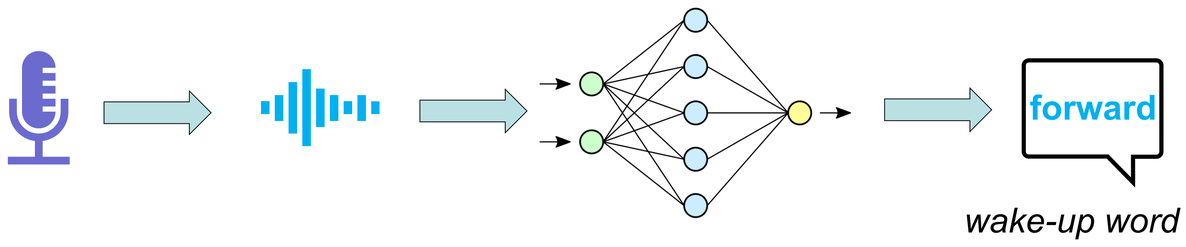
Towards Verification-Aware Neural Networks for Speech Recognition

Syed Ali Asadullah Bukhari, Barak A. Pearlmutter and Rosemary Monahan
Department of Computer Science
National University of Ireland, Maynooth

1. Overview

Motivation:

- Speech recognition is one of the most widely used AI applications
- On-device recognition:
  - Uses a scaled-down and quantized model
  - Works on a limited input subset
  - More prone to perturbations in inputs



Must be formally verified against robustness

Challenge:

- Evaluating verification tools for a custom application
- Adapt neural network for verifiers' compatibility

2. Case Study

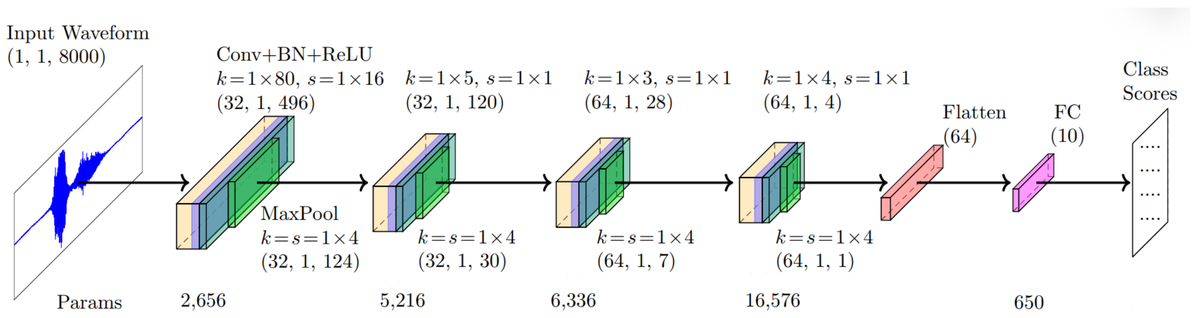
- Use neural network verification tools to verify robustness of speech recognition neural networks against input perturbations

Data Preprocessing:

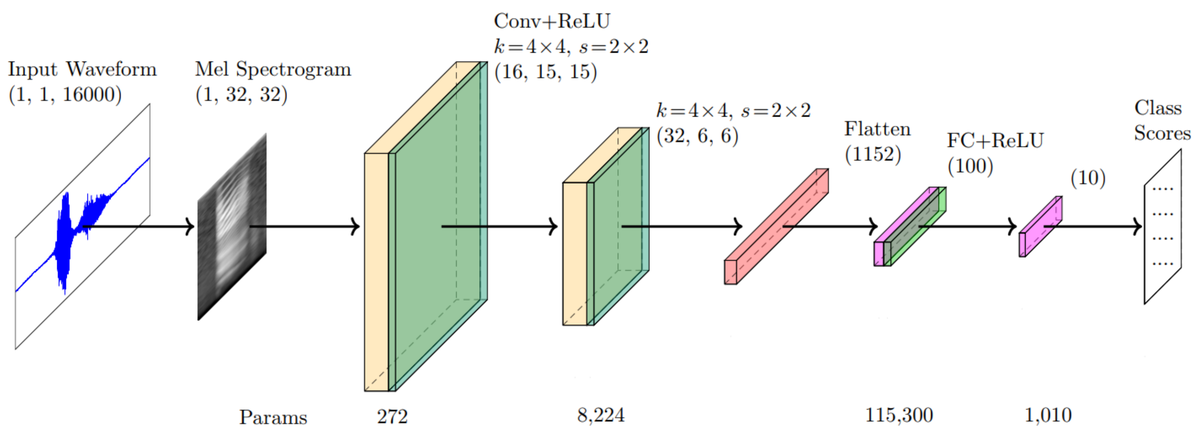
- Digits (0-9) subset from Google Speech Command Dataset
- Conversion to Mel Spectrogram → Image Classification
- Down sampling from 16KHz to 8KHz → Raw waveform classification

Neural Networks:

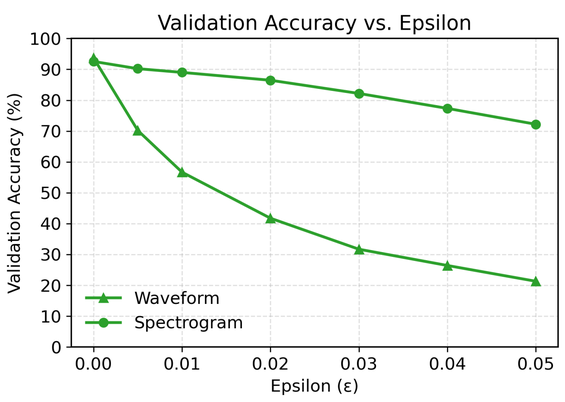
- Waveform Classification

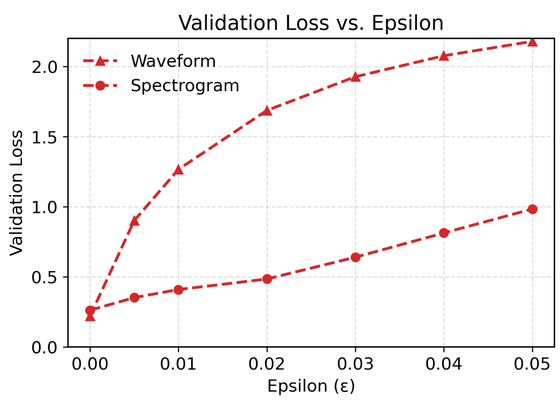


- Spectrogram Classification



PGD-Based Adversarial Training:





3. Verification Property

Local Robustness:

$$(x_0, \varepsilon, \delta) \coloneqq \forall x. ||x_0 - x||_\infty \leq \varepsilon \Rightarrow ||\mathcal{N}(x_0) - \mathcal{N}(x)||_\infty \leq \delta$$

- For slight perturbations x within some bound ε of a given input x_0, the neural network $\mathcal{N}$ should give roughly the same output,

Robustness Property in VNNLIB format:

```
; I/P, O/P Declaration
(declare-const X_0 Real)
(declare-const X_1 Real)
.....
(declare-const Y_0 Real)
(declare-const Y_1 Real)
......

; Input Assertions
(assert (<= X_30 0.18725))
(assert (>= X_30 0.08725))
.....
```

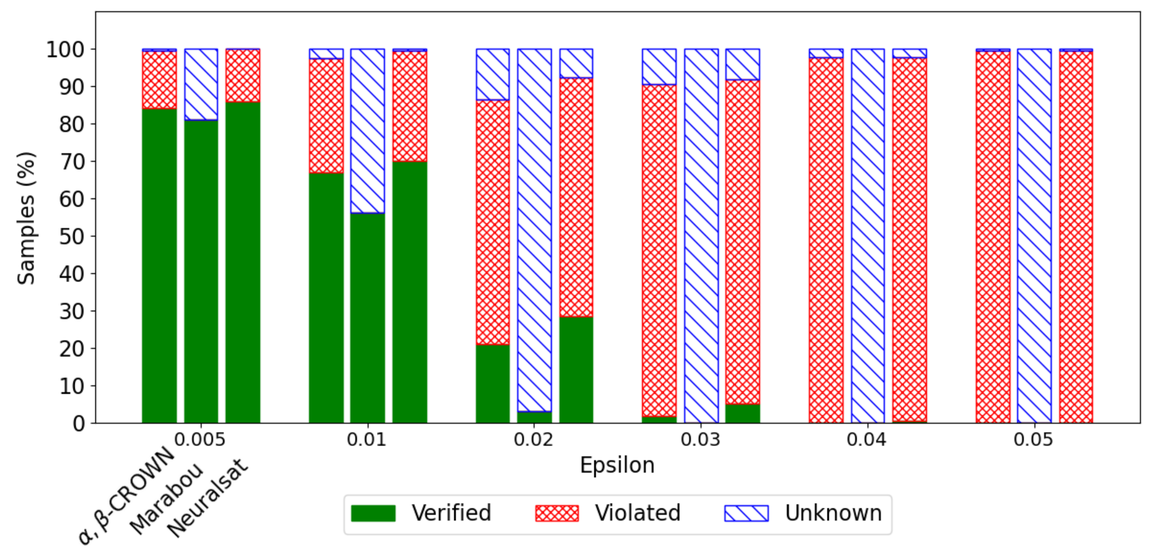
```
; Output Assertions
(assert (or
    (and (>= Y_1 Y_0))
    (and (>= Y_2 Y_0))
    (and (>= Y_3 Y_0))
    (and (>= Y_4 Y_0))
    (and (>= Y_5 Y_0))
    (and (>= Y_6 Y_0))
    (and (>= Y_7 Y_0))
    (and (>= Y_8 Y_0))
    (and (>= Y_9 Y_0))
))
```

4. Verification Results & Discussion

Verification Setup:

- Neural Network Verifiers: α, β-CROWN, Marabou, Neuralsat
- Verification Samples: 1000 (100 per class)
- Timeout: 60 secs
- Model format: onnx
- Property specification: vnnlib

Performance Comparison of Neural Network Verifiers:



Findings:

- α, β-CROWN and Neuralsat perform better than Marabou
- Challenging to formalize robustness with complex data preprocessing
- Tools Limitations:
  - Limited support for 1D convolution layers
  - Restricted network design:
    - Only square kernels, no asymmetric pooling layer
- Ease-of-Use Issues
  - Resource-intensive
  - No standard interface

Future Work:

- Design of Verification-Aware Neural Networks
- Adaption of Neural Network Verifiers for LLMs

Funding

Taighde Éireann Research Ireland

Contact

ali.bukhari@mu.ie
barak.pearlmutter@mu.ie
rosemary.monahan@mu.ie