# SHE: A Fast and Accurate Deep Neural Network for Encrypted Data

Qian Lou, Lei Jiang
Indiana University Bloomington, USA

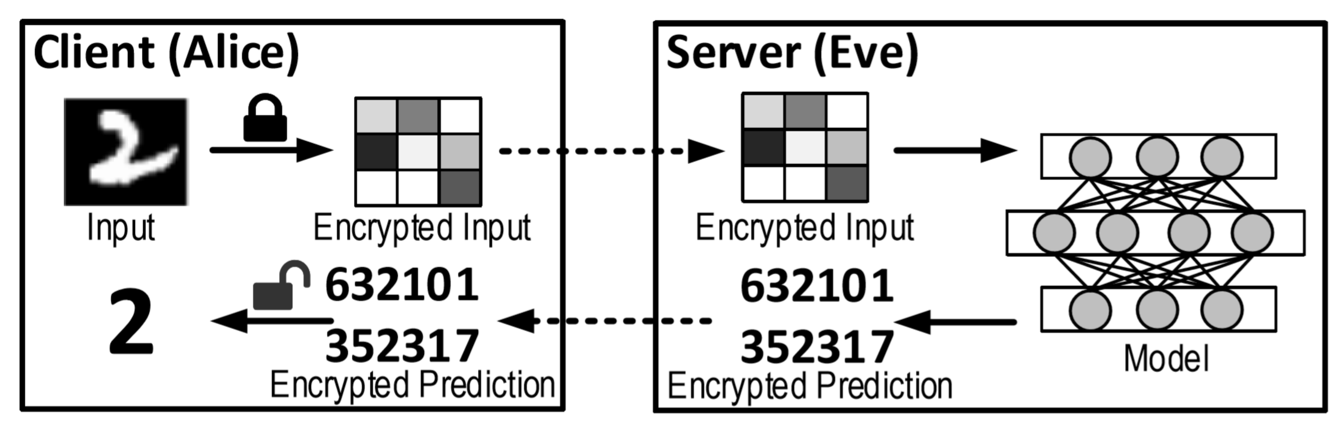
ICML | 2019

## Executive Summary

- **Need**: Fast and accurate deep Learning over encrypted data
- **Opportunities** to improve privacy-preserving deep learning by the co-design of Homomorphic Encryption scheme and neural network optimization:
  - Binary bits-operations-friendly TFHE encryption scheme
  - Shift-Accumulation based quantization for neural network
- **Problem**: Previous works stacked **multiple** & **inaccurate** ReLU activation and max pooling layers (Polynomials approximation): Accuracy ↓ & overhead ↑ & shallow networks topology
- **Key Idea**: Directly implementing ReLU and max using TFHE [1]; Using cheap Shift-Accumulation to support deeper neural networks other than acceleration.
- **SHE**: Accuracy-lossless CNN, performance ↑76.12%, the first to support modern deep learning like AlexNet on MNIST.

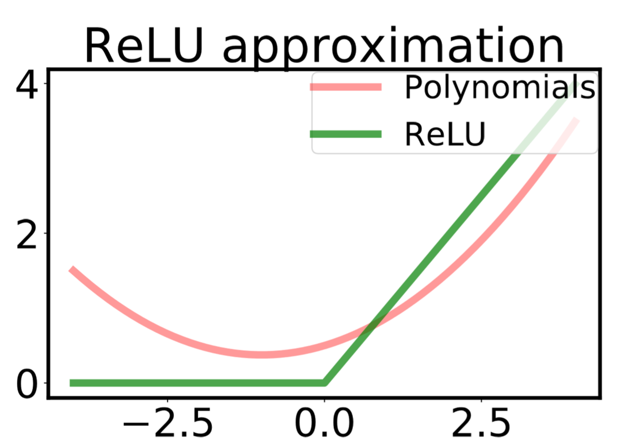
## Introduction

- **Thread Model**
  - Untrusted servers may lead to data leakage where the data from client-side users;
  - Results sent to clients from servers be illegally utilized by adversarial.
- **Private Neural Networks by Homomorphic Encryption**

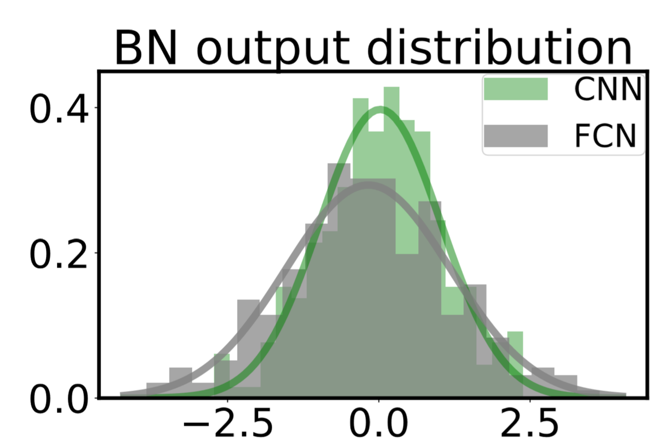



- Untrusted servers learn encrypted data and output encrypted prediction
- Only clients with private key can decrypt the encrypted prediction

## Problems

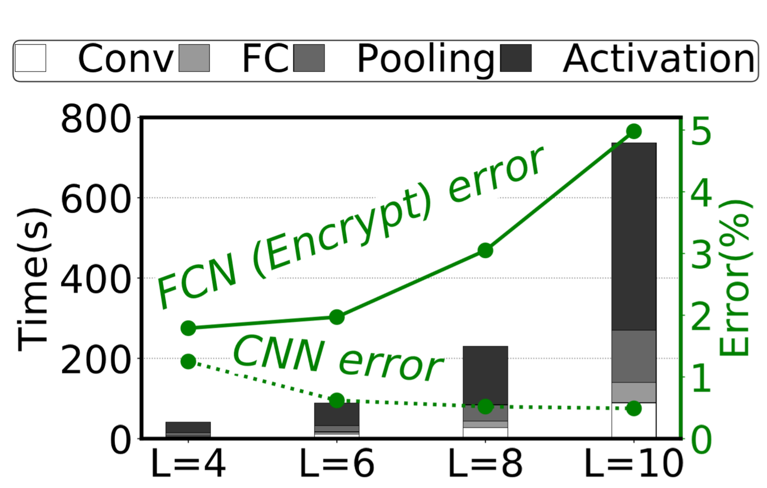



Inaccurate ReLU approximation




Distribution distortion

- Traditional Homomorphic Encryption schemes (B/FVs and HEAAN) uses polynomials like y=0.125x^2+0.25X+0.5 to approximate ReLU with large **errors**. Stacked multiple such inaccurate layers bring distribution **distortion** of latent variables.




More errors with deeper neural network

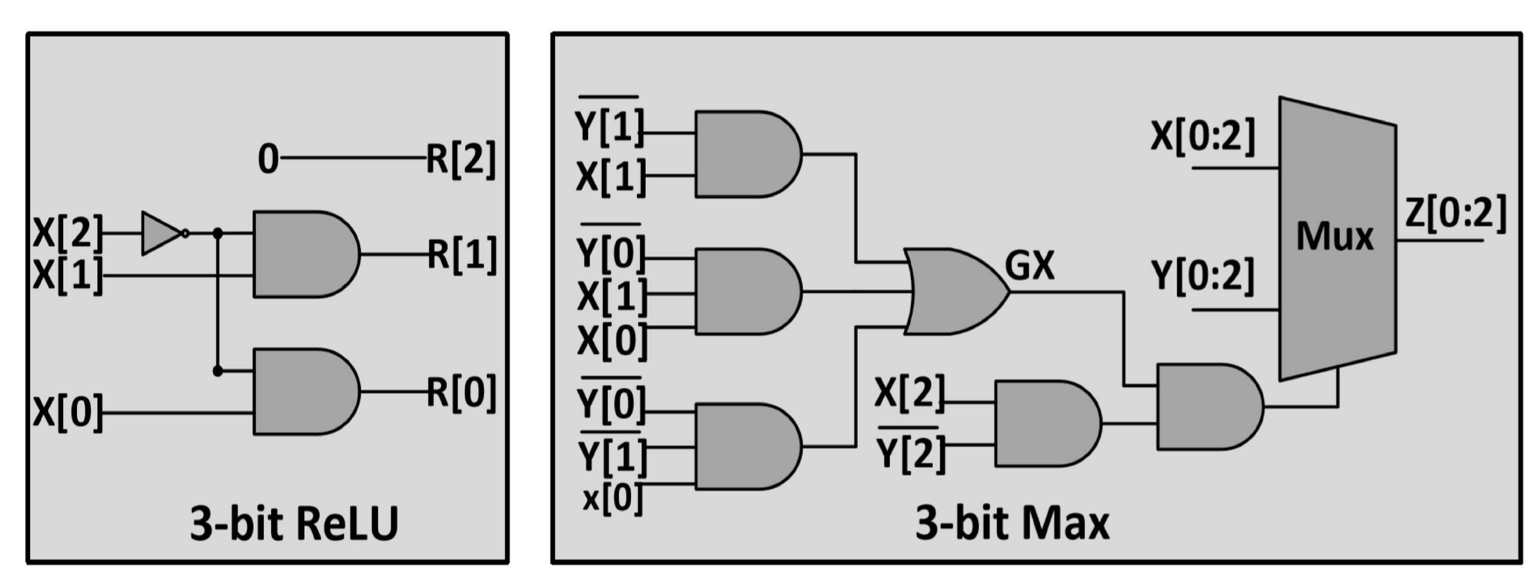
| Name | Total Depth | Accuracy (%) |
|---|---|---|
| FCN[3] | 21K | 98.71 |
| DiNN[5] | 0.8K | 93.34 |
| SHE | 2.0K | 99.54 |
| DSHE | 6.2K | 99.77 |

Prior works require bigger circuits depth

- Because of the distribution distortion of latent variables, even deeper neural network leads to more errors. Also, previous works like CNT [2], FCN [4] and NED [3] requires bigger circuits depth than SHE.

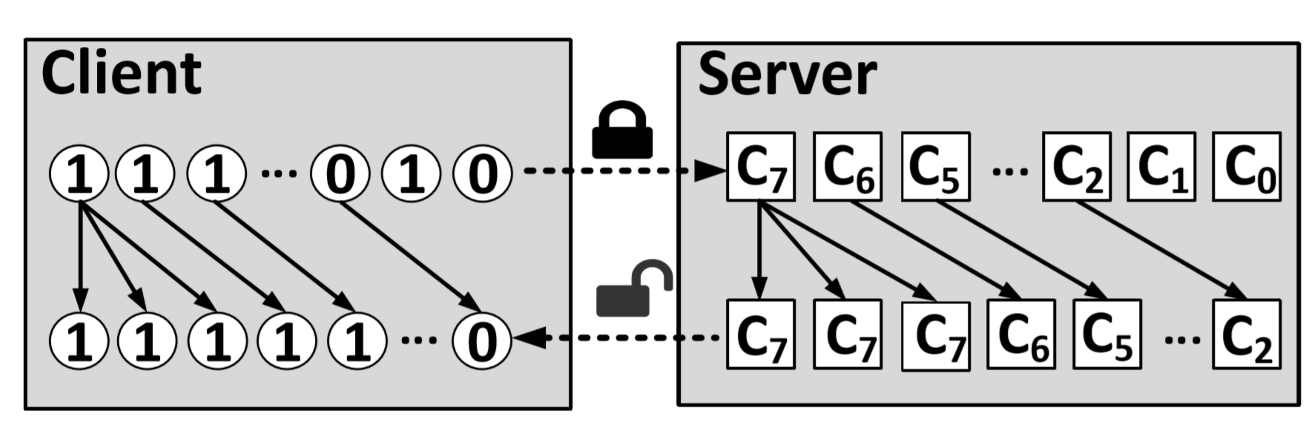
## SHE Overview

- **SHE**
  - Support **accuracy-lossless** ReLU and Max ⟶ **Accurate**
  - Logarithmic Quantization: Convolution to Shift-Accumulation ⟶ **Fast**
  - TFHE scheme (Binary bits-operations and shift-operations friendly) ⟶ **fast & deeper neual networks.**

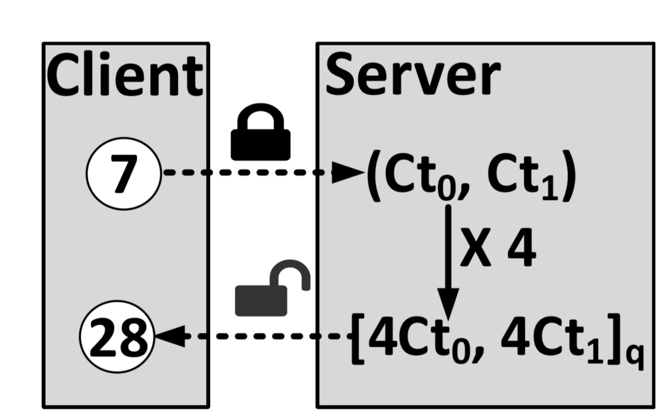



- Using TFHE scheme to implement ReLU and Max
- Left figures show 3-bit ReLU and Max

Accuract-lossless 3-bit ReLU and Max




A TFHE shift

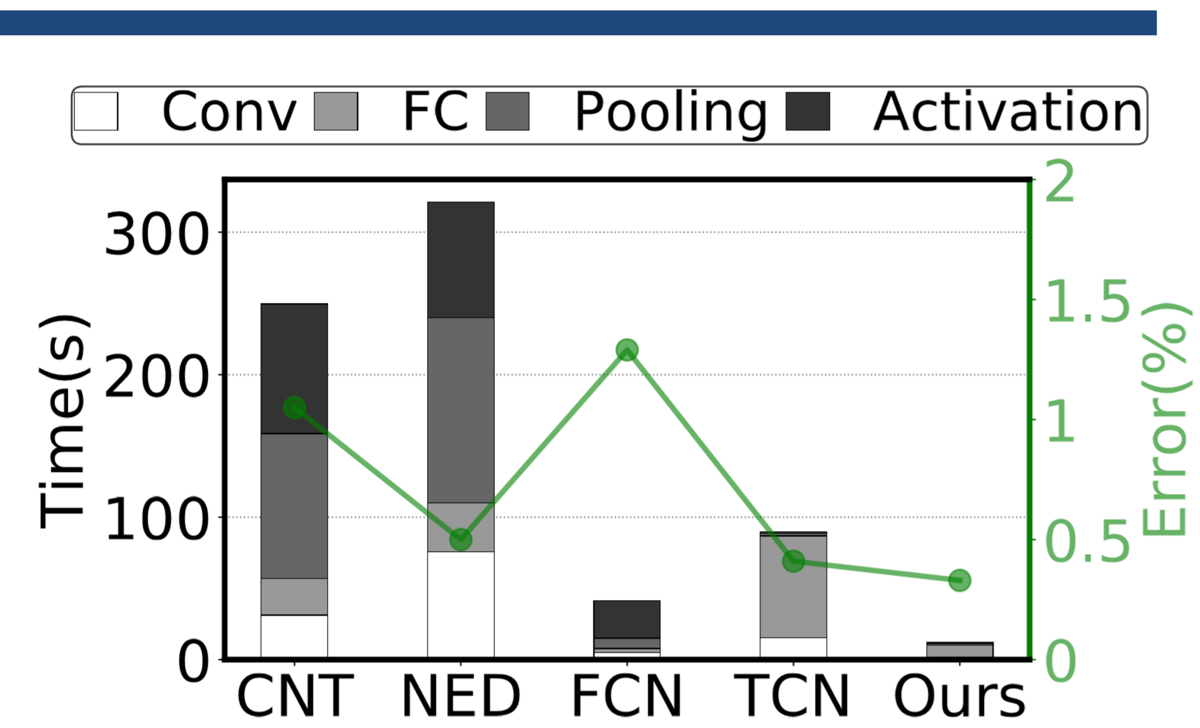



A shift in other schemes

## Result

| Name | Support ReLU | Support Max Pooling | No Convs | Deep Neural networks |
|---|---|---|---|---|
| CNT[2] | ✗ | ✗ | ✗ | ✗ |
| NED[3] | ✗ | ✗ | ✗ | ✗ |
| FCN[4] | ✗ | ✗ | ✗ | ✗ |
| DiNN[5] | ✗ | ✗ | ✓ | ✗ |
| SHE | ✓ | ✓ | ✓ | ✓ |

The comparison between previous works and our work




The performance and accuracy comparisons

## References


[1] Chillotti, et al. Tfhe: Fast fully homomorphic Encryption over the torus. IACR Cryptology ePrint Archive, 2018

[2] Dowlin, et al. Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy. In ICML 2016

[3] Chou et al. Faster cryptonets: Leveraging sparsity for realworld encrypted inference. arXiv 2019.

[4] Hesamifard, et al. Deep neural networks classification over encrypted data. In ACM CDASP 2019.

[5] Bourse,et al. Fast homomorphic evaluation of deep discretized neural networks. In CRYPTO 2018.


ISE IU Intelligent Systems Engineering

**Qian Lou**
Email: louqian@iu.edu, Phone:+18125586704