MODELLING FONTS WITH CONVOLUTIONAL NEURAL NETWORKS

Progress presentation

Philip Bouman

29 juni 2017

Universiteit van Amsterdam

RESEARCH QUESTION

- Institute for Logic, Language and Computation (ILLC)
- Bouba-Kiki effect (iconicity)
- Grapheme to phoneme classification
- Convolutional Neural Network

DATA

- 340 scripts collected from Omniglot.com
- Alphabets, Abigudas, Abjads
- 280 scripts in dataset
- Letter-IPA pairs

Ҳә ҳә	[ħʷ]
Ҿ ҿ	[t͡sʼ]
ь	[ʲ]
Б б	[b]
Е е	[ɛ]
К к	[kʼ]
Н н	[n]
Ҭ ҭ	[tʰ]

Кіь	[1, 0, 0, 0, 0, 0, 0, 0, 2, 1, 1, 0, 1, 1]
Р	[1, 1, 1, 2, 1, 1, 0, 0, 0, 0, 2, 2, 2, 2]
Х	[1, 0, 1, 1, 0, 0, 0, 0, 2, 1, 1, 0, 0, 0]
Щ	[1, 0, 1, 1, 0, 0, 0, 0, 1, 1, 1, 0, 1, 0]
Б	[1, 0, 0, 0, 0, 0, 0, 1, 2, 0, 2, 2, 2, 2]
Ё	[0, 1, 1, 2, 1, 0, 0, 0, 2, 1, 1, 0, 1, 0]
Кв	[1, 0, 0, 0, 0, 0, 0, 1, 2, 1, 1, 0, 0, 0]
Л	[1, 1, 1, 2, 1, 0, 1, 0, 0, 0, 2, 2, 2, 2]
С	[1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 2, 2, 2, 2]
Хв	[1, 0, 1, 1, 0, 0, 0, 1, 2, 1, 1, 0, 0, 0]

DATA

- Images (32x64) with Google NOTO fonts
- Phonological encoding scheme
- Training (80 %), validation (10 %), test (10 %)
- Two Datasets (language and random character)

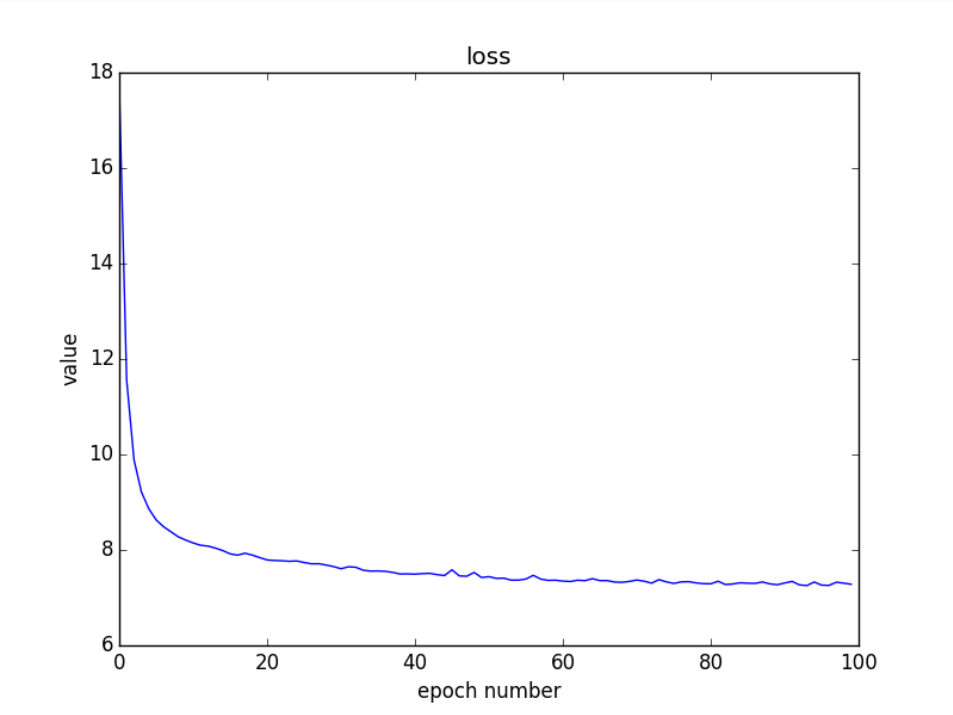
IPA	consonantal	sonorant	continuant	delayed release	approximant	nasal	labial	rounded	strident	height	low	front	back	tense
a	-	+	+	0	+	-	-	-	0	-	+	-	-	0
æ	-	+	+	0	+	-	-	-	0	-	+	+	-	0
o	-	+	+	0	+	-	+	+	0	-	-	-	+	+
œ	-	+	+	0	+	-	+	+	0	-	-	+	-	-
e	-	+	+	0	+	-	-	-	0	-	-	+	-	+
ø	-	+	+	0	+	-	+	+	0	-	-	+	-	+
u	-	+	+	0	+	-	+	+	0	+	-	-	+	+

METHOD

- Convolutional neural network
  - State of the art
  - Handwritten character recognition (MNIST)
  - New for this task
- Model with Keras (Theano)

EXPERIMENTS

- Two different models/architectures
  - 2 Convolutional layers, fixed kernel size (MNIST)
  - 3 Convolutional layers, decreasing kernel size (Image classification)
- Both datasets
- 100 epochs, 10 runs



RESULTS

- Accuracy on test set (language dataset)

feature	Model 1	Model 2
approximant	0.593352192442	0.649646393295
back	0.456718529059	0.582036775191
consonantal	0.589533239114	0.681471004328
continuant	0.612446959024	0.639179632291
delayed release	0.482319660622	0.52022630843
front	0.483168316904	0.591937765285
height	0.483168316904	0.596181046756
labial	0.966053748232	0.965912305516
low	0.493352192438	0.583451202348
nasal	0.781188118812	0.973833097595
round	0.579490806279	0.722206506407
sonorant	0.437906647884	0.543140028369
strident	0.586421499293	0.661386138614
tense	0.505940594135	0.585572843079

CONCLUSION

- Model 2 yields higher accuracy overall
- Continuant best predictive feature for model 1
- Round best predictive feature for model 2
- Round, continuant, consonantal, strident, approximant best predictive features overall

DISCUSSION/FUTURE WORK

- Discussion
  - More depth improves accuracy
  - Continuant and round best features for this task
- Future work
  - Investigate intermediate feature representations
  - Different features/ feature combinations
  - Experiment with deeper architectures