# The Speech Processing Lab NN Toolkit (SPLNNTK)

Sivanand Achanta

June 14, 2017

## Contents

# 1   Introduction

The goal of building this toolkit is to make a wide variety of neural network (NN) architectures accessible to those who are more familiar with MATLAB/OCTAVE than other programming languages like python. This document describes various features of the toolkit and also how to train different models.

The following architectures are present in the toolkit.

- DNN

- RNN

- GRU

- LSTM

- Highway Layer

- CBHG [1]

- Tacotron (seq2seq) [2]

# 2   Directory Structure

There are three main directories in the toolkit.

- general_neuralnet_modules

- models

- tutorial

## 2.1   general_neuralnet_modules:

This directory has functions that are useful as building blocks to all neural architectures. There are several sub-directories in this folder a brief description of each of them is given below.

activation_functions

attention_functions

configninit_fns

fp_bp_fns

gc_fns

generate_data

get_oplayer_error

get_XY

gradient_operations

layer_index_functions

loss_functions

make_batches

normalization_functions

optim_methods

read_data

update_params_fns

weight_operations

wtinit_fns

## 2.2  models:

This directory is mainly for developers. Back-propagation check using numerical gradients is carried out for each of the neural architectures before moving them to tutorial directory.

## 2.3  tutorial:

This is the working directory useful for training and testing NNs.

- generic : contains general call functions for training neural net (mostly these are architecture independent)

- basic : contains DNN, RNN, DRNN (2layers), LSTM, LSTM-2l, GRU, GRU-2l, BLSTM, HL

- adv : contains CBHG, Tacotron (seq2seq)

# 3  Example Run

# References

[1] J. Lee, K. Cho, and T. Hofmann, "Fully character-level neural machine translation without explicit segmentation," *CoRR*, vol. abs/1610.03017, 2016. [Online]. Available: http://arxiv.org/abs/1610.03017

[2] Y. Wang, R. J. Skerry-Ryan, D. Stanton, Y. Wu, R. J. Weiss, N. Jaitly, Z. Yang, Y. Xiao, Z. Chen, S. Bengio, Q. V. Le, Y. Agiomyrgiannakis, R. Clark, and R. A. Saurous, "Tacotron: A fully end-to-end text-to-speech synthesis model," *CoRR*, vol. abs/1703.10135, 2017. [Online]. Available: http://arxiv.org/abs/1703.10135