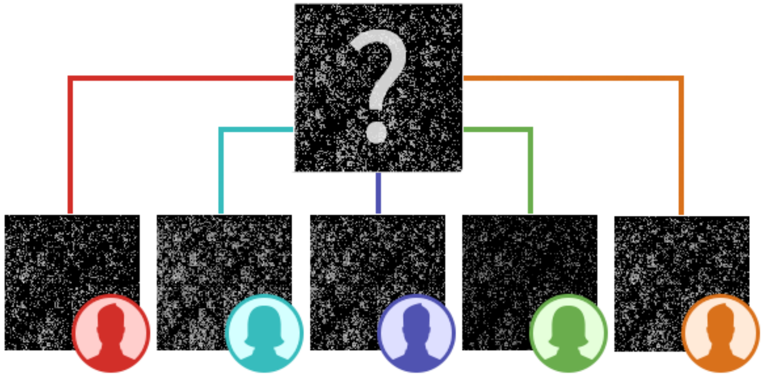
Discovery and Scaling of Text Models for Authorship Attribution

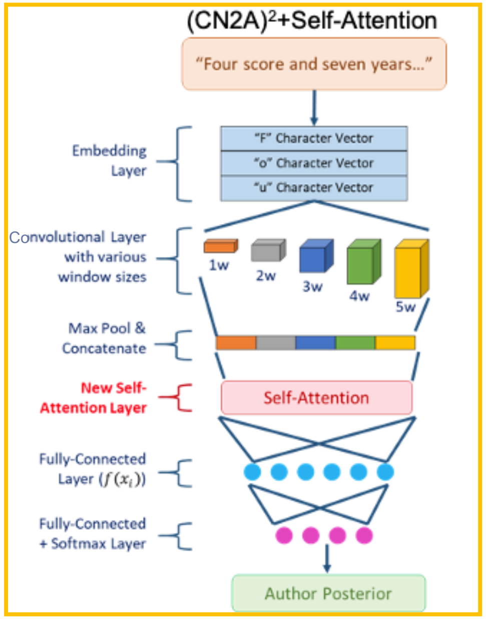
Porter Bagley
Computing - DSSI
Collaborators: Juanita Ordonez, Rafael Rivera Soto

LLNL-PRES-813215
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Introduction

- Motivation:
  - As more of us rely on online sources for news and information, it becomes increasingly important to verify their authenticity
  - Authorship attribution can help us determine trustworthiness



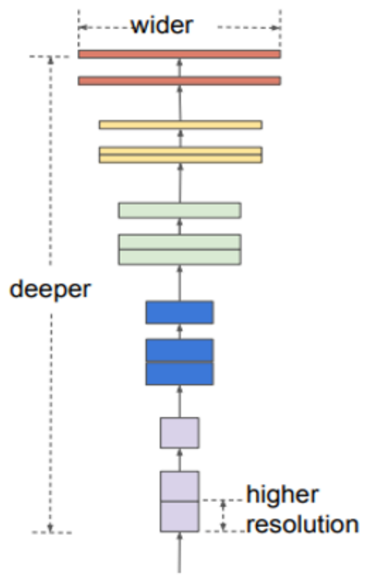


Our team created a CNN architecture which used Attention and performed well

- **Questions:**
  - How to scale a CNN model strategically?
  - Is this the best architecture?

Methods

- Compound Scaling
  - "EfficientNet: Rethinking Model Scaling for CNNs" (Tan & Le, 2019)
  - Rather than searching for the best dimensions with a huge model, find the best dimensions on a small model, then scale up proportionally



- We apply the methods from EfficientNet to text-based data

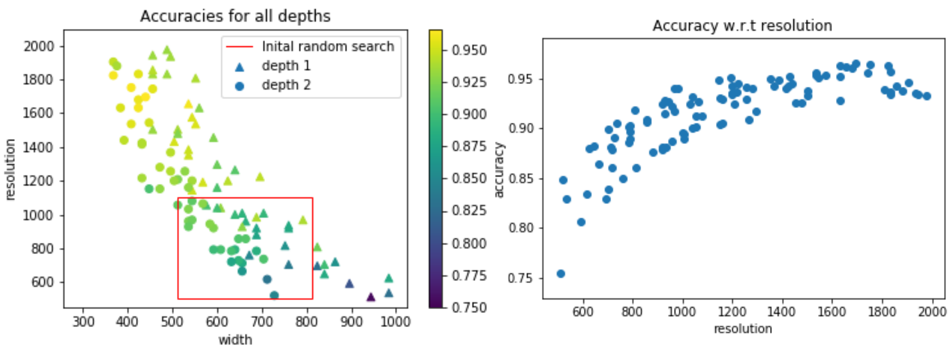
$$
\begin{aligned}
\text{depth: } & d = \alpha^{\phi} \\
\text{width: } & w = \beta^{\phi} \\
\text{resolution: } & r = \gamma^{\phi} \\
\text{s.t. } & \alpha \cdot \beta^2 \cdot \gamma \approx 2 \\
& \alpha \geq 1, \beta \geq 1, \gamma \geq 1
\end{aligned}
$$

- Dataset
  - 100 authors from Reddit
  - Documents of 100-500 characters
  - Concatenated according to timestamp and subreddit to make 2000 & 5000 character length datasets

Results and Conclusion

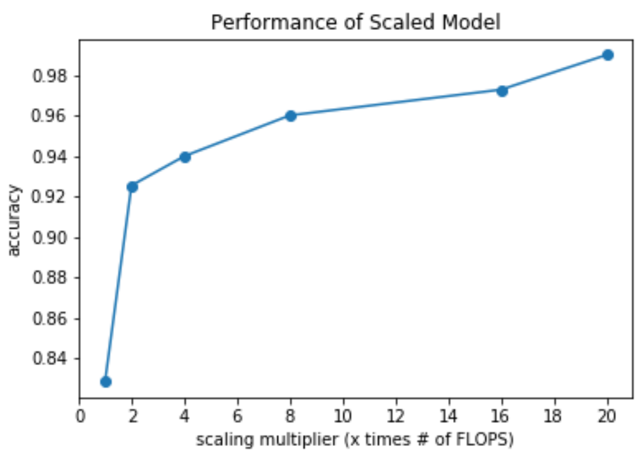
- Results

Random search results:



Top model results:

Scaling amount	Average Author Accuracy
1.0 (original)	0.8286
2.0	0.9254
4.0	0.9399
8.0	0.9601
16.0	0.9728
20.0	0.9900



- Conclusion
  - Compound scaling is effective for text-based CNNs
  - Increasing the resolution of text-based data can lead to large gains in performance
  - Achieved 99.0% (+16.1%) accuracy on authorship attribution task

- Next Steps
  - Discover alternate architectures using TextNAS (Wang et al., 2019)
  - Apply similar methods to open-world authorship attribution task

Lawrence Livermore National Laboratory