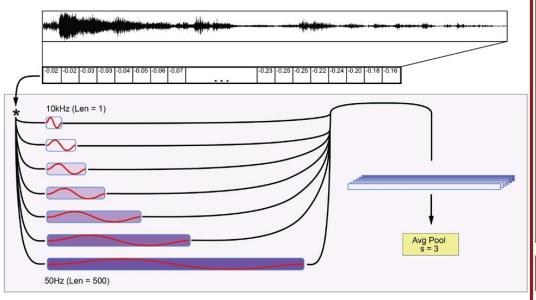
Dense Hair Net **CS230** Final Project

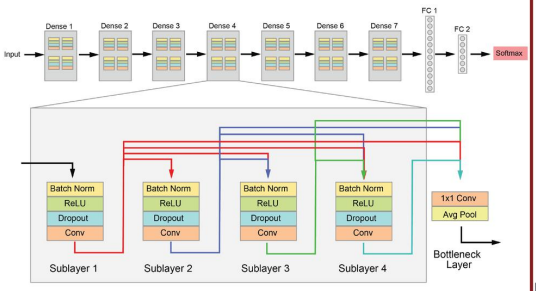
Exploration of Modern Convolutional Architectures using 1Dimensional Sound

Eric Loreaux eloreaux@stanford.edu
Yuri Zaitsev yz4321@stanford.edu

HairNets as an input[1]



1DenseNet[2] vs. ConvNet[3]



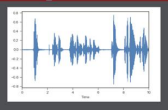
We revisit the use of 1D convolutional networks for auditory tasks using raw audio waveforms, and we attempt to boost their performance using techniques and architectures that we design with inspiration from some of the latest convolutional neural network research.

We design a new, Inception-style front-end component (HairNet) that parses raw audio using different filters. We also design an entirely distinct, DenseNet-style architecture (1DenseNet) that is intended to replace the standard 1D convolutional architecture. These designs are compared to a control benchmark CNN.

Both of these designs are intended to improve our capability to learn features from raw audio.

Data Set - Presidential Speech

Obama: "Good evening. As we speak, our nation faces a multitude of challenges. At home, our top priority is to recover..."



```
array([ 1.82474760e-05,  3.60148738e-06, -3.22135602e-05, ...,
       -1.21263368e-02, -7.64999213e-03, -4.43573901e-03], dtype=float32)
```

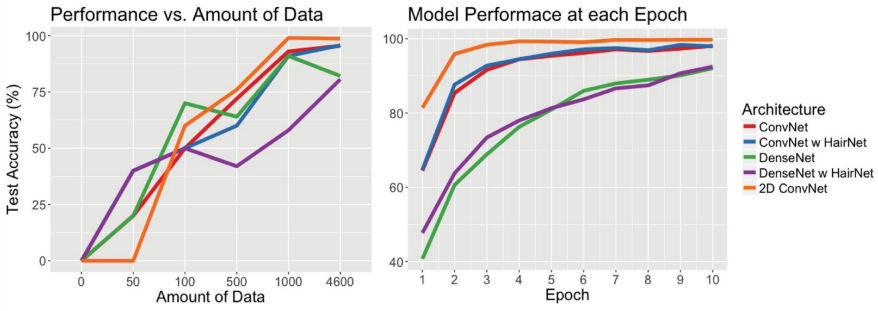
Dataset = Recorded speeches from the last 5 Presidents.
~ 3 hours of audio per President
Audio = 10 sec. clips | 16 bit mono 22050 Hz .WAV

Experimental Set Up

	w/o HairNet	HairNet
ConvNet	ConvNet	ConvNet w/ HairNet
DenseNet	DenseNet	DenseNet w/ HairNet

Trained for 10 epochs with 16 minibatch size.
Learning rate = 0.001
Use AdamOptimization

Results



2D ConvNets outperform 1D Nets,
but 1D Nets do not require data preprocessing.

1DenseNet did worse than ConvNet but is not as deep and has far fewer parameters. Can increase channels and layers to match size.

HairNet may be artificially limiting the scope of possible learned features by setting a certain filter size. Standard conv filter may learn more sophisticated features in the front end.

References

[1] C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 1-9.
doi: 10.1109/CVPR.2015.7298594

[2] G. Huang, Z. Liu, L. v. d. Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 2261-2269.
doi: 10.1109/CVPR.2017.243

[3] P. Rajpurkar, A. Hannun, M. Haghpanahi, C. Bourn, A. Ng, "Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks," arXiv:1707.01836, 2017.