## **Andrew Smith**

- Application of Artificial Neural Networks (ANN) to Automatic Speech Recognition (ASR) on a Novel Dataset created using You Tube
- Approach involves using Connectionist Temporal Classification and LSTM units similar to the architecture of Google's Deep Speech 2.

0000000

0000000

0000000

0000000

0000000

0000000

0000000

0000000

0000000

Spectrogram

Recurrent

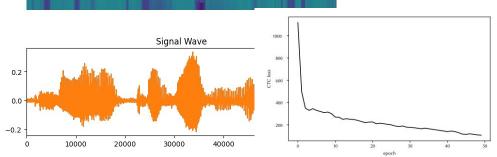
GRU

(Bidirectional)

Invariant Convolution

- Data were obtained from YouTube, transformed using Discrete Fourier Transformation, and partitioned accordingly for supervised learning.
- The model with approximately 26 million parameters was trained on a 3090ti.
- 50 epochs of training were conducted on 10 hours of speech data achieving a WER of 66.9.
- Size of model occupied GPU memory and forced small batch size and increased length of training therefore constraining the feasible size of the dataset
- Could have just used Deep Speech 2 and evaluated it on the novel You Tube dataset, but I wanted to learn about CTC and train the model myself, so it was a nice investigation.

so the observable universe is the distance that light has



- Target: in eighteen eleven williams who murdered the marrs in ratcliffe highway having committed suicide in jail to escape hanging
   Prediction: in site a leven wilcome we murated the mars and
- Prediction: in aite a leven wiloms wo murered the mars and relit hywa hing commisted suasid in jail to sca aniing
- Target: on his release an uncle a slopseller in chatham gave him a situation as barker or salesman
- Prediction: on his relys and oncal ay sloselr in shaterm ga himmasituation as barker or salsmon
- 5) Target: even after sentence and until within a few hours of execution he was buoyed up with the hope of reprieve
- (6) Prediction: the vin after senten sand intil within of forsobexcution he as boio with thu hal profrprve
- (7) Target: oswald's marine training
- (8) Prediction: oswalds moain traning
- (9) Target: these materials are used by the animal in the manufacture of new protoplasm to take the place of that which has been used up

  Prediction: these meterials are used by the animal in the many-

Prediction: thes meterials ar us by the animol in the manyfactur of n protoplasm to tak the place of that which has beedusdu

CSCE 771: Computer Processing of Natural Language