Weight Agnostic Neural Networks

Author: Gaier et al.

created by paper2slides

Uploaded to arxiv on 2019-06-11

https://arxiv.org/abs/1906.04358

Executive Summary

- Investigates the importance of neural network architecture compared to weight parameters.
- Proposes a method to find architectures that perform tasks well with randomly assigned weights, termed Weight Agnostic Neural Networks (WANNs).
- WANNs are found to perform several reinforcement learning tasks and MNIST classification without weight training, displaying a high innate capability.
- Further performance improvements are possible with simple weight tuning or full weight training.

Introduction

- Biology shows some species are born with innate abilities, suggesting architecture can encode behavior.
- Conventional methods focus on training weights within fixed architectures.
- This work explores if architectures alone, without learning specific weights, can solve tasks.

Weight Agnostic Neural Networks (WANNs)

- Search for architectures to perform tasks well across a range of shared weight values.
- Using a single shared weight value for all connections of the network during evaluation.
- Performance is assessed over several weight values to emphasize architecture robustness.

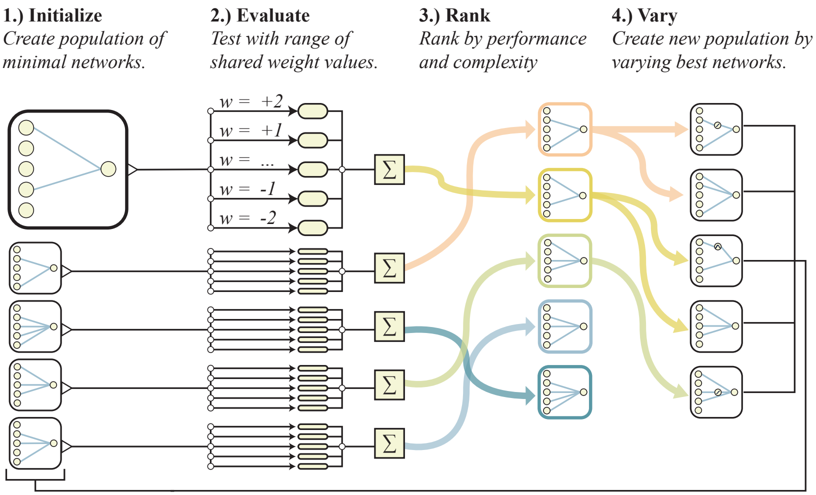

Figure: Weight Agnostic Neural Network search method overview.

Topology Search

- Inspired by NEAT algorithm for evolving neural networks.
- Starts with minimal networks and evolves by adding nodes, connections, or changing activation functions.
- Activation functions include varied types (e.g., linear, ReLU, Gaussian).

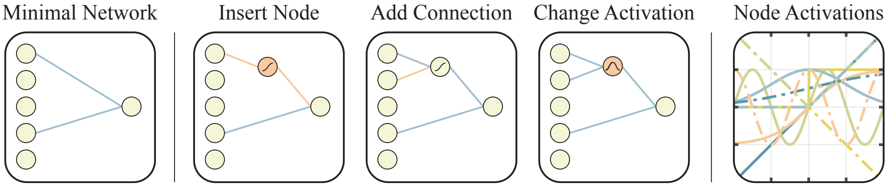

Figure: Operators for searching network topologies.

Advantages of WANN Approach

- Simplifies the evaluation process by ignoring individual weight optimization.
- Encourages finding architectures with strong inductive biases.
- Produces networks that perform well even without training specific weights.

Experimental Results: Continuous Control

- Evaluated on CartPoleSwingUp, BipedalWalker-v2, and CarRacing-v0 tasks.
- WANNs compared with hand-designed networks.
- Performance assessed with random weights, random shared weight, tuned shared weight, and fully tuned weights.

Results: Continuous Control Tasks

- WANNs perform better than fixed topologies with random shared weights.
- Significant performance improvement with tuned shared weights.
- Comparable to state-of-the-art after full weight tuning.

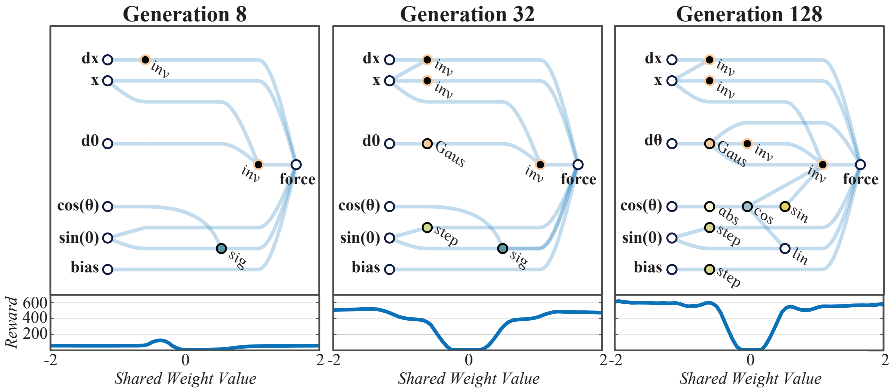

Figure: Development of WANN topologies over generations for CartPoleSwingUp.

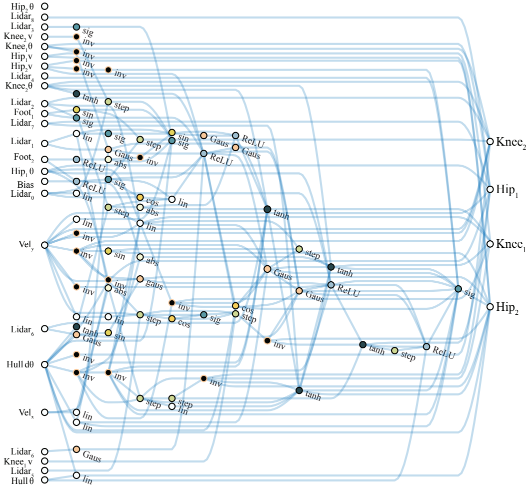

Figure: Champion network for BipedalWalker-v2.

Experimental Results: MNIST Classification

- Evaluated WANN approach on the MNIST dataset, a high-dimensional image classification task.
- Compared WANNs' performance with that of a single layer trained by gradient descent.

Results: MNIST Classification

- WANNs perform well with random weights, significantly better using an ensemble of weight values.
- Accuracy comparable to gradient-based methods with ensemble approach.

WANN	Test Accuracy
Random Weight	*82.0% ± 18.7%*
Ensemble Weights	*91.6%*
Tuned Weight	*91.9%*
Trained Weights	*94.2%*

ANN	Test Accuracy
Linear Regression	*91.6%* [62]
Two-Layer CNN	*99.3%* [15]

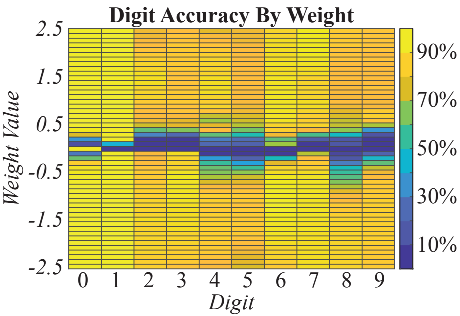

Figure: Classification accuracy on MNIST dataset using WANNs.

Conclusion and Future Work

- Demonstrated that neural network architectures with strong inductive biases can solve tasks without weight training.
- Encourages further research exploring novel neural network building blocks.
- Potential applications in few-shot learning, continual learning, and development of new building blocks.