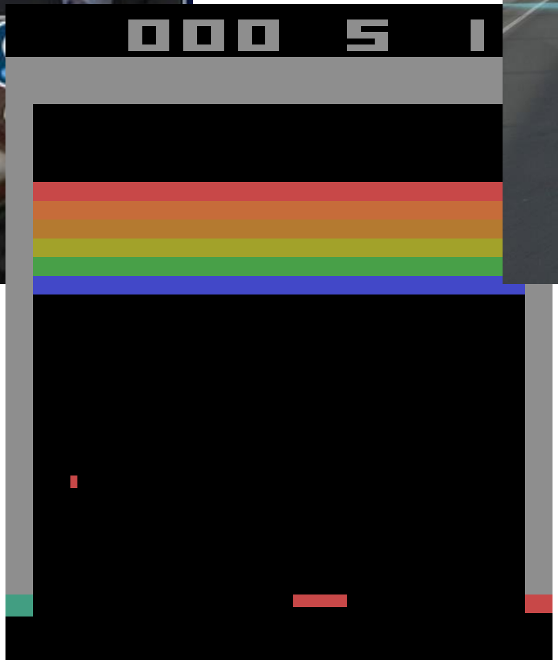
Transfer Learning for Related Reinforcement Learning Tasks via Image-to-Image Translation

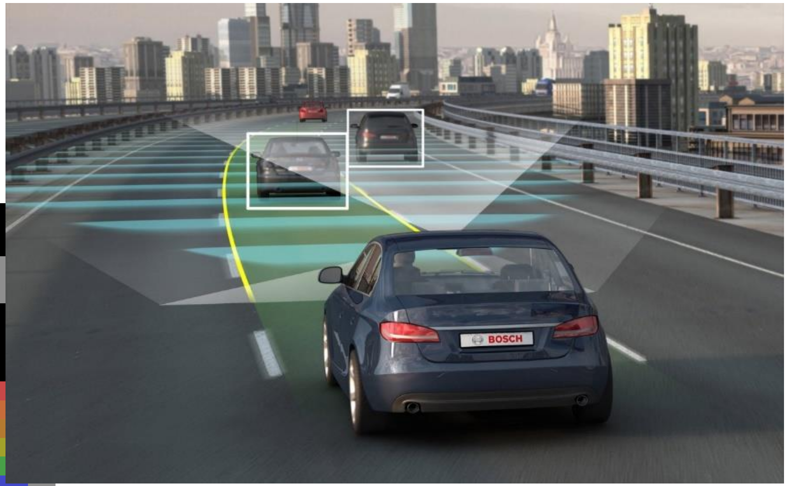
Shani Gamrian Yoav Goldberg

Bar-Ilan University

ICML 2019, Long Beach

Deep Reinforcement Learning





Transfer Learning

- Deep Reinforcement Learning is effective but fails to generalize.

Can we **TRANSFER** knowledge between related RL tasks?

Generalization Failures of Deep-RL

Breakout

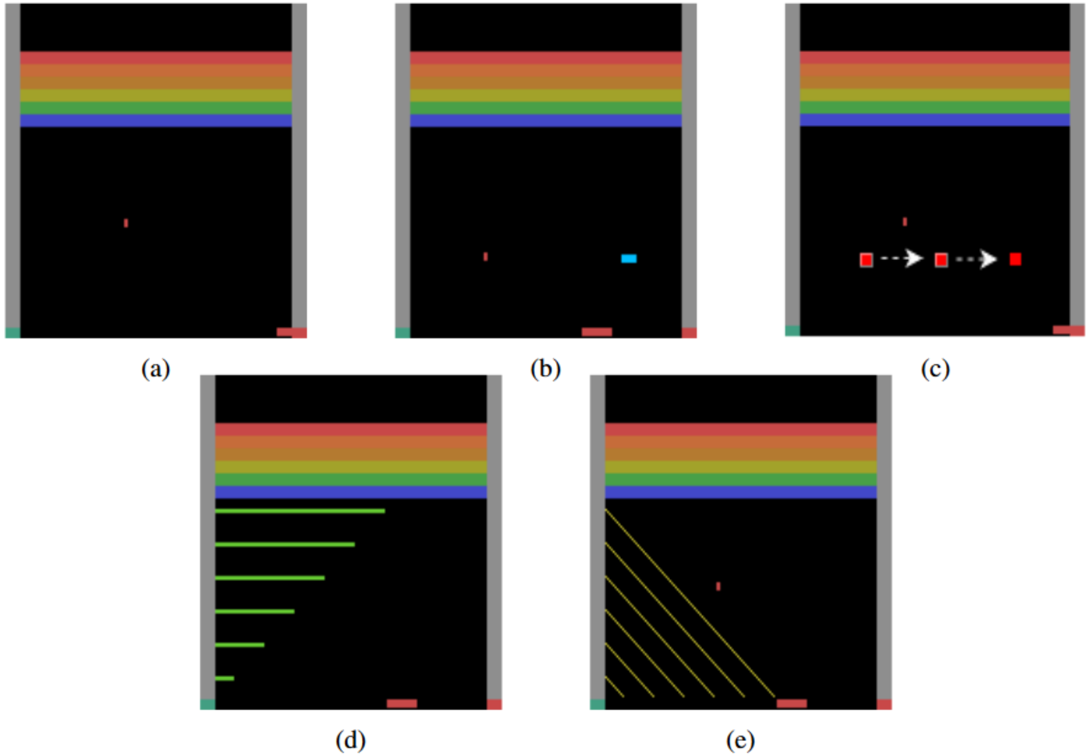
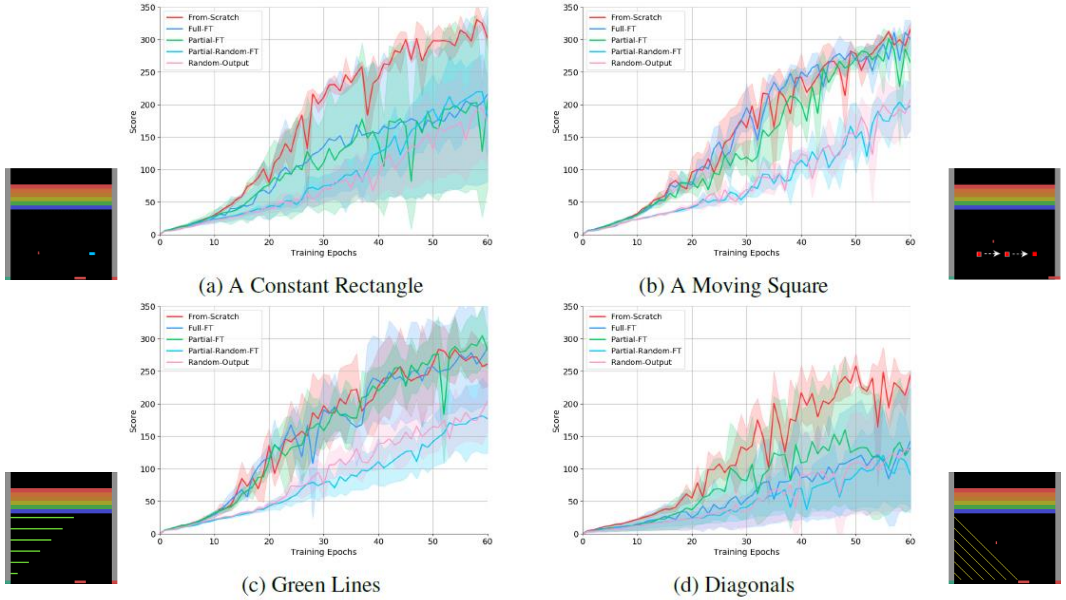


Figure 1: Various variations of the Breakout game: (a) Standard version, (b) A Constant Rectangle, (c) A Moving Square, (d) Green Lines, (e) Diagonals.

Generalization Failures of Deep-RL
Transfer Learning via Finetuning

- The results show that fine-tuning takes as long or longer than training from scratch!



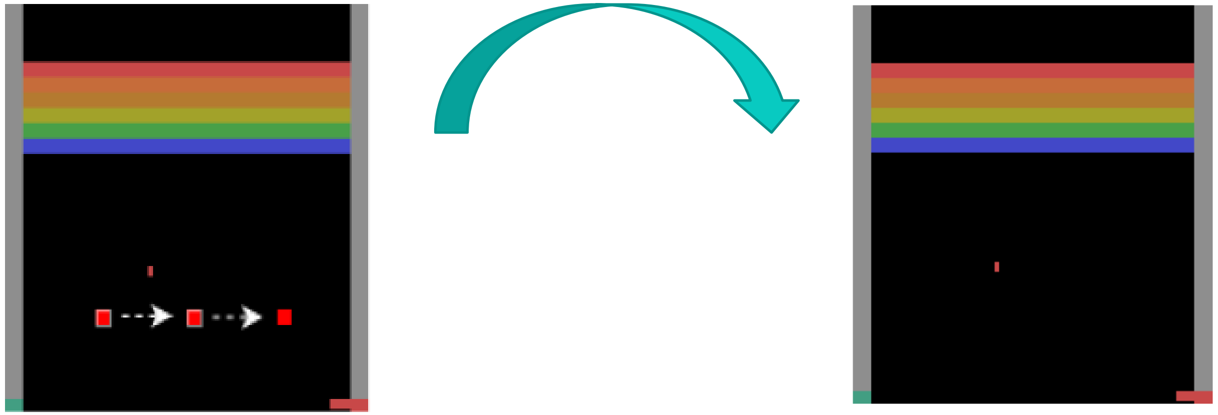
(a) A Constant Rectangle (b) A Moving Square

(c) Green Lines (d) Diagonals

Figure 2: A comparison between the different baselines on Breakout. The y-axis on each one of the plots shows the average reward per episode of Breakout during training. The x-axis shows the total number of training epochs where an epoch corresponds to 1 million frames. The plots are averaged on 3 runs with different random seeds.

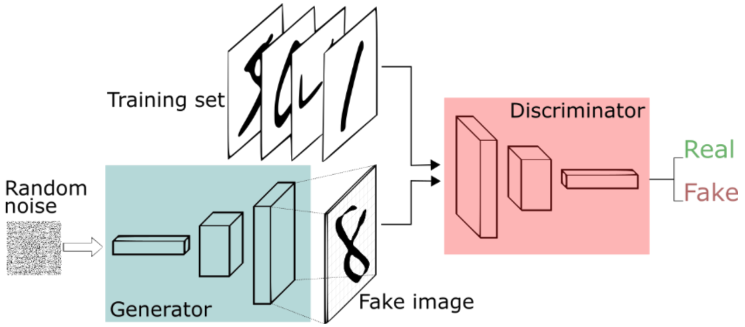
Analogy-based Zero-Shot Transfer with GANs

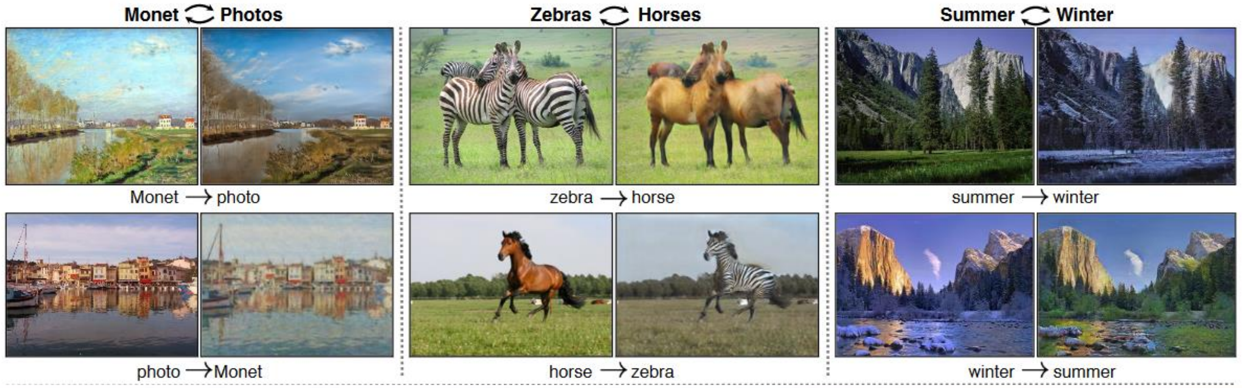
- **Problem:** finetuning fails to transfer between related tasks.
- **Our Solution:** Transfer by visual mapping.
- **How?:** map the input images from the target task to the source task.



UNsupervised Image-to-Image Translation (UNIT)

- Generative Adversarial Networks (GANs)



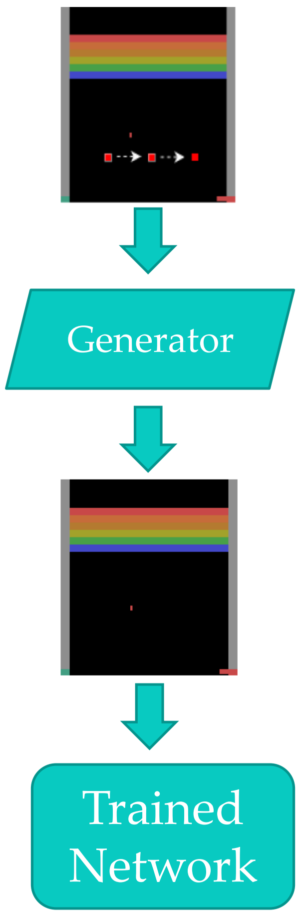


https://deeplearning4j.org/generative-adversarial-network

Analogy-based Zero-Shot Transfer with GANs

Experiments

- ❑We initialize the layers with the values of the trained network.
- ❑We run the game and translate each image from the target task to source task.
- ❑Our model accuracy is the score of the game.

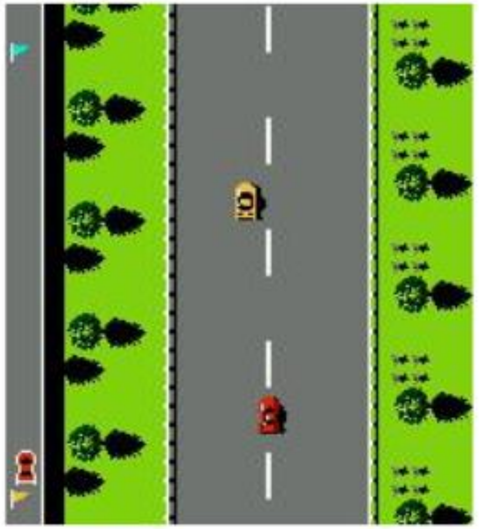
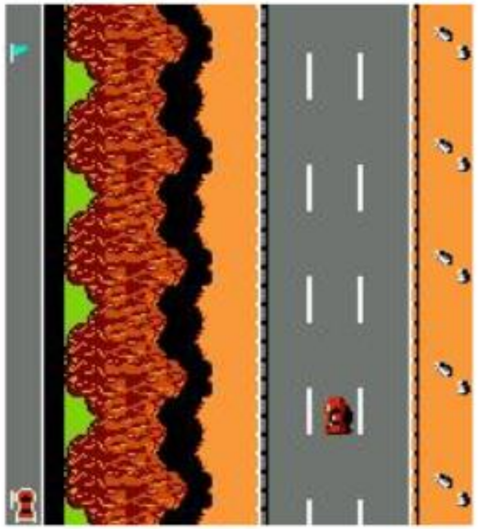


Analogy-based Zero-Shot Transfer with GANs
Breakout

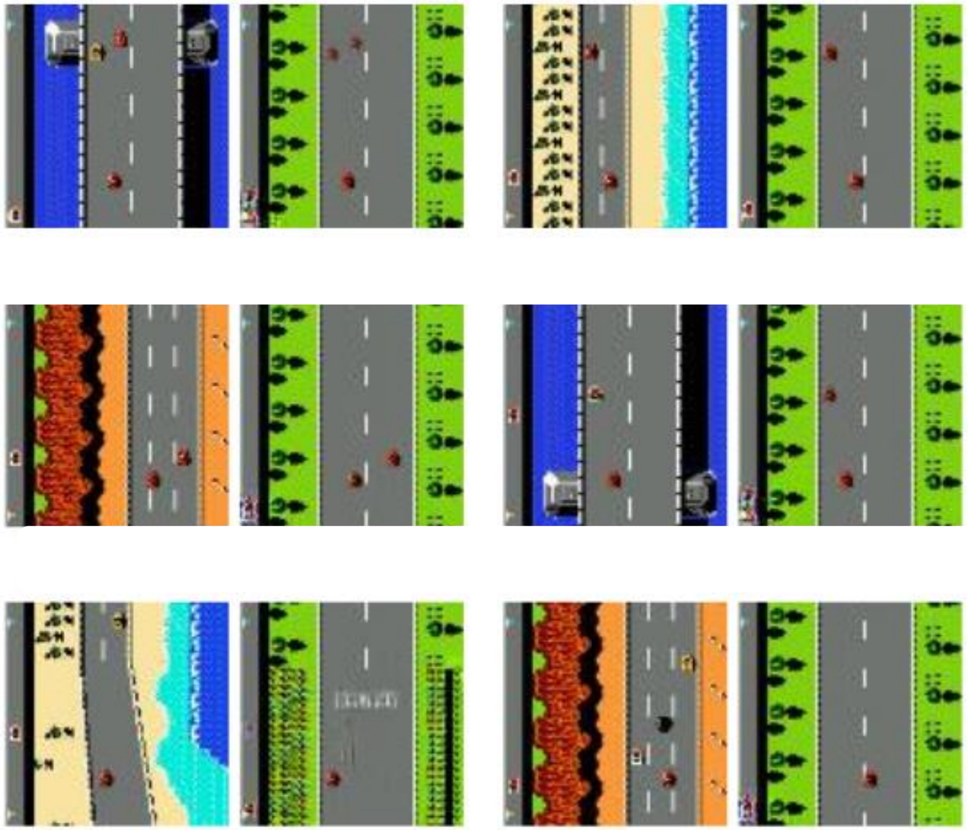
	Source		Target		Target with GANs	
	Frames	Score	Frames	Score	Frames	Score
A Constant Rectangle	43M	302	122	0	**260K**	**362**
A Moving Square	43M	302	100	0	**384K**	**300**
Green Lines	43M	302	186	2	**288K**	**300**
Diagonals	43M	302	100	0	**383K**	**330**

Our method is 100x more data efficient than training from scratch!

Road Fighter

Analogy-based Zero-Shot Transfer with GANs
Road Fighter



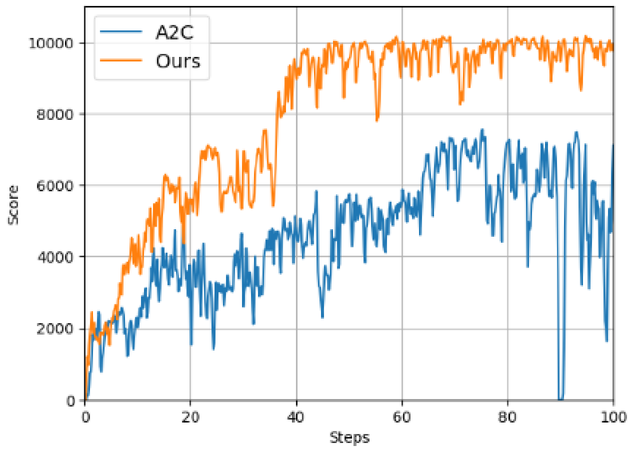
Analogy-based Zero-Shot Transfer with GANs
Road Fighter

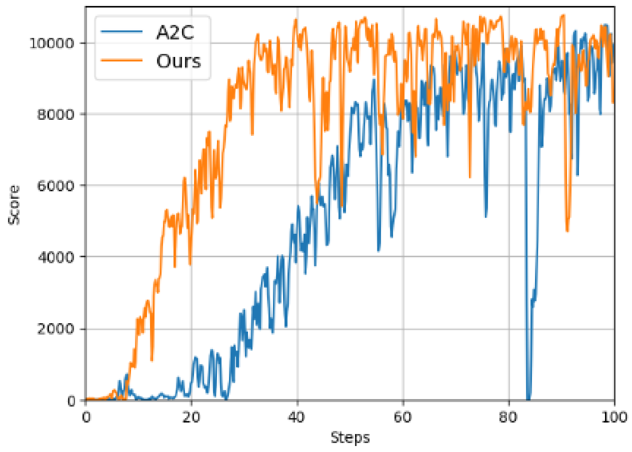
	Score (no transfer)	Score (analogy transfer)	# Frames (analogy)	# Frames (from scratch)
Level 2	0	5350	250K	12.4M
Level 3	0	5350	250K	31M
Level 4	0	2050	250K	13.6M

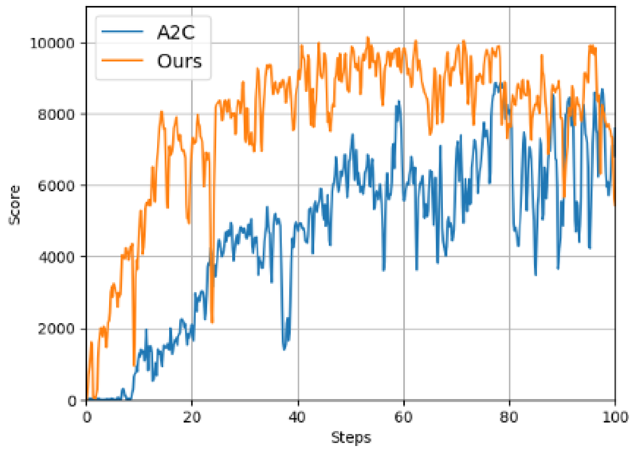
Accelerating RL with Imitation Learning

- Our transfer method is limited by the imperfect GAN generation and generalization abilities.
- We propose to use the visual-transfer based policy as imperfect demonstrations.
- We combine **off-policy supervised** updates and **on-policy RL** updates to accelerate the training process.
- We apply this method on *Road Fighter*.

Road Fighter
Accelerating RL with Imitation Learning







Road Fighter
Results

Poster #185

	Score (no transfer)	Score (analogy transfer)	# Frames (analogy)	# Frames (from scratch)	Score (+imitation)	#Frames (imitation)	# Frames (from scratch)
Level 2	0	5350	250K	12.4M	10230	38.6M	159M
Level 3	0	5350	250K	31M	10300	21M	54.4M
Level 4	0	2050	250K	13.6M	10460	13.4M	111M

With transfer + imitation learning, agent manages to complete the levels with just 20% of the needed frames.