# AI Model Testing Fuzzing: Research Project Proposal

Yuqing Yang

Nanjing University

November 30, 2018

## Outline

1 Backgrounds
- AI fuzzing
  - Fuzzing for AI
  - AI for fuzzing
- Testing
- Coverage
- CNN,RNN

2 Related Works
- DeepXPlore
- DeepGauge
- ReluVal

3 Research Goal

4 Refereences

## AI fuzzing

Fuzzing:

- an automated software **testing** technique
- provides invalid, unexpected, or random data as inputs to a computer program
- monitors exceptions such as crashes, failing built-in code assertions, or potential memory leaks

Two main categories:

- Fuzzing for AI
- AI for fuzzing

**Fuzzing for AI**
Target: AI Components

- neuron coverage[1]
- layer coverage[2]
- formal security method[3]

[1] [Kexin Pei et al., 2017]DeepXplore: Automated Whitebox Testing of Deep Learning Systems

[2] [Lei Ma et al., 2018]DeepGauge: Multi-Granularity Testing Criteria for Deep Learning Systems

[3] [Shiqi Wang et al., 2018]Formal Security Analysis of Neural Networks using Symbolic Intervals

**AI for fuzzing**

Method: AI

- RNN-based(LSTM+AFL)[4]
- CNN-based(CNN+gradient descent)[5]
- RL(Q-Learning)[6]

[4] [Mohit Rajpal et al., 2017]Not all bytes are equal: Neural byte sieve for fuzzing

[5] [Dongdong She et al., 2018]NEUZZ: Efficient Fuzzing with NeuralProgram Smoothing

[6] [Konstantin Bottinger et al., 2018]Deep Reinforcement Fuzzing

## Testing

- Blackbox, testing functions without peering into internal structures or workings
- Whitebox, testing internal structures or workings of an application
- Greybox, tests improper structure-caused defects, if any

**Coverage**

**Software testing measurement** for describing the degree to which the source code of a program is executed

- Edge Coverage
- Function Coverage
- Statement Coverage

**CNN,RNN**

Short view:

- MLP: Simplest DNN with fully-connected layers
- CNN: +Hypo:Space-correlation, everywhere in CV
- RNN: +Hypo:Time-correlation, usually used in Speech Analytics

## Outline

1 Backgrounds
   - AI fuzzing
     - Fuzzing for AI
     - AI for fuzzing
   - Testing
   - Coverage
   - CNN,RNN

2 Related Works
   - DeepXPlore
   - DeepGauge
   - ReluVal

3 Research Goal

4 Refereences

## DeepXPlore[1]

- Neuron coverage: coverage of neurons with outputs exceeding preset thresholds
- Goal: Optimize neuron coverage
- How: Gradient Descending aiming to find maximal value

[1] [Kexin Pei et al., 2017]DeepXplore: Automated Whitebox Testing of Deep Learning Systems

## DeepGauge[2]

- Neuron coverage is not enough:
  - k-multisection Neuron Coverage
  - Neuron Boundary Coverage (Corner Region Coverage)
  - Strong Neuron Activation Coverage (Corner Case Coverage)
- Layer coverage:
  - Top-k Neuron Coverage
  - Top-k Neuron Patterns

[2] [Lei Ma et al., 2018]DeepGauge: Multi-Granularity Testing Criteria for Deep Learning Systems

## ReluVal[3]

- Formal Security: Mathematically declared secure properties
- Goal: Achieve a exhaustive, high-performance analysis method
- How: Symbolic intervals and Interval analysis

[3] [Shiqi Wang et al., 2018]Formal Security Analysis of Neural Networks using Symbolic Intervals

## Outline

1 Backgrounds
- AI fuzzing
  - Fuzzing for AI
  - AI for fuzzing
- Testing
- Coverage
- CNN,RNN

2 Related Works
- DeepXPlore
- DeepGauge
- ReluVal

3 Research Goal

4 Refereences

- Goal: To explore new efficient way of fuzzing for AI components
- Target: Exsisting AI components
- How: Explore by adopting, analysing, optimizing exsisting fuzzing methods
- How: Optimize by combining suitable AI methods

## Outline

1 Backgrounds
- AI fuzzing
  - Fuzzing for AI
  - AI for fuzzing
- Testing
- Coverage
- CNN,RNN

2 Related Works
- DeepXPlore
- DeepGauge
- ReluVal

3 Research Goal

4 Refereences

- [Kexin Pei et al., 2017]DeepXplore: Automated Whitebox Testing of Deep Learning Systems
- [Lei Ma et al., 2018]DeepGauge: Multi-Granularity Testing Criteria for Deep Learning Systems
- [Shiqi Wang et al., 2018]Formal Security Analysis of Neural Networks using Symbolic Intervals
- [Mohit Rajpal et al., 2017]Not all bytes are equal: Neural byte sieve for fuzzing
- [Dongdong She et al., 2018]NEUZZ: Efficient Fuzzing with NeuralProgram Smoothing
- [Konstantin Bottinger et al., 2018]Deep Reinforcement Fuzzing