

Crow

- [Deep Speech import](#)
- [Crow res](#)

Deep Speech import

```
Loading TSV file: /home/kenneth/Projects/JSALT_NPLM_data/Speech/Deep_Speech/cro/train.tsv
Saving new DeepSpeech-formatted CSV file to:
/home/kenneth/Projects/JSALT_NPLM_data/Speech/Deep_Speech/cro/clips/train.csvImporting mp3
files...Progress
|#####
100% completedWriting CSV file for DeepSpeech.py as:
/home/kenneth/Projects/JSALT_NPLM_data/Speech/Deep_Speech/cro/clips/train.csvProgress
|#####
100% completed
Imported 50975 samples.
Skipped 278 samples that failed on transcript validation.Skipped 32 samples that were too
short to match the transcript.
Skipped 254 samples that were longer than 10 seconds.Final amount of imported audio: 52:28:24.
Loading TSV file: /home/kenneth/Projects/JSALT_NPLM_data/Speech/Deep_Speech/cro/test.tsv
Saving new DeepSpeech-formatted CSV file to:
/home/kenneth/Projects/JSALT_NPLM_data/Speech/Deep_Speech/cro/clips/test.csvImporting mp3
files...Progress
|#####
| 99% completedWriting CSV file for DeepSpeech.py as:
/home/kenneth/Projects/JSALT_NPLM_data/Speech/Deep_Speech/cro/clips/test.csvProgress
|#####
100% completed
Imported 6374 samples.
Skipped 23 samples that failed on transcript validation.Skipped 2 samples that were too short
to match the transcript.
Skipped 31 samples that were longer than 10 seconds.Final amount of imported audio: 6:34:09.
Loading TSV file: /home/kenneth/Projects/JSALT_NPLM_data/Speech/Deep_Speech/cro/dev.tsv
Saving new DeepSpeech-formatted CSV file to:
/home/kenneth/Projects/JSALT_NPLM_data/Speech/Deep_Speech/cro/clips/dev.csvImporting mp3
files...Progress
|#####
| 99% completedWriting CSV file for DeepSpeech.py as:
/home/kenneth/Projects/JSALT_NPLM_data/Speech/Deep_Speech/cro/clips/dev.csvProgress
|#####
```

100% completed

Imported 6371 samples.

Skipped 36 samples that failed on transcript validation. Skipped 2 samples that were too short to match the transcript.

Skipped 34 samples that were longer than 10 seconds.

Final amount of imported audio: 6:34:15.

Unique characters to add to alphabet

```
['b', 'm', 'f', ' ', 'e', 'r', 'ú', 'l', 'd', '\uf009', 'k', 'c', 'g', 'i', 'a', 'h', 'p',  
'u', 'o', 'w', 'á', 't', 's', 'n', 'x']
```

the kenlm model needs to be built. Use the data fran sent and check this out

<https://github.com/mozilla/DeepSpeech/issues/1411>

Crow res

After setting alpha and beta parameters for language model to 0 (effectively negating the language model)

```
--lm_weight .1 \  
--lm_alpha 0\  
--drop_source_layers 1 \      --source_model_checkpoint_dir  
"${SOURCE_MODEL}/deepspeech-0.5.1-checkpoint/" \  
--n_hidden 2048 \  
--epoch -10 \  
--earlystop_nsteps 5 \  
--train_batch_size 30\  
--dev_batch_size 48 \  
--test_batch_size 48 \  
--learning_rate 0.001 \  
--dropout_rate 0.2 \  

```

Test on /home/kenneth/Projects/JSALT_NPLM_data/Speech/Deep_Speech/cro/clips/test.csv - WER:
0.990699, CER: 0.871154, loss: 27.086073

Examples:

```
-----WER:  
2.000000, CER: 0.800000, loss: 10.316745  
- src: "grgacgicr " - res: "r r"  
-----WER:  
1.000000, CER: 0.666667, loss: 1.255609  
- src: "grg" - res: "r"  
-----WER:  
1.000000, CER: 0.400000, loss: 1.491453  
- src: "bggrp" - res: "bgg"  
-----WER:  
1.000000, CER: 1.000000, loss: 1.747670  
- src: "cg" - res: ""
```

```

-----WER:
1.000000, CER: 1.000000, loss: 2.060837
- src: "gahg" - res: ""
-----WER:
1.000000, CER: 1.000000, loss: 2.078961
- src: "cg" - res: ""
-----WER:
1.000000, CER: 0.500000, loss: 2.292280
- src: "gg" - res: "bgg"
-----WER:
1.000000, CER: 1.000000, loss: 2.531346
- src: "cg" - res: "r"
-----WER:
1.000000, CER: 1.000000, loss: 2.713193
- src: "ccg" - res: ""
-----WER:
1.000000, CER: 1.000000, loss: 2.719913
- src: "gg" - res: ""
-----

```

```

--lm_weight .1 \
--lm_alpha 0\
--drop_source_layers 1 \      --source_model_checkpoint_dir
"${SOURCE_MODEL}/deepspeech-0.5.1-checkpoint/" \
--n_hidden 2048 \
--epoch -10 \
--earlystop_nsteps 5 \
--train_batch_size 30\
--dev_batch_size 48 \
--test_batch_size 48 \
--learning_rate 0.0001 \
--dropout_rate 0.2 \

```

Test on /home/kenneth/Projects/JSALT_NPLM_data/Speech/Deep_Speech/cro/clips/test.csv - WER:
0.989741, CER: 0.857567, loss: 25.778952

```

-----WER:
2.000000, CER: 0.750000, loss: 7.365357
- src: "grgágr " - res: "r r"

```

-----WER:
1.000000, CER: 0.666667, loss: 0.978152
- src: "grg" - res: "r "
-----WER:
1.000000, CER: 0.666667, loss: 1.270408
- src: "grg" - res: "r "
-----WER:
1.000000, CER: 1.000000, loss: 1.365655
- src: "gsa " - res: "b"
-----WER:
1.000000, CER: 0.750000, loss: 1.641714
- src: "gahg" - res: "bg"
-----WER:
1.000000, CER: 0.400000, loss: 1.725150
- src: "bgga " - res: "bgg"
-----WER:
1.000000, CER: 1.000000, loss: 1.827598
- src: "ggl " - res: "b"
-----WER:
1.000000, CER: 1.000000, loss: 1.866848
- src: "gacw " - res: "b"
-----WER:
1.000000, CER: 0.714286, loss: 1.869485
- src: "gáicdi " - res: "di"
-----WER:
1.000000, CER: 0.400000, loss: 1.982887
- src: "bggrp" - res: "bgg"
