Journal First Track

An annotation-based approach for finding bugs in neural network programs

Journal of Systems and Software

Mohammad Rezaalipour, Carlo A. Furia

Software Institute – USI Università della Svizzera italiana
Lugano, Switzerland

Neural Networks are implemented programs


```
def dense_block(
   x, nb layers, nb channels, growth rate,
   dropout_rate=None, bottleneck=False,
):

   x_list = [x]
   for i in range(nb layers):
       cb = convolution block(x, growth rate,
                              Dropout rate,
                              bottleneck)
       x_list.append(cb)
       x = Concatenate(axis=-1)(x_list)
       nb channels += growth_rate
   return x, nb_channels
```

Written by domain experts who may not be professional programmers

- An empirical study on TensorFlow program bugs *(Zhang et al. 2018)*
- A Comprehensive Study on Deep Learning Bug Characteristics *(Islam et al. 2019)*
- Taxonomy of Real Faults in Deep Learning Systems *(Humbatova et al. 2019)*

"A noticeable percentage [...] threw runtime exceptions due to code or script defects [...]"

Objective:

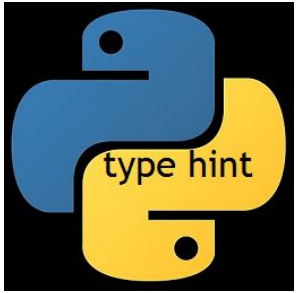
Automated Test Generation

Python is dynamically typed language

Complicated parameter types

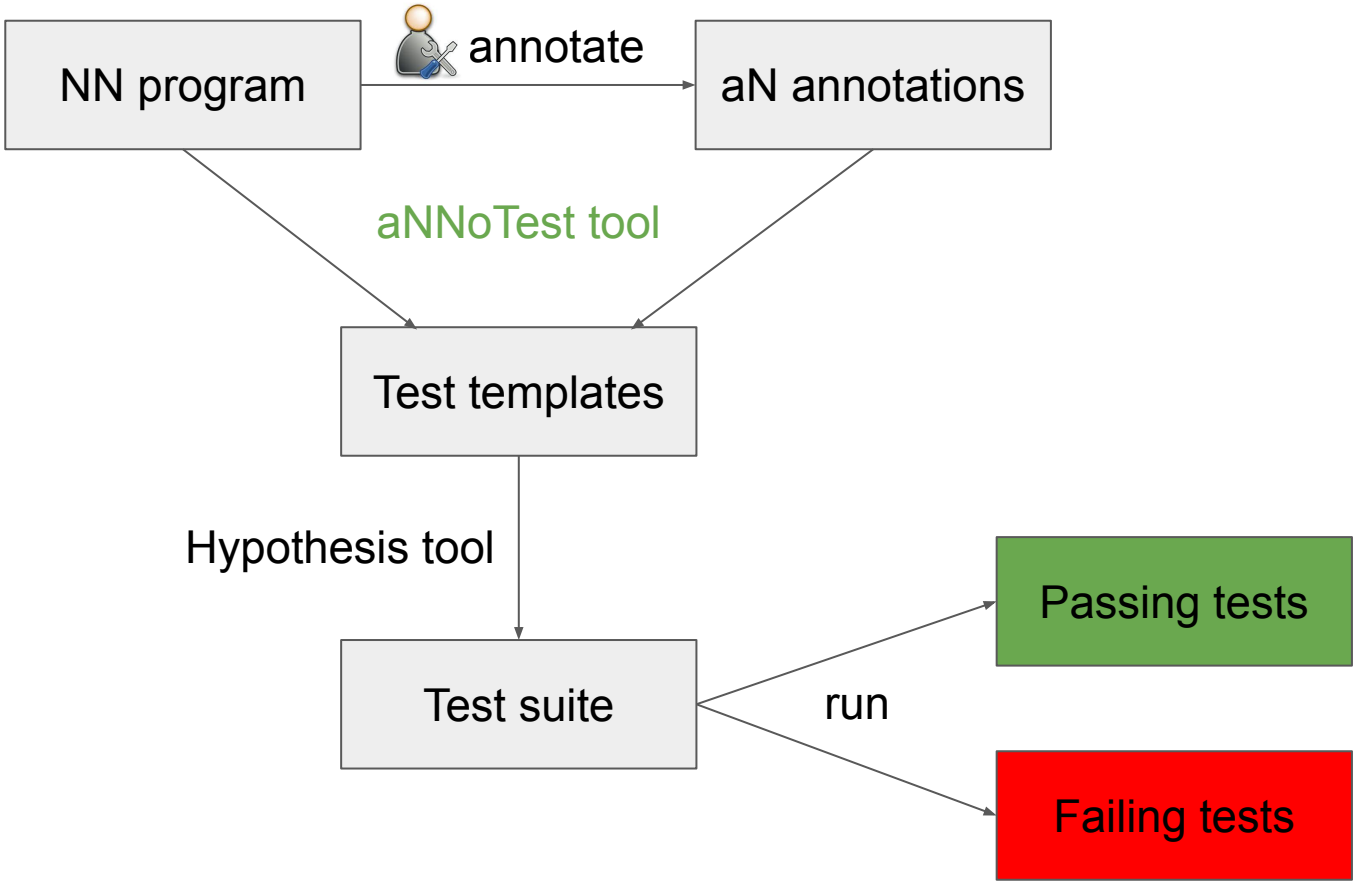
```
def DenseNet(input_shape=None, dense_blocks=3, dense_layers=-1,
             growth_rate=12, nb_classes=None, dropout_rate=None,
             bottleneck=False, compression=1.0, weight_decay=1e-4,
             depth=40):
```

https://github.com/cmasch/densenet/


type hint


Pynguin


DEAL


NN program
annotate
aN annotations
aNNoTest tool
Test templates
Hypothesis tool
Test suite
run
Passing tests
Failing tests

```
def DenseNet(input_shape=None, dense_blocks=3, dense_layers=-1,
             growth_rate=12, nb_classes=None, dropout_rate=None,
             bottleneck=False, compression=1.0, weight_decay=1e-4,
             depth=40):

    if nb_classes==None:
        raise Exception('Please define number of classes.')

    if compression <=0.0 or compression > 1.0:
        raise Exception('Compression must be between 0.0 and 1.0.')

    if type(dense_layers) is list:
        if len(dense_layers) != dense_blocks:
            raise AssertionError('Dense blocks must be the same as layers')
    elif dense_layers == -1:
        dense_layers = (depth - 4)/3

     ...
```

Bug

https://github.com/cmasch/densenet/

```
@arg(input_shape): tuples(ints(min=20, max=70),
                          ints(min=20, max=70),
                          ints(min=1, max=3))

@arg(dense_blocks): ints(min=2, max=5)
```

Experimental Evaluation

- RQ1. Precision
- RQ2. Recall
- RQ3. Amount of Annotations
- RQ4. Comparison to Generic Test-Case Generators
- RQ5. Code Coverage

Experimental Subjects

Projects	LOC	Total Functions	Tested Functions	Avg. Number of Annotations	Known Bugs
2	3917	249	105	1.33	-
19	14219	735	24	6.00	81

Experiment Results

Projects	LOC	Known Bugs	Found Bugs	Spurious	Precision	Recall
2	3917	-	50	6	89%	-
19	14219	81	63	0	100%	78%

An annotation-based approach for finding bugs in neural network programs

Mohammad Rezaalipour, Carlo A. Furia

mohrez86.github.io

aNNoTest Tool Demo:

Fri. 6 Oct. (11:13 AM)

Software Testing 2 - Session 1 Room

https://github.com/atom-sw/annotest