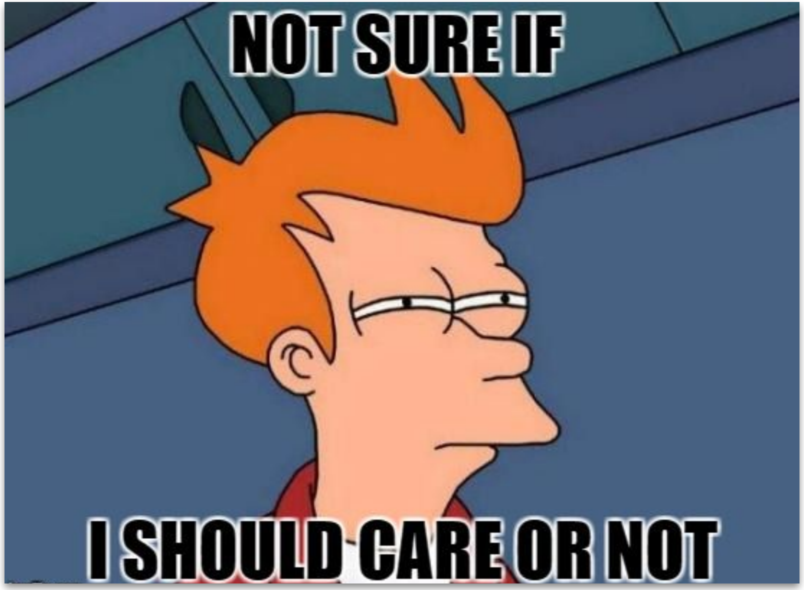
Leaky Tiling Activations: A Simple Approach to Learning Sparse Representations Online

Authors: Yangchen Pan, Kirby Banman, Martha White

Presentation by Erfan Miahi

What is the problem!?

Catastrophic Forgetting (Interference): the tendency of neural networks to forget what they have learned by learning from new data points



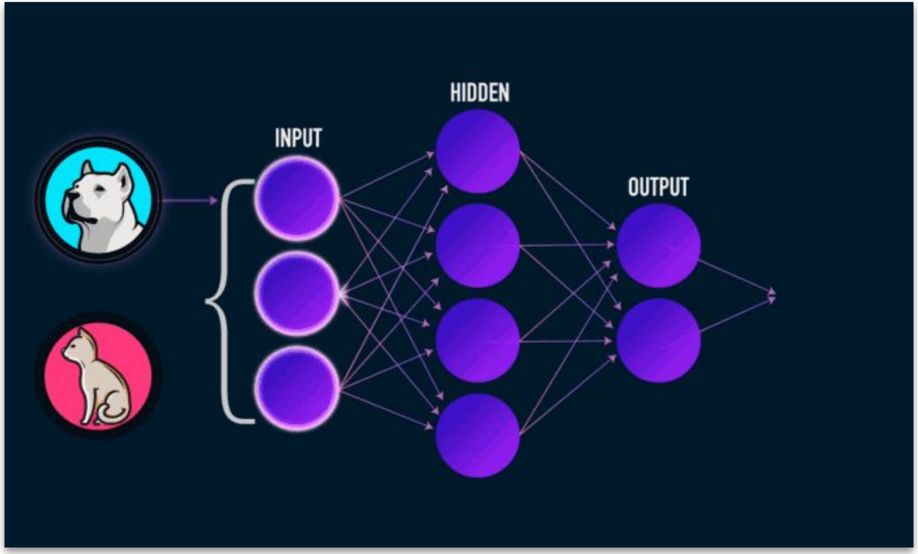
Source: Link

Catastrophic Forgetting is a Serious Problem in Online Supervised Learning!

Example: Online classification of cat and dog:

1. System shows hundred cat examples at first
2. It shows only dog examples Afterward

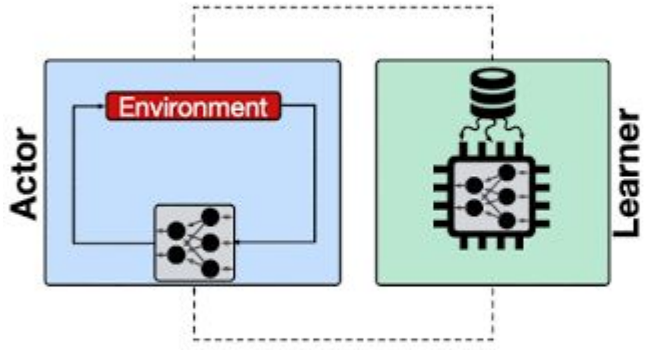
The system will eventually **forget** about the cat examples because we have **updated all the weights** that previously used for prediction of cat



Source: Link

Catastrophic Forgetting in RL is **Much More Serious**

- Data is Temporally Correlated



Source: Link

- Magnified by Bootstrapping

$$\mathbf{w}_{t+1} \doteq \mathbf{w}_t + \alpha[\underbrace{R_{t+1} + \hat{q}\left(S_{t+1}, A_{t+1}, \mathbf{w}_t\right)}_{\text{Bootstrapped Targets}} - \hat{q}\left(S_t, A_t, \mathbf{w}_t\right)]\nabla\hat{q}\left(S_t, A_t, \mathbf{w}_t\right)$$

How we usually approach this problem?

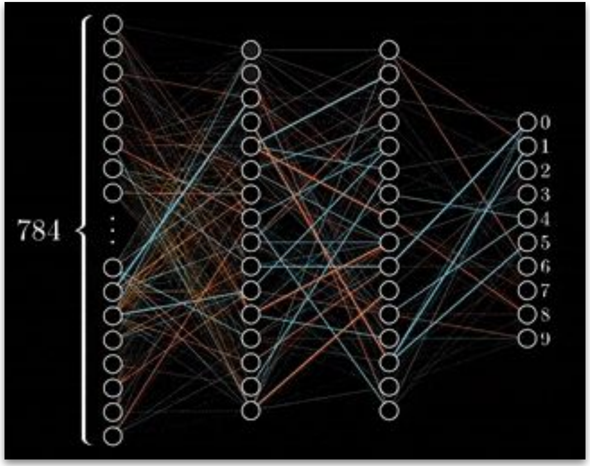
- Target Network
  - Moves us further from online reinforcement learning
  - Can further slow learning by using outdated estimates of value in the targets
- Experience Replay Buffer
  - Is incompatible with online reinforcement learning
  - Cannot scale to continuing environments

Source: Link

Kim, Seungchan, et al. "Deepmellow: removing the need for a target network in deep Q-learning." Proceedings of the Twenty Eighth International Joint Conference on Artificial Intelligence. 2019.

A Better Approach: Injecting Sparsity Into the Representation!

- A small number of features are **active**, for each input.
- Each update only **impacts** a small number of weights
- So it is less likely to **interfere** with many state values.



Source: Link

Why there is still a need for a new sparsity technique?

Existing techniques, at least, suffer from one of **these problems**:

- Causing dead neurons
- Being non-differentiable
- Not being simple to implement and understand
- Being unable to control the sparsity level
- Having mixed results

Outline

1. ~~The Importance of Developing a New Sparsity Technique~~
2. **Leaky Tiling Activation (LTA)**
3. Evaluating Stability and Performance of LTA
4. Final Reflections on Their Experimental Methodology and LTA

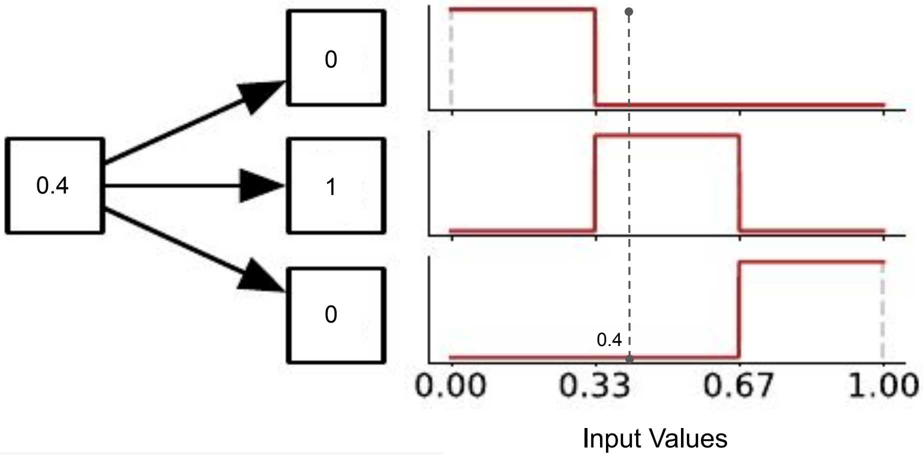
Tiling **(Binning!)** Activation

Parameters required to specify tiling activation:

- Range of input values: ***[l, u]***
- Number of tiles (bins): ***k***
- Bin size: ***$\delta = (u-l)/k$***

Parameters for this example:

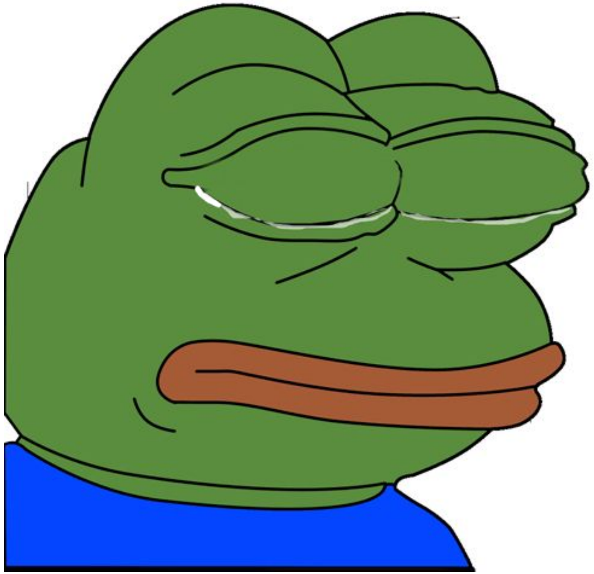
- **[l, u] = [0, 1]**
- **k = 3**
- **$\delta = (1-0)/3 = 0.33$**



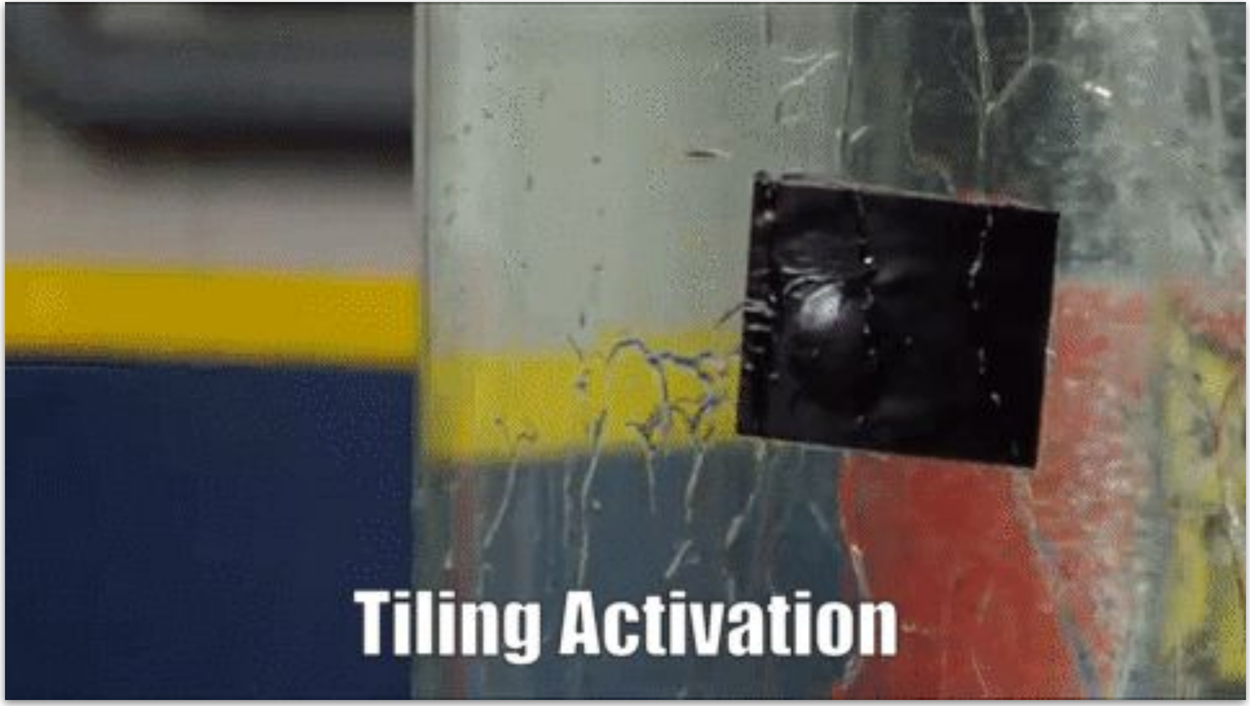
Pan, Yangchen, et al. Leaky Tiling Activations: A Simple Approach to Learning Sparse Representations Online. International Conference on Learning Representations, 2021.

Tiling Activation does not work. **Why!?**

- A loss of precision due to aggregation
- Zero gradient almost everywhere

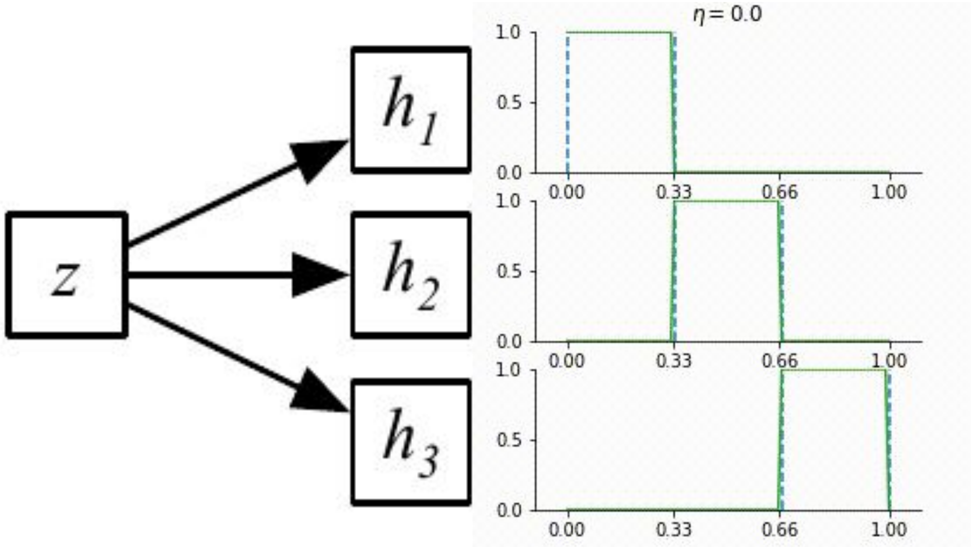


Leaky Tiling Activation (LTA)



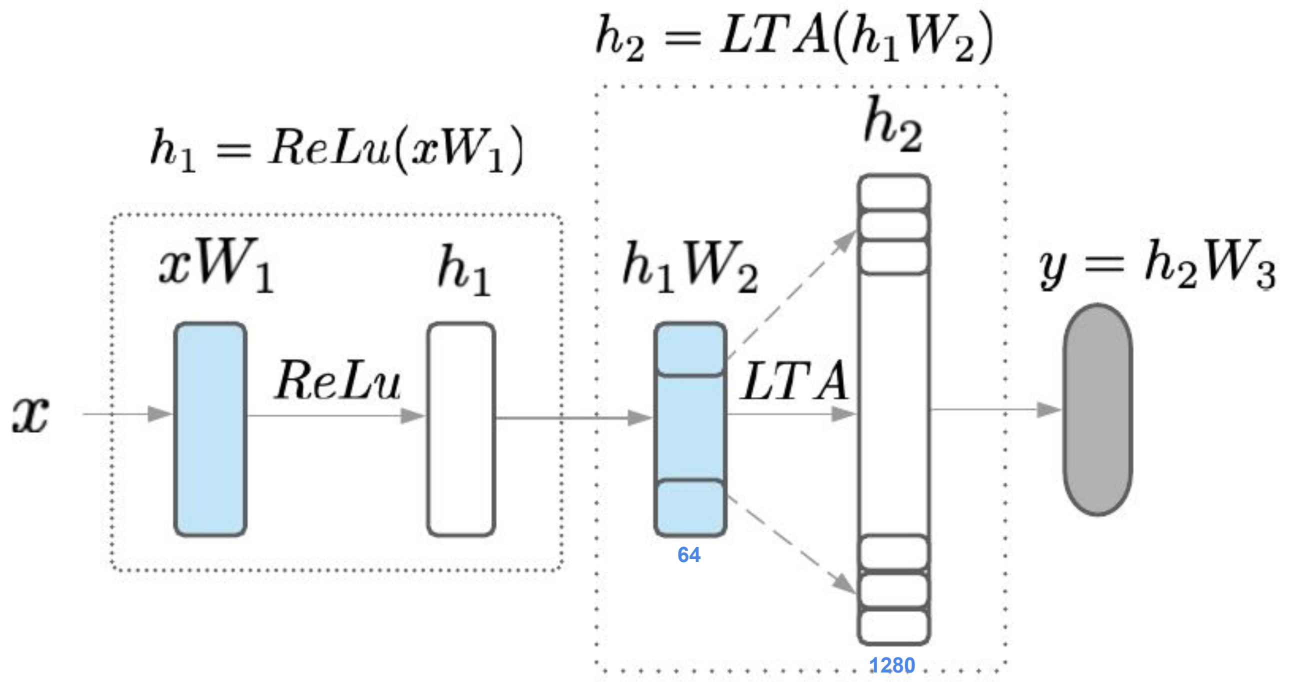
Leaky Tiling Activation (LTA)

- Introduces a new parameter: η
- This parameter determines the sparsity level
- Or the level of leakage from one bin to the neighboring bins



Pan, Yangchen, et al. Leaky Tiling Activations: A Simple Approach to Learning Sparse Representations Online. International Conference on Learning Representations, 2021.

A Visualization of a Neural Network with an LTA layer



K = 20

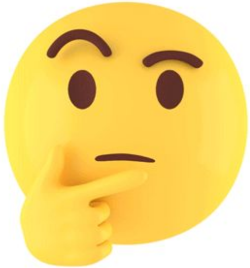
What problems

LTA is

- Differentiable
- Simple to use and implement
- Able to control the sparsity level

But, one question still remains: Does this technique produce

- Mixed results?
- Dead neurons?



Outline

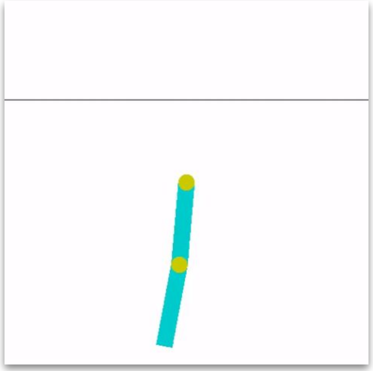
1. ~~The Importance of Developing a New Sparsity Technique~~
2. ~~Leaky Tiling Activation (LTA)~~
3. **Evaluating Stability and Performance of LTA**
   a. **Overall Performance in environments with discrete and continuous action spaces**
   b. Comparison with other Sparse Approaches
   c. Testing Stability in a Simulated Autonomous Driving Domain
4. Final Reflections on Their Experimental Methodology and LTA

Goals for Evaluation of Overall Performance

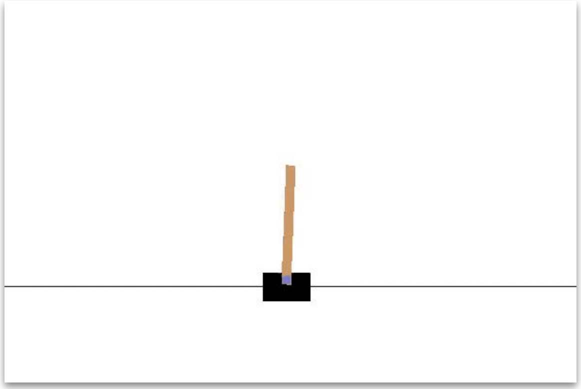
- Obtain improved performance with LTA, **with fixed parameter choices** across different domains
- **Improve stability in learning with LTA**
- See if we can **remove the need to use target networks** with LTA

Source: Link

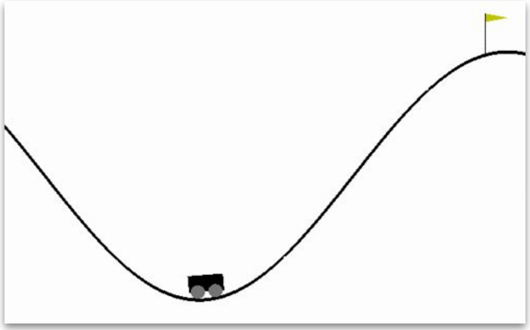
Environments (Discrete Action Space)



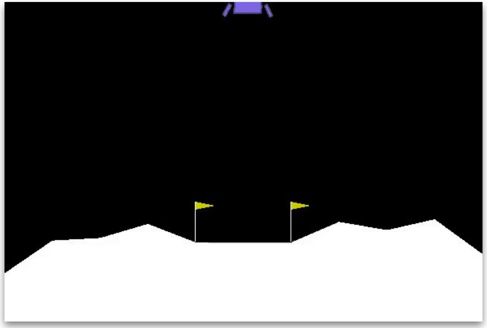
Acrobot-v1



CartPole-v1



MountainCar-v0



LunarLander-v2

Brockman, Greg, et al. "Openai gym." arXiv preprint arXiv:1606.01540 (2016).

LTA and Baselines (Discrete Action Space)

- DQN-LTA:
  - ***64x1280*** Hidden Units
  - LTA only used on the last hidden layer
    - Ranges: ***[-20, 20]***
    - Number of bins: ***k = 20***
    - Sparsity Parameter: $\boldsymbol{\eta = 2}$
- Baselines:
  - DQN: DQN with tanh or ReLU on the last layer (Best Parameter Reported)
    - ***64x64*** Hidden Units
    - ReLU units
  - DQN-Large: DQN, but with the last layer of same size as DQN-LTA.
    - ***64x1280*** Hidden Units
    - ReLU units

Pan, Yangchen, et al. Leaky Tiling Activations: A Simple Approach to Learning Sparse Representations Online. International Conference on Learning Representations, 2021.

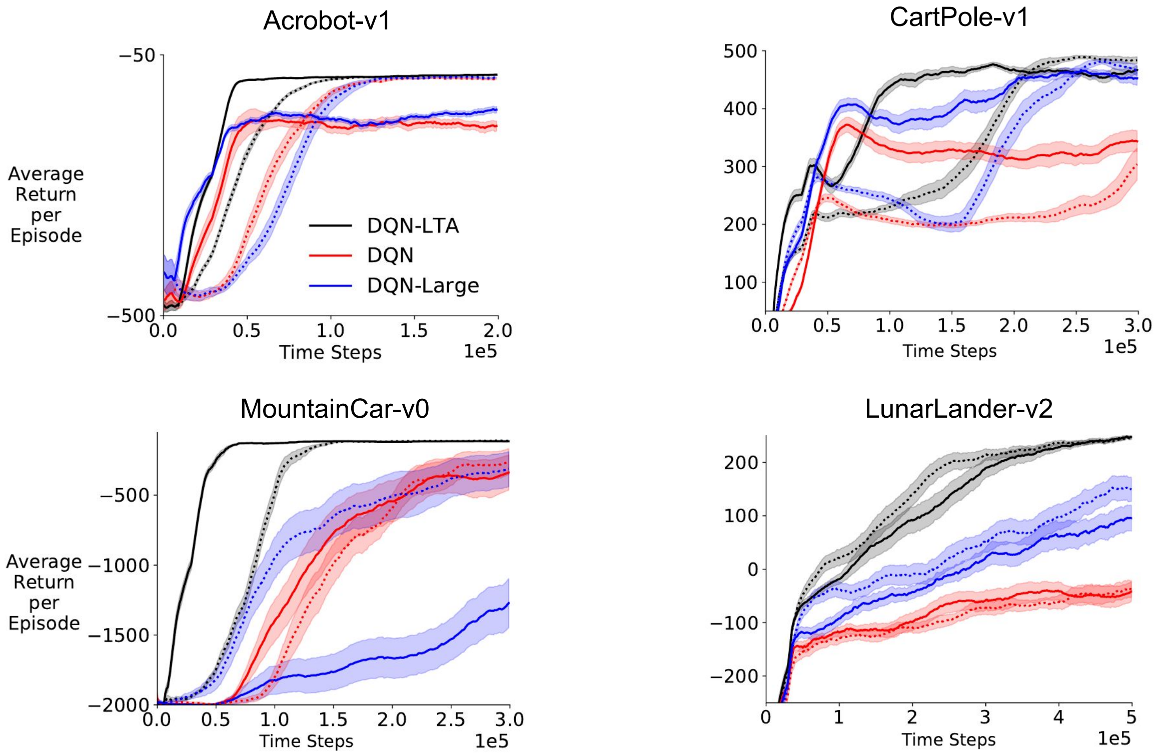
Shared Settings for DQN Experiments

- Adam optimizer
- Learning rate is **0.0001**
- Xavier Initialization
- The number of warm-up steps is **5000**
- Target network update frequency: each **1000** time steps
- Mini-batch size of **64**
- Experience replay buffer size of **100,000**
- Epsilon-greedy exploration with constant **Epsilon = 0.1**
- Discounting factor is $\boldsymbol{\gamma = 0.99}$

Evaluation Settings (Discrete Action Space)

- **Runs**: **20**
- **Offline evaluation**: Every **1000** training/environment time steps
  - **Epsilon = 0.05**
- **Reports**: Learning-curves with mean and standard error
  - The learning curve is smoothed over a **window of size 30** before averaging across runs

Overall Performance (Discrete Action Space)

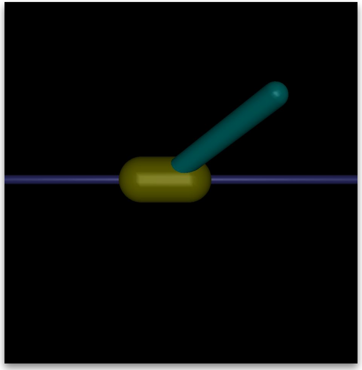


The **dotted line** indicates algorithms trained **with target networks.**

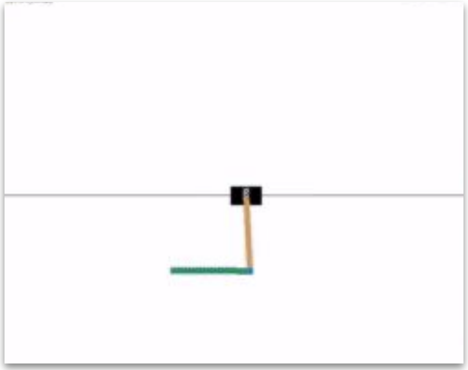
Paper's Conclusions on Overall Performance (Discrete Action Space)

- With or without using a target network, DQN with LTA can significantly outperform the version without using LTA.
- LTA has significantly lower variability across runs (smaller standard errors) in most of the figures.
- DQN-LTA trained without a target network outperforms DQN-LTA trained with a target network, which indicates a potential gain by removing the target network.
- Without using LTA, DQN trained without a target network cannot perform well in general (remember this part), providing further evidence for the utility of sparse feature highlighted in previous works.
- Simply using a larger neural network does not obtain the same performance improvements, and in some cases significantly degrades performance.

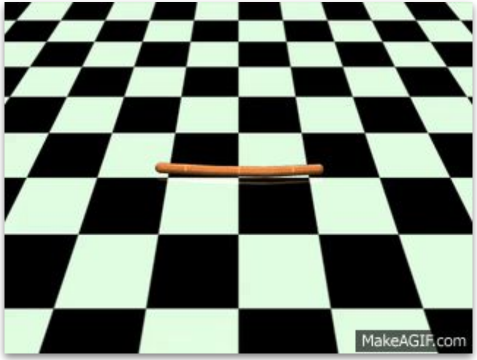
MuJoCo Environments (Continues Action Space)



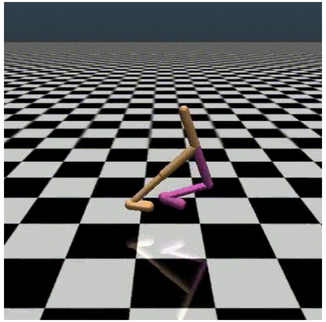
Inverted Pendulum
(Link)



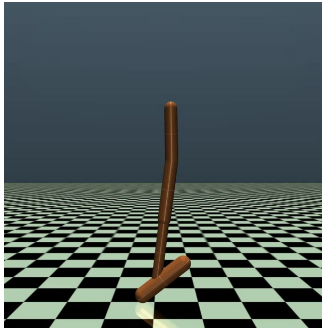
Double Inverted Pendulum
(Link)



Swimmer
(Link)



Walker 2D
(Link)



Hopper
(Link)

LTA and Baselines (Continues Action Space)

- DDPG-LTA:
  - ***200x2000*** Hidden Units
  - LTA only used on the last hidden layer
    - Ranges: ***[-20, 20]***
    - Number of bins: ***k = 20***
    - Sparsity Parameter: $\boldsymbol{\eta = 2}$
- Baselines:
  - DDPG:
    - ***200x100*** Hidden Units
    - ReLU units
  - DDPG-Large: DDPG, but with the last layer of same size as DDPG-LTA.
    - ***200x2000*** Hidden Units
    - ReLU units

Pan, Yangchen, et al. Leaky Tiling Activations: A Simple Approach to Learning Sparse Representations Online. International Conference on Learning Representations, 2021.

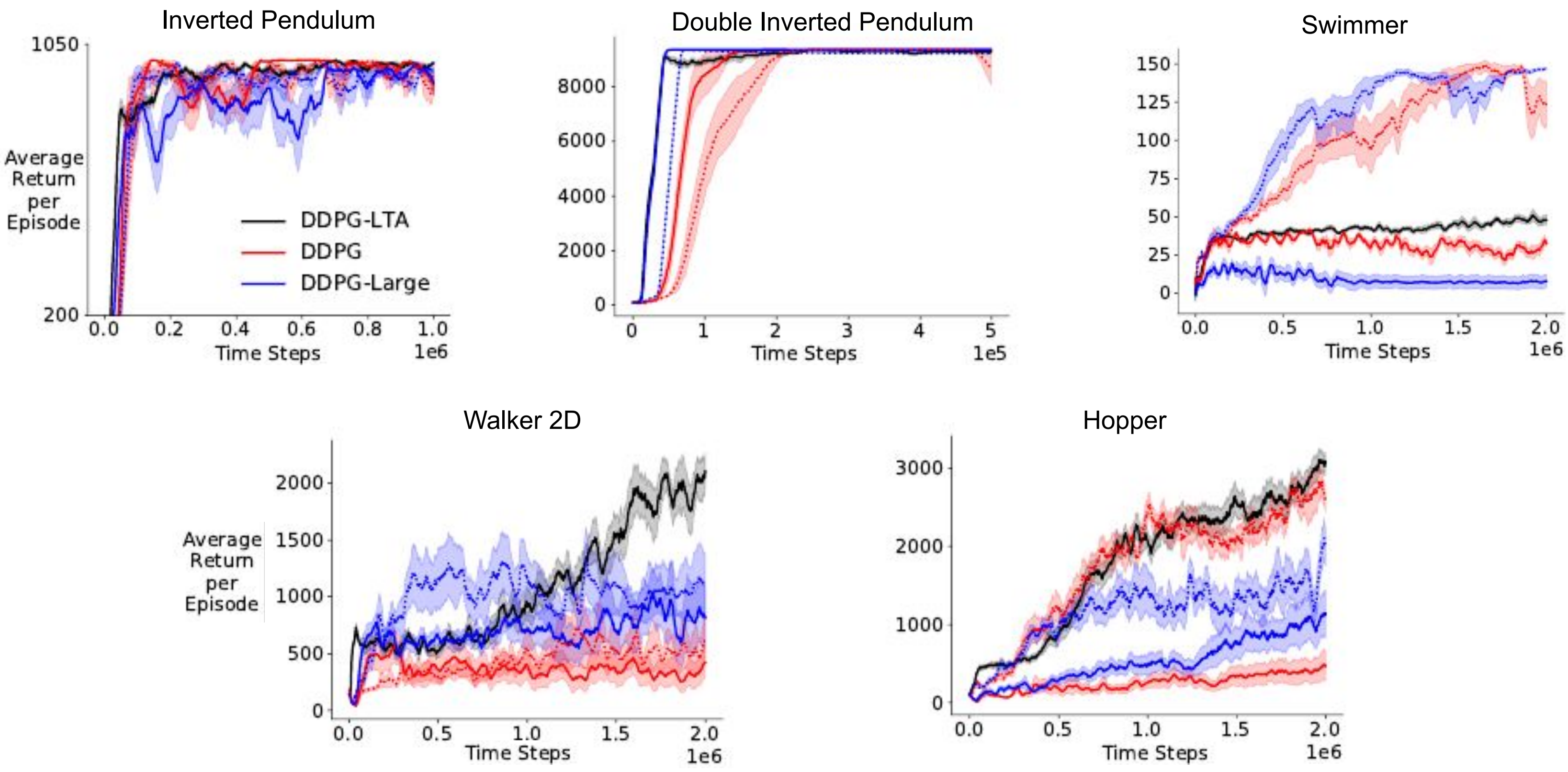
Shared Settings for DDPG Experiments (Continues Action Space)

- Adam optimizer
- **Actor network learning rate is 0.0001**
- **Critic Network learning rate is 0.001**
- Xavier Initialization
- **The number of warm-up steps is 10,000**
- **Target network moving rate** is **0.001**
- Mini-batch size of **64**
- Experience replay buffer size **100k**
- Discounting factor is $\gamma = 0.99$

Evaluation Settings (Continues Action Space)

- **Runs**: **20**
- **Offline evaluation**: Every **1000** training/environment time steps
- **Reports**: Learning-curves with mean and standard error
    - The learning curve is smoothed over a **window of size** **10** before averaging across runs

Overall Performance (Continues Action Space)



The **dotted line** indicates algorithms trained **with target networks.**

Paper's Conclusions on Overall Performance (Continuous Action Spaces)

- The results are qualitatively similar to the discrete-action environments (Are they!?), except in one domain (Swimmer).
- In all other domains, DDPG equipped with LTA, without target networks, achieves **comparable** and sometimes **significantly better performance** to DDPG.
- Swimmer highlights that LTA is not always sufficient on its own to overcome instabilities, and could be complemented by strategies such as using mellowmax

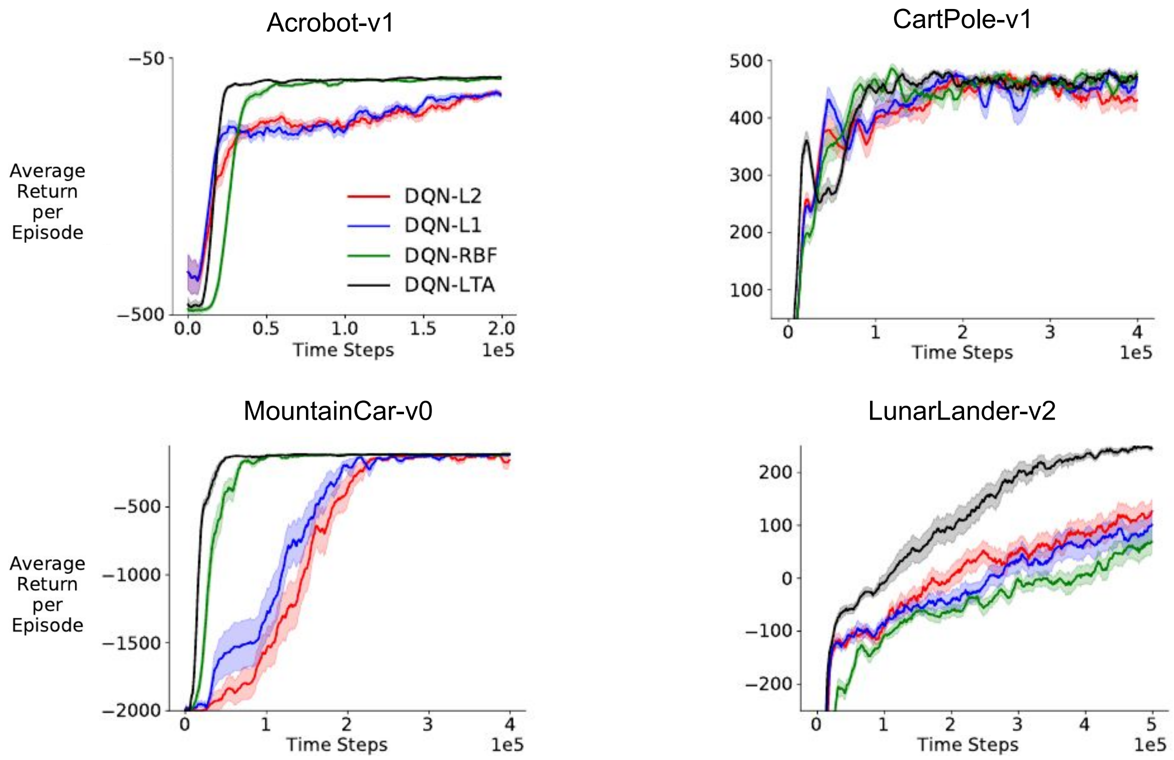
Outline

1. ~~The Importance of Developing a New Sparsity Technique~~
2. ~~Leaky Tiling Activation (LTA)~~
3. **Evaluating Stability and Performance of LTA**
    a. ~~Overall Performance in environments with discrete and continuous action spaces~~
    b. **Comparison with other Sparse Approaches**
    c. Testing Stability in a Simulated Autonomous Driving Domain
4. Final Reflections on Their Experimental Methodology and LTA

Baselines

- DQN-RBF: DQN using radial basis functions (RBFs) on the last layer
  - ***64x1280*** Hidden Units
- DQN-L1/L2: L1 and L2 only used on the activation function of the final hidden layer
  - ***64x1280*** Hidden Units
  - L1 and L2 are swept over **{0.1, 0.01, 0.001, 0.0001}** on MountainCar, then they fix the chosen optimal weight **0.01** across all domains.

Comparison with other Sparse Approaches



Paper’s Conclusions on this Comparison

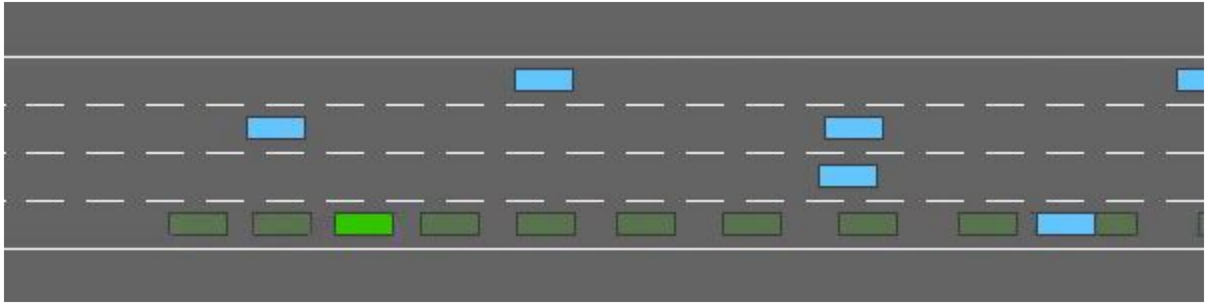
- LTA performs consistently well across all environments using a fixed parameter setting
- None of the other approaches achieve consistent performance, even though we tuned their parameters per environment!.
- Both the L1 and L2 approaches have a high variance across different random seeds.
- The RBF variant can do better than the L1 and L2 approaches but is worse than our algorithm.

Outline

1. ~~The Importance of Developing a New Sparsity Technique~~
2. ~~Leaky Tiling Activation (LTA)~~
3. **Evaluating Stability and Performance of LTA**
    a. ~~Overall Performance in environments with discrete and continuous action spaces~~
    b. ~~Comparison with other Sparse Approaches~~
    c. **Testing Stability in a Simulated Autonomous Driving Domain**
4. Final Reflections on Their Experimental Methodology and LTA

Simulated Autonomous Driving Environment (Highway)

- Goal: To show that DQN-LTA is more Stable than DQN
- Reward System
  - Having high speed is positively rewarded
  - Collisions with neighbouring vehicles are negatively rewarded
  - Driving on the right (bottom) side of the road is also rewarded.
- Observations are 25-dimensional
- Action space is discrete



Source: Link

Leurent, E. An environment for autonomous driving decision-making, 2018

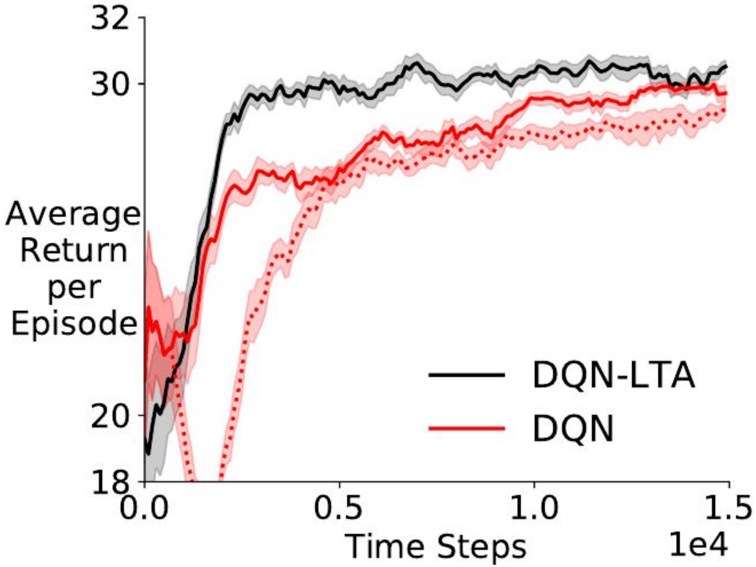
LTA and Baselines (Discrete Action Space)

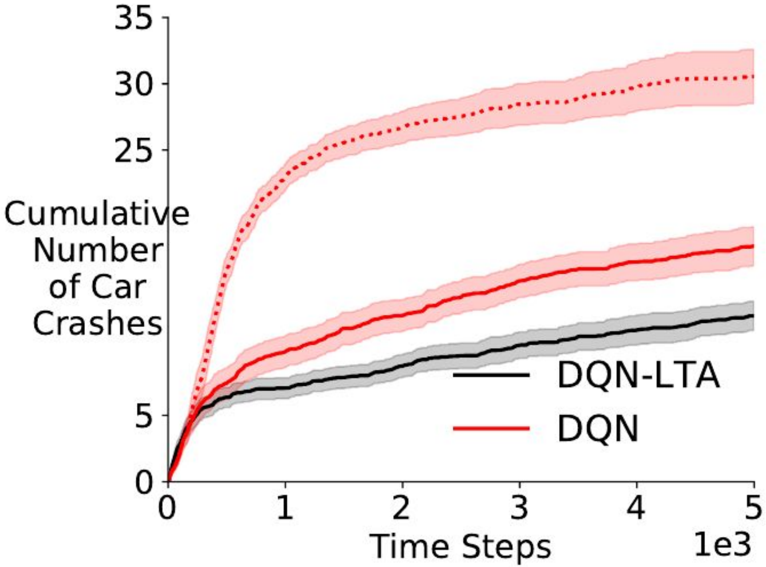
- DQN-LTA:
  - ***64x1280*** Hidden Units
  - LTA only used on the last hidden layer
    - Ranges: ***[-20, 20]***
    - Number of bins: ***k = 20***
    - Sparsity Parameter: $\eta = 2$
- Baseline:
  - DQN: DQN with tanh or ReLU on the last layer (Best Parameter Reported)
    - ***64x64*** Hidden Units
    - ReLU units

Pan, Yangchen, et al. Leaky Tiling Activations: A Simple Approach to Learning Sparse Representations Online. International Conference on Learning Representations, 2021.

Testing Stability

Runs: 30





The **dotted line** indicates algorithms trained **with target networks.**

Paper's Conclusions on Stability

- LTA learns faster, with significantly fewer car crashes incurred during the evaluation time steps.
- Target networks are harmful in this environment, potentially because they slow early learning, so the agent will accumulate a significant number of crashes before improving.
- They previously claimed that DQN generally cannot do well with target network, but it is evident from the previous plots that, in some cases, it can.

Outline

1. ~~The Importance of Developing a New Sparsity Technique~~
2. ~~Leaky Tiling Activation (LTA)~~
3. ~~Evaluating Stability and Performance of LTA~~
4. **Final Reflections on Their Experimental Methodology and LTA**

Pros and Cons of their Experimental Methodology

Pros:

- Used a good variety of environments to evaluate LTA
- Ran their algorithm for 20 number of runs
- Reported all the parameters

Cons:

- Did not sweep over target network update frequency
- Were not clear about some of their choices:
  - Switching from 20 runs to 30 runs in highway
  - Usage of LTA in DDPG
  - Usage of DDPG neither in Highway or when comparing it with other sparse techniques
- Stopped reporting performance of DQN-LTA with target network
- Inacceptable approach for sweeping over L1 and L2 parameters
- Used only one environment (Highway) with limited number of experiments to show the stability of their algorithm

Final Reflections on LTA

LTA is

- Differentiable
- Simple to use and implement
- Able to control the sparsity level

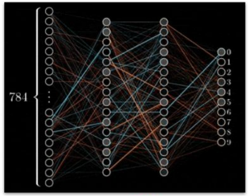
However, it

- Has mixed results on its efficacy, especially in environments with continues actions spaces
- Is not clear whether it causes dead neurons.
- Needs a parameter study because it introduces 4 new parameters

Conclusion

A Better Approach: Injecting Sparsity!

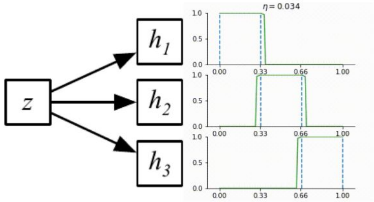
- A small number of features are **active**, for each input.
- Each update only **impacts** a small number of weights
- So it is less likely to **interfere** with many state values.



Source: Link

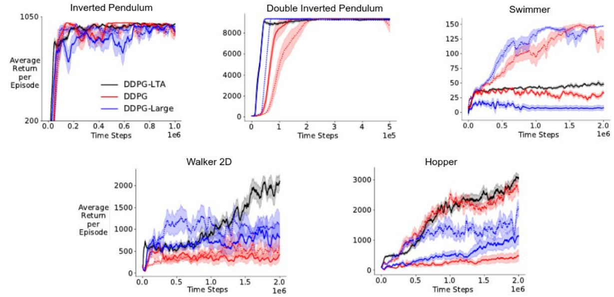
Leaky Tiling Activation

- Introduces a new parameter: η
- This parameter determines the sparsity level
- Or the level of leakage from one bin to the neighboring bins



Pan, Yangchen, et al. Leaky Tiling Activations: A Simple Approach to Learning Sparse Representations Online. International Conference on Learning Representations, 2021.

Overall Performance (Continues Action Space)



The **dotted line** indicates algorithms trained **with target networks.**

Pros and Cons of their Experimental Methodology

Pros:

- Used a good variety of environments to evaluate LTA
- Ran their algorithm for 20 number of runs
- Reported all the parameters

Cons:

- Did not sweep over target network update frequency
- Were not clear about some of their choices:
  - Switching from 20 runs to 30 runs in highway
  - Usage of LTA in DDPG
  - Usage of DDPG neither in Highway or when comparing it with other sparse techniques
- Stopped reporting performance of DQN-LTA with target network
- Inacceptable approach for sweeping over L1 and L2 parameters
- Used only one environment (Highway) with limited number of experiments to show the stability of their algorithm