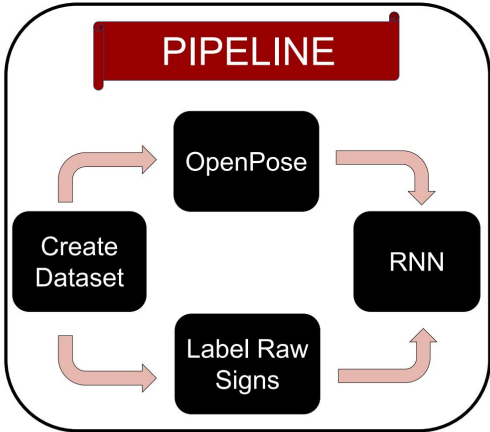




SignNet

Grant Fisher and Daniel Book
gfisher7@stanford.edu dbook@stanford.edu

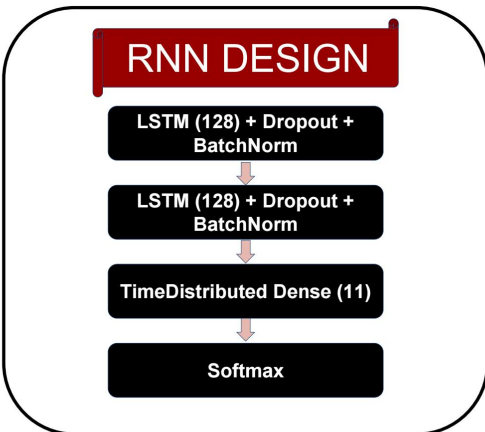


DATASET

20 videos per sign separated into sequences of still images.
~20,000 images total

OPENPOSE

Output: JSON file with 390 entries for pose, face, right hand, and left hand



RESULTS

After training the model for 20 epochs:

- Train Categorical Accuracy: 91%
- Test Categorical Accuracy: 90%
- Loss Function: Categorical Cross Entropy (0.322)

FUTURE

- More signs
- User interface
- “Alexa” for sign language

REFERENCES

- OpenPose
 - Z. Cao, T. Simon, S. Wei and Y. Sheikh, "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," 2017.
 - T. Simon, H. Joo, I. Matthews and Y. Sheikh, "Hand Keypoint Detection in Single Images using Multiview Bootstrapping," 2017.
 - S.Wei, V. Ramakrishna, T. Kanade and Y. Sheikh, "Convolutional pose machines," 2016
- ASL-LEX
 - N. Caselli, Z. Sevcikova Selyr, A. Cohen-Goldberg and K. Emmorey, "ASL-LEX: A lexical database for American Sign Language. Behavior Research Methods," 2016.