervised Learning

- Adversarial Learning

What about RL?

Can we use RL?

- Significant progress in complex environments.

- Yes, but there is a catch.

- A straightforward implementation of RL is not adequate.

- Action variables are not fully known in super resolution environment.

One way is to parameterize action variables by matrices and train our policy network using Monte-Carlo sampling.

Single Image Super Resolution

LR → HR

Popular Approaches

- Supervised Learning

- Adversarial Learning

What about RL?

- Significant progress in complex environments.

Can we use RL?

- Yes, but there is a catch.

- A straightforward implementation of RL is not adequate.

- Action variables are not fully known in super resolution environment.

> One way is to parameterize action variables by matrices and train our policy network using Monte-Carlo sampling.

Grid World

Initial state

Goal state

Grid World

Initial state

Goal state
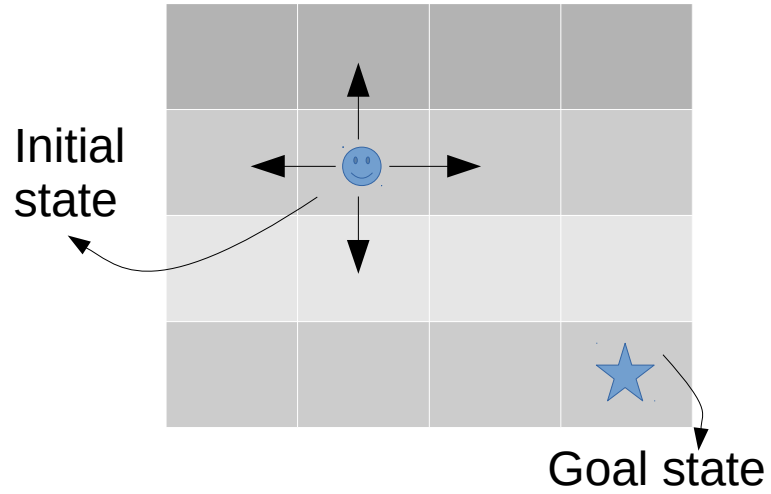
Known Action Variables

← Left

→ Right

↑ Up

↓ Down

Grid World
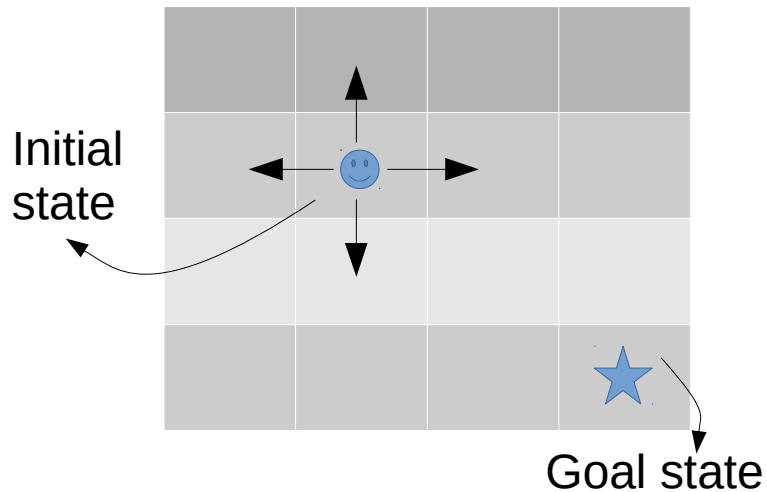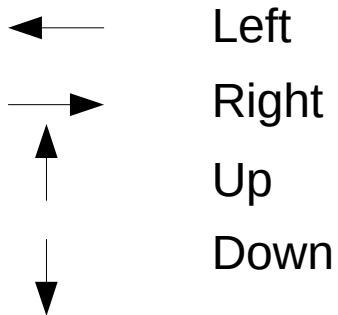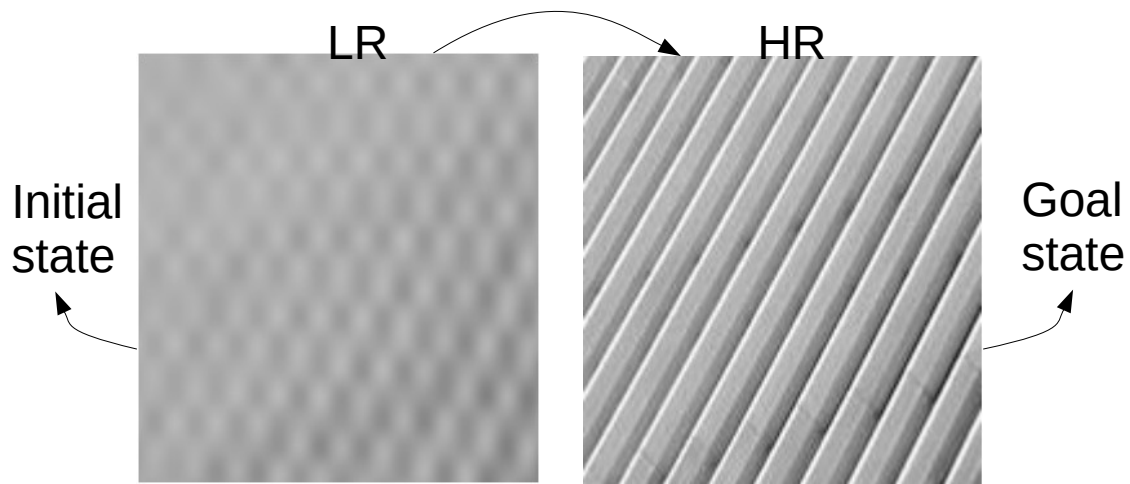
Initial state

Goal state

Up

Down

Known Action Variables

←——— Left

——→ Right

↑ Up

↓ Down

Super Resolution

LR   HR

Initial state

Goal state

Unknown Action Variables

Parameterized by matrices

Could represent
horizontal edge detection
or vertical edge detection

Grid World

Initial state

Goal state

Known Action Variables

← Left

→ Right

↑ Up

↓ Down

Super Resolution

LR

HR

Initial state

Goal state

Unknown Action Variables

Parameterize by matrices

Could represent
horizontal edge detection
or vertical edge detection

Grid World

Initial state

Goal state

Known Action Variables

← Left

→ Right

↑ Up

↓ Down

Super Resolution

LR

HR

Initial state

Goal state

Unknown Action Variables

- Parameterize by matrices

Could represent
horizontal edge detection
or vertical edge detection

Grid World

Super Resolution

Initial state

Goal state

LR

HR

Initial state

Goal state

Known Action Variables

← Left

→ Right
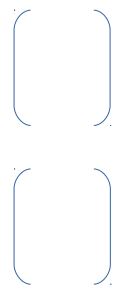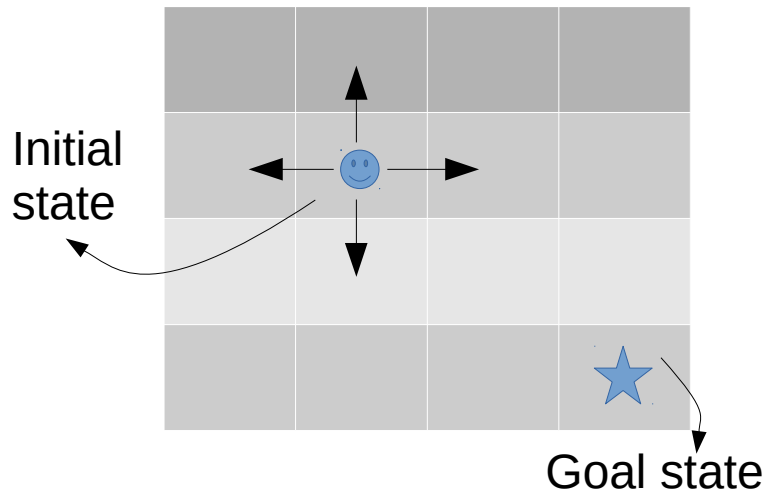
↑ Up

↓ Down
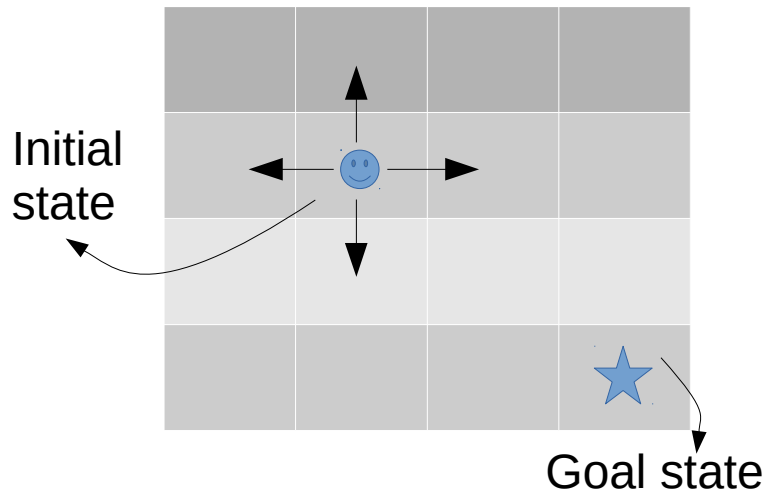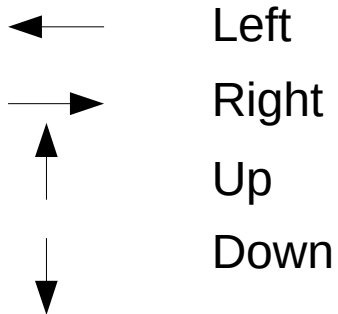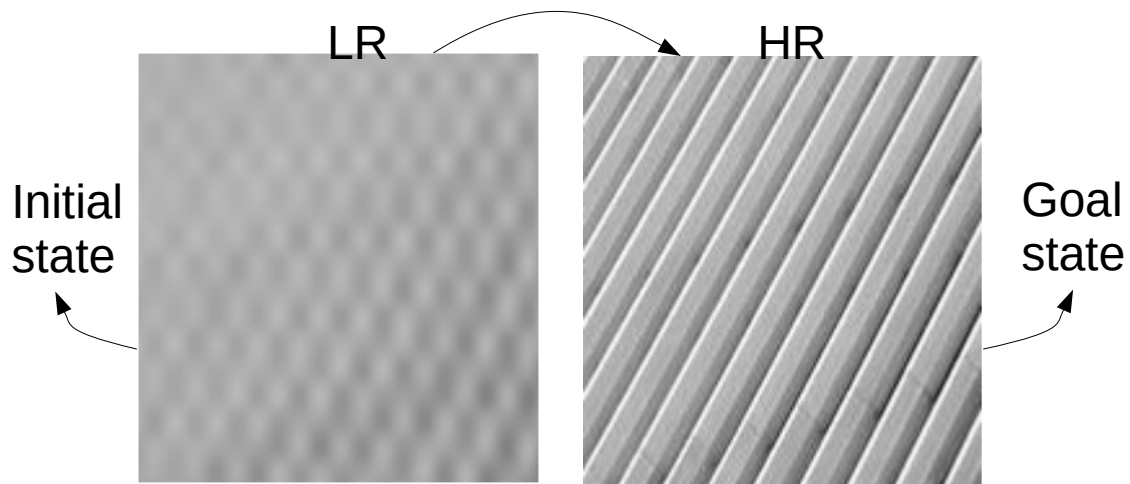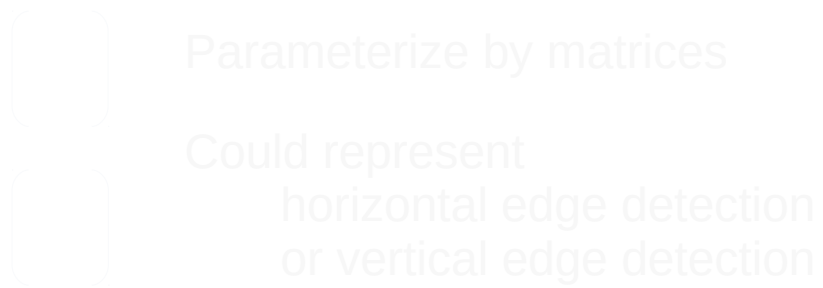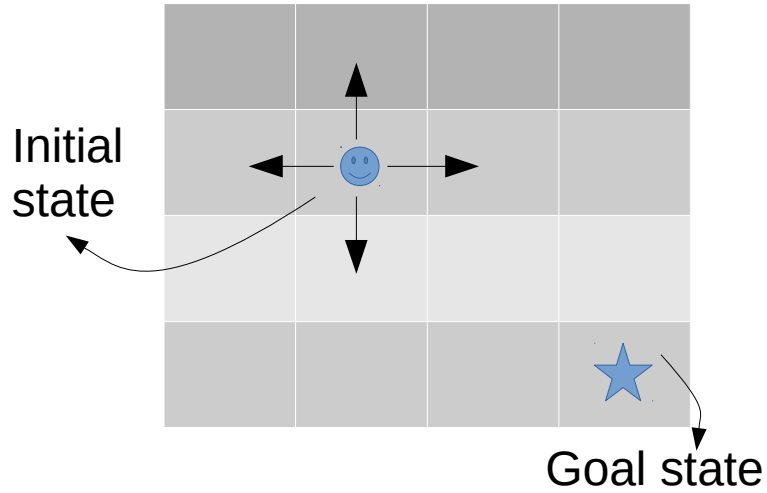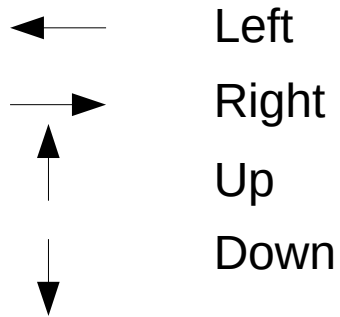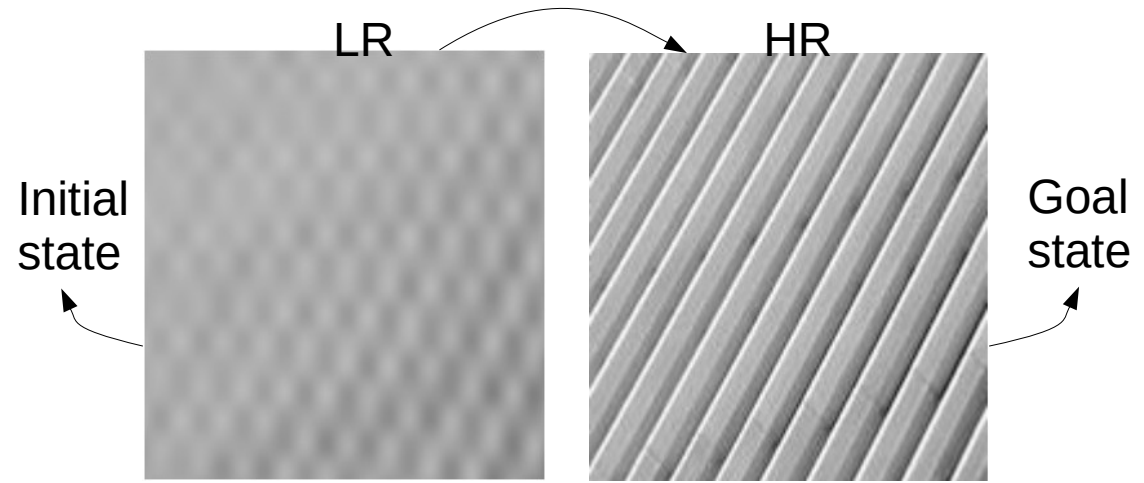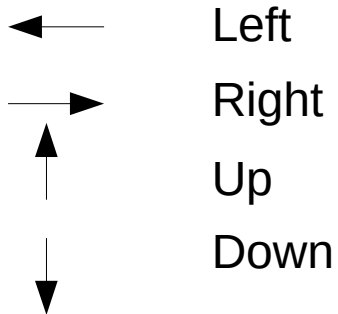
Unknown Action Variables
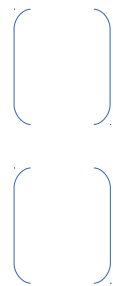
- Parameterize by matrices

- Could represent
    - horizontal edge detection
    - Or vertical edge detection
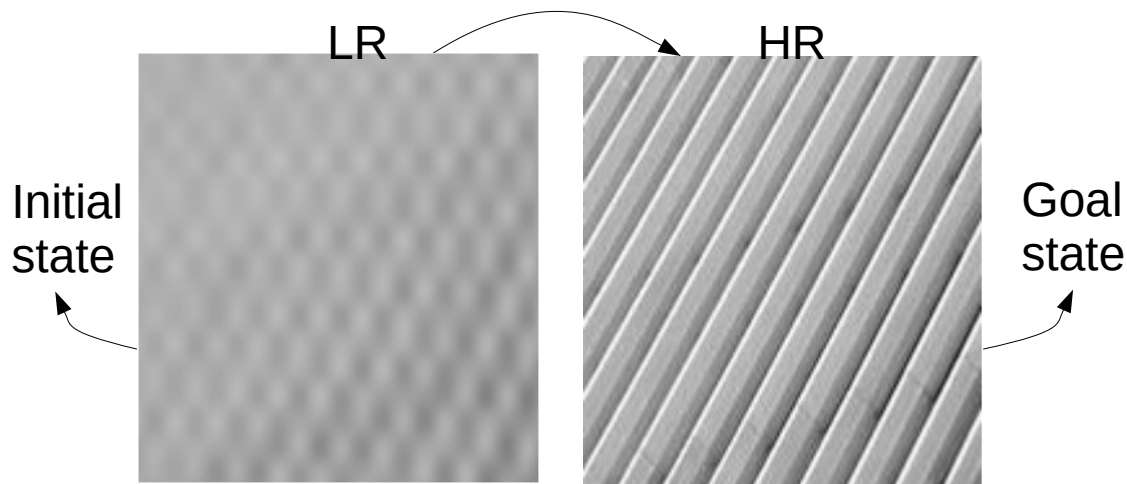
Super Resolution Environment
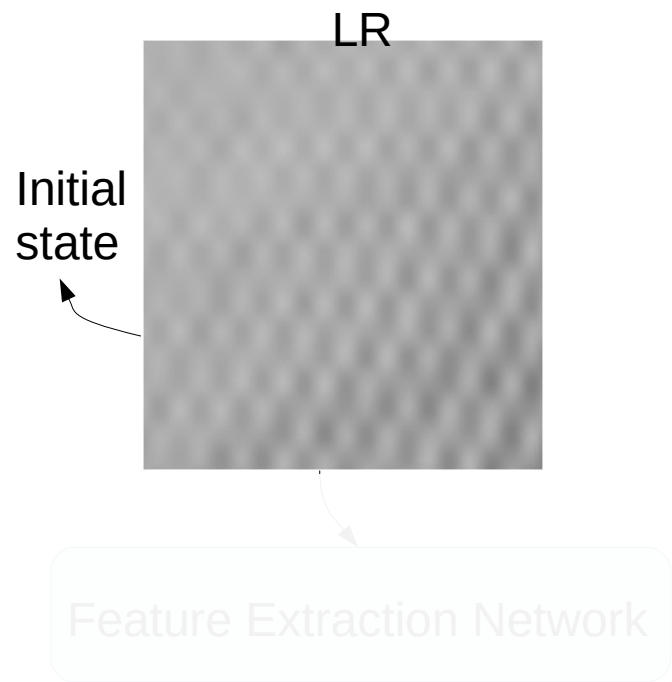
LR

Initial
state

Feature Extraction Network

Super Resolution Environment

LR

Initial
state

Feature Extraction Network

Siamese Policy on Actor

Siamese Policy Network

Actor Network

Super Resolution Environment

LR

HR

Initial
state

Goal
state

Observed Reward

Feature Extraction Network

Reward Model

Siamese Policy on Actor

Siamese Policy Network

Actor Network

Super Resolution Environment

LR

Initial state

HR

Goal state

Observed Reward

Feature Extraction Network

Reward Model

Siamese Policy on Actor

Siamese Policy Network

Actor Network
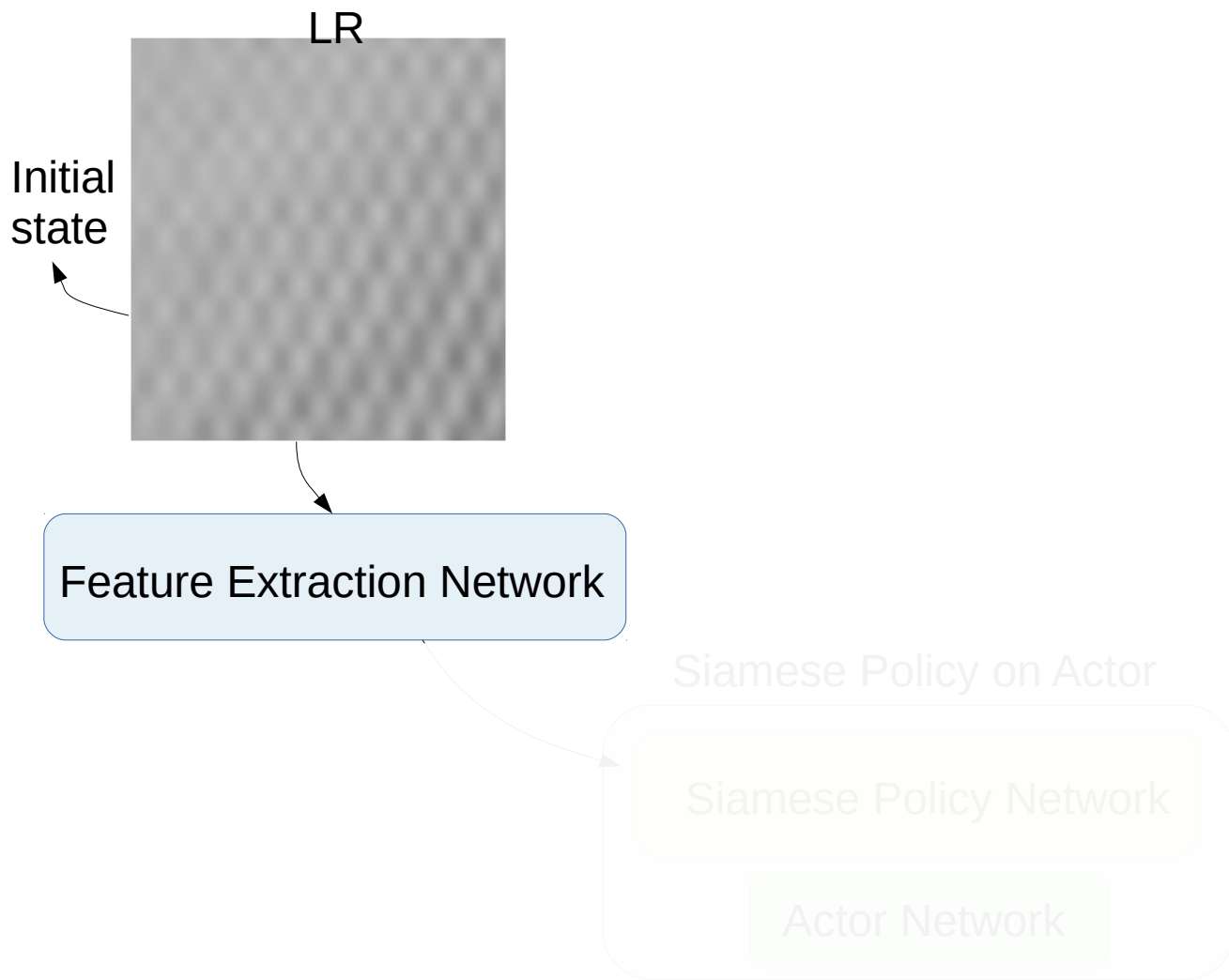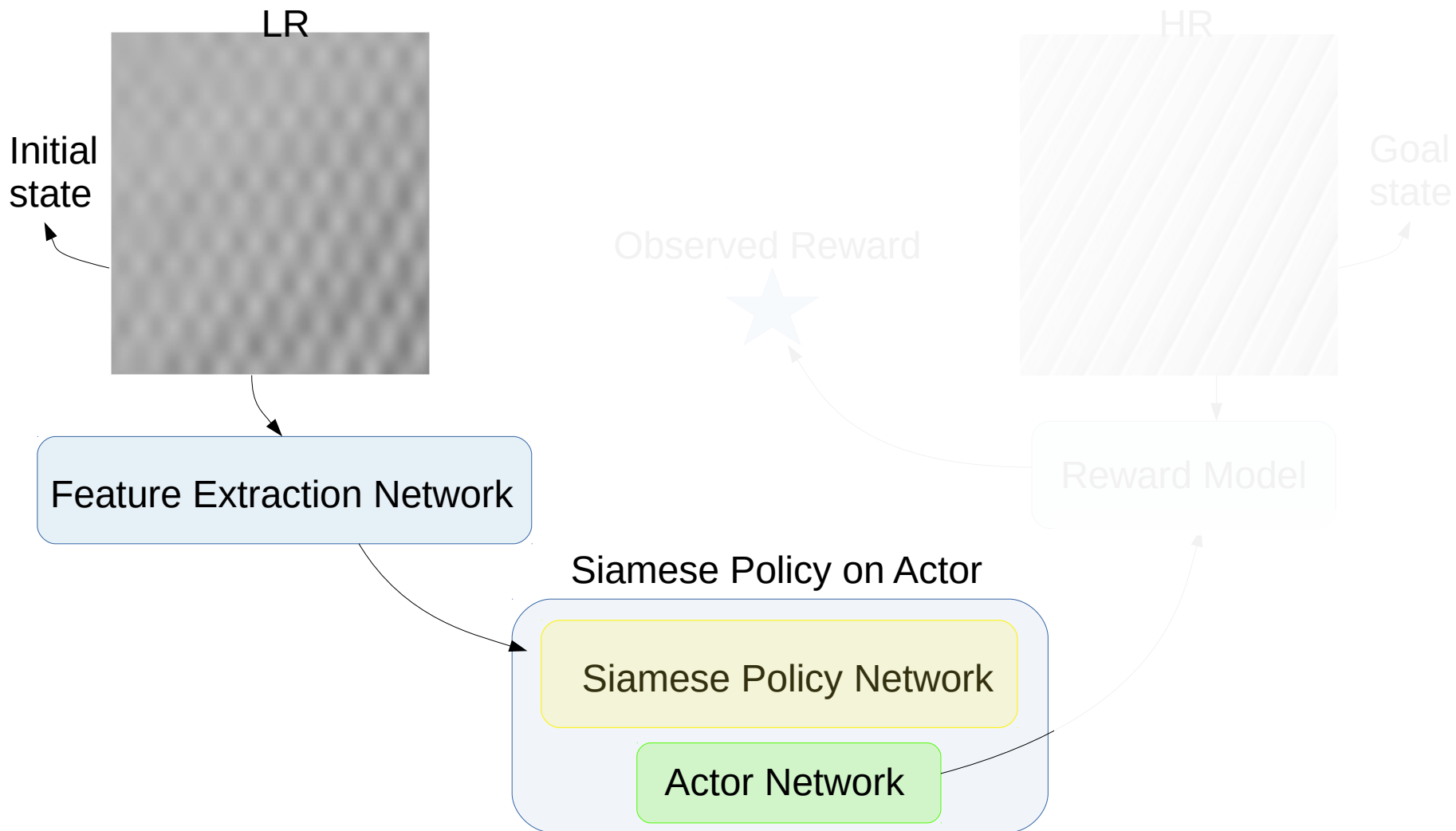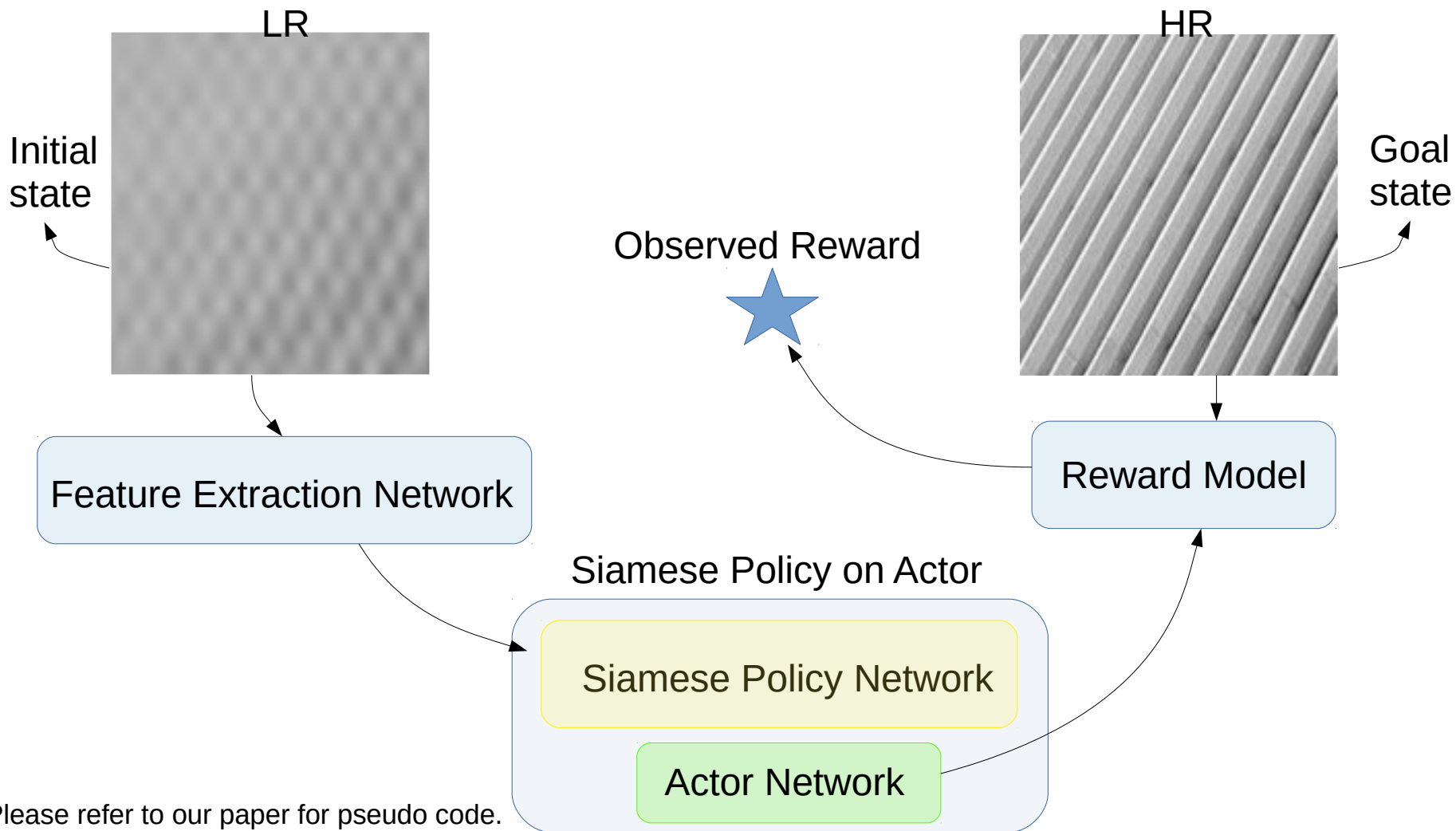
N.B. Please refer to our paper for pseudo code.

Exploration-Exploitation Paradigm
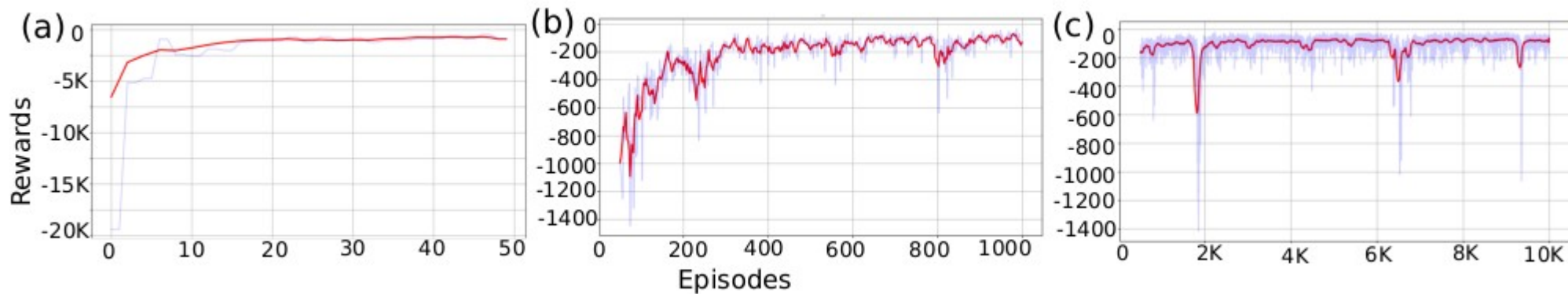
- Explores various action variables

- Exploits selected action variables after sufficient exploration

- Repeatedly performs selected actions

Experimental results on CelebA Dataset

| Ground Truth PSNR (dB)/ SSIM | BiCubic 25.91/0.82 | SRCNN 28.98/0.90 | SPOA(our) 29.73/0.92 |

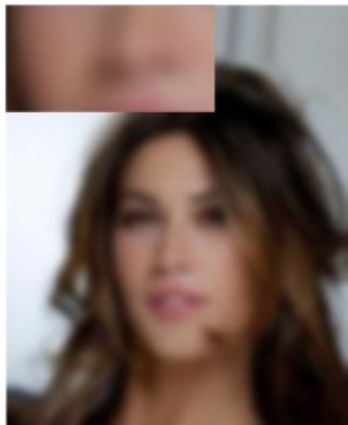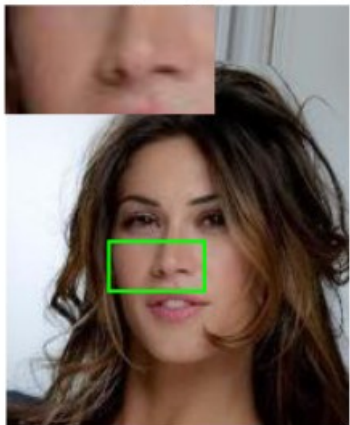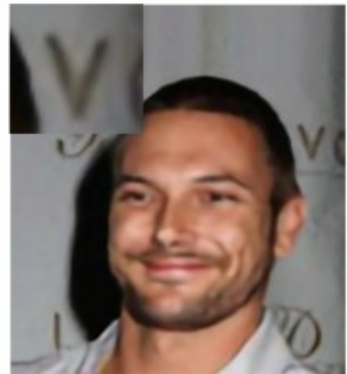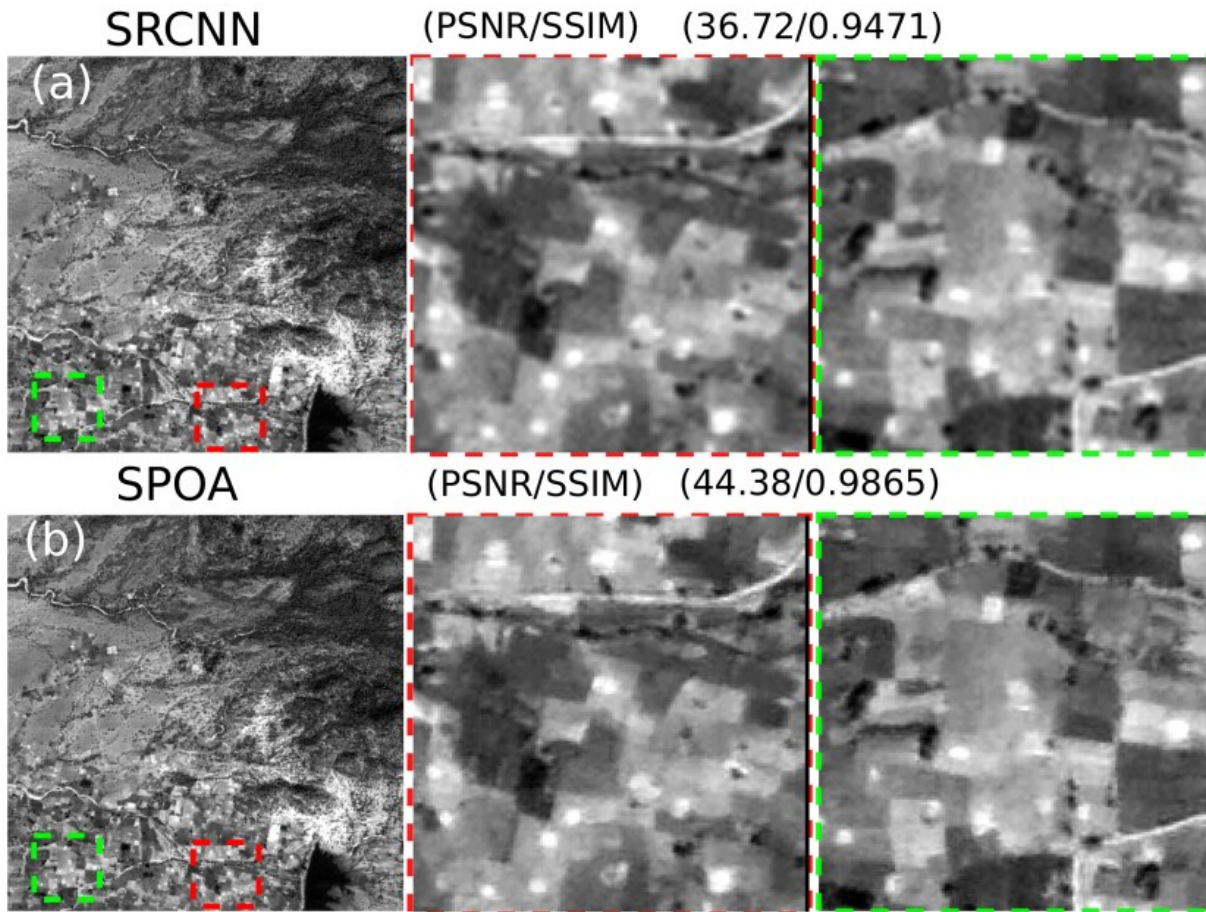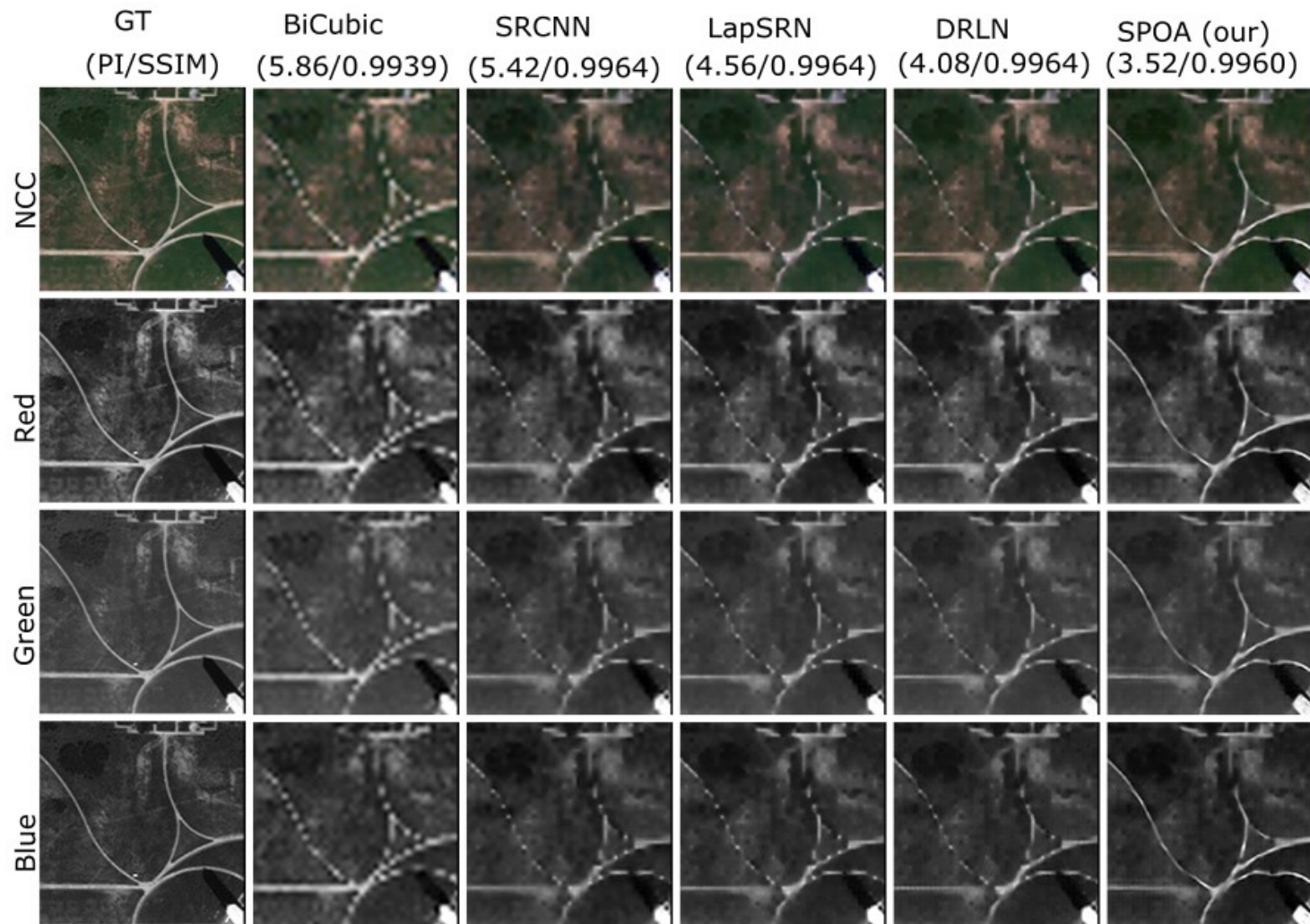| PSNR (dB)/ SSIM | 26.25/0.84 | 29.6/0.92 | 30.12/0.94 |
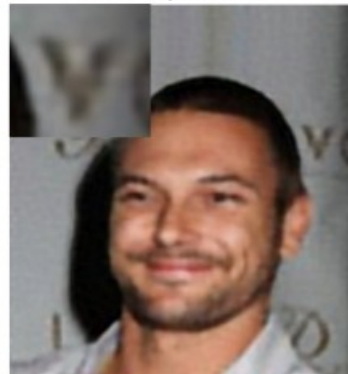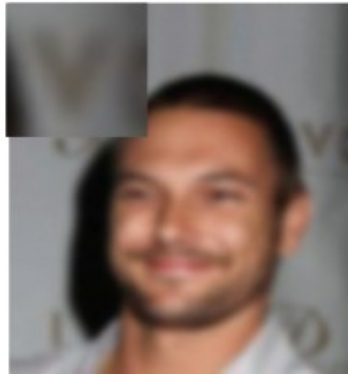
Experimental results on IRS Dataset

SRCNN   (PSNR/SSIM)   (36.72/0.9471)

(a)

SPOA   (PSNR/SSIM)   (44.38/0.9865)

(b)

IRS: Indian Remote Sensing Satellite

Comparison with State-of-the-art Methods on WorldView-2

Comparison with State-of-the-art Methods on WorldView-2

Metrics	PSNR	SSIM	SRE	SAM	NIQE	Ma's	PI
BiCubic	57.51	0.9939	46.48	17.25	5.50	3.77	5.86
SRCNN [14]	59.15	0.9964	48.10	14.14	5.73	4.88	5.42
LapSRN [28]	59.31	0.9964	48.08	13.98	5.08	5.96	4.56
DRLN [2]	59.32	0.9964	48.10	13.97	4.21	6.03	4.08
SPOA(DRLN)	58.89	0.9960	47.94	14.69	3.65	6.60	3.52
SPOA(DRLN)+SA	59.33	0.9966	48.20	13.81	5.02	5.54	4.74
SPOA(DRLN)+SA+VGG	59.22	0.9963	48.23	14.13	4.30	6.20	4.05
SPOA(DRLN)+VGG	58.98	0.9961	47.94	14.60	4.16	6.56	3.80
GT	-	-	-	-	2.05	7.01	2.52

Proposed Approach

- SPOA(DRLN) achieves state-of-the-art result in perceptual quality.

- SPOA(DRLN)+SA achieves state-of-the-art result in distortion metrics.

Yochai Blau and Tomer Michaeli. The perception-distortion tradeoff. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 6228–6237, 2018.

Summary

- Explored plausible usage of deep reinforcement learning in super resolution.

- Introduced SPOA to circumvent tractability issues in RL based super resolution.

- Key ingredients of SPOA:
  - Feature Extraction Network
  - Siamese Policy Network
  - Actor Network

- Provided pseudo code for training SPOA in an end-to-end fashion.

- Experimented on multiple datasets:
  - CelebA
  - IRS-1C/1D
  - WorldView-2

- Investigated perception-distortion tradeoff.

Few Noteworthy Extensions

- <u>Extension of SPOA</u> to wide variety of problems which are currently solved using supervised learning.

- Explore broad <u>spectrum of reinforcement learning</u> algorithms in this framework.

- Study how well SPOA figures out <u>matrix representation of actions</u> by hiding known action variables in <u>RL benchmarks</u>.