Brian Landy
Deep Learning
3/27/2020

Milestone 2

Title:

Examination of Translation Performance with Various Input Feature Types on a Continuous Sign Language Translation Model

One Liner:

The goal of this set of experiments is to use a Continuous Sign Language Translation model with different input features from AlexNet, OpenPose, Resnet as well as Resnet3D to try and improve upon or reproduce already stated results in terms of ROGUE and BLEU scores. The model goes by the name Neural SLT and the dataset used RWTH-PHOENIX-Weather 2014T, which is a continuous sign language dataset.

Sources and One-liners

Paper for NSLT:

http://openaccess.thecvf.com/content_cvpr_2018/papers/Camgoz_Neural_Sign_Language_CVPR_2018_paper.pdf

Camgoz et al. [1] proposes a model that aims to perform the task of continuous sign language translation from images to text which, unlike much prior work, has the goal of translating everything including order and grammar in a sentence.

Background papers:

Motivation, current methods and history:

https://watermark.silverchair.com/enj003.pdf?token=AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAlgwggJUBgkqhkiG9w0BBwagggJFMIICQQIBADCCAjoGCSqGSIb3DQEHATAeBglghkgBZQMEAS4wEQQMzlxzzaJ4d_UF_msqAgEQgIICC6_d24zVhJ4VM7Ysi0qFPyJHOUj-ovRA7uV0lPN5JCCKZMszGKlAlF0CcX0o6K5ZvBPoEishZRjuNH7C-Ld_KI0vRL8c4eCuRm78wiMr8-4SloBb0dOJZWKcPf5SBLr_kiuanyQQQC0AIi2LdCfq7NKwWSWoNRKG7KDFZ6EK-4Wk-VNUccIKCaov0m7UoUYVPlR7tU5opn0nrunv_311u0Rv0jYtKJH--b3pR75hWqyMIV09Gf-vdWOC6bLlSVxZJAkmbJkwbXK2vMeWFfg9WVhTmjUOBNiDOdj-4HVJnk_0BKZ3ZVqlJgTb90w7SwKJs2u-tXe7By9xHMyPslCUAKgv62F3Qsyh1NvDo4k376iSAe9R9aEXsJQt_qV1gRQKFHm6CEbhMe4QYwV50xADSy7V-USbvitf_POUUqC3Oflb6YGKi3ygpuacDMFUf0QzglwQ_1UYdyBrLfOaEOqoaQSbzDI4Fca6MgjYti7NASJTgMumUo9vSGU1zC8hurKoWnt_kuNsIV4AR9HyHVryQ3FbPmO-Xa4DQ00SV03CW6ecPwouaP8L8kyvqQKpZbah_L5TFsTe494KUZ03iOfeWhbI-ZnhIJ_r5EOV_NExk2tBMqKDwXQpeIIX8pTlpw4-NZ88sHtUDkGxANGiTgoA868TMGq33RbuE6p4zSxaGmXrUbeLFCGBHC0OBLo

Parton et al. [2] covers a broad variety of sign language translation work and research from original notation in the 1960s to more modern NLP approaches as well as applications of sign language translation(SLT), which will help show the need for SLT tools.

https://arxiv.org/pdf/1908.08597.pdf

Bragg et al. [3] discusses the results of a large sign language conference from 2019 that was focused on sign language translation technologies, applications of such technologies, deaf experiences, available sign language research resources, as well as current challenges in translation.

Model building block papers:

Understanding RNNS:
Bidirectional RNN and LSTM:
https://arxiv.org/pdf/1506.00019.pdf

Lipton et al. [4] dives deeply, and very informatively, into the category of Recurrent Neural Networks (RNNs) and discusses 2 important concepts from the NSLT model. These are Bi-directional RNNs (BRNNs) and Long Short Term Memory (LSTM) cells, which are both found in the NSLT model experiments.

Another paper on GRU/LSTM: for discussing the use of gru vs lstm, build understanding with this
https://arxiv.org/pdf/1412.3555.pdf

Chung et al. [5] discuss the differences between LSTM cells and Gated Recurrent Unit Cells (GRU) in tasks where recurrent models are used; this paper will serve as a good explanation about GRU. They show that LSTM and GRU perform better than traditional methods in tasks in the area of speech modeling.

Understanding sequence to sequence:
Rnns video to text
https://arxiv.org/pdf/1505.00487.pdf

Venugopalan et al. [6] propose a method to provide text captions for videos by using a mix of Convolutional Neural Networks and LSTMs. This is a sequence to sequence framework similar to what is deployed in NSLT.

Understanding attention:
Align and translate paper
https://arxiv.org/pdf/1409.0473.pdf

Bahdanau et al. [7] implemented an attention method that allows for variable length sequences to be selected automatically and translated. This is one type of attention used in NSLT.

https://arxiv.org/pdf/1508.04025.pdf

Luong et al. [8] has sought to improve translation by incorporating attention scoring to neural machine translation; this allows for less important information in a sentence to be excluded while more important information is emphasized. NSLT experiments with this type of attention scoring.

Similar papers:

KETI paper
https://arxiv.org/pdf/1811.11436.pdf

There are many similar experiments with variations that are always appearing in the domain of neural machine translation. In this case, Ko et al. [9] use a similar model to NSLT with the modification being that OpenPose keypoints are fed into the model instead and the internal units are GRU cells.

Method of sourcing information:

The papers in the following section are a mix of building block papers to gain knowledge on the modules that make up the NSLT model, the actual NSLT paper itself as well as similar experiments, and papers that discuss motivation and progress for continuous sign language translation.

Other useful links for understanding concepts and finding papers/info not allocated yet:
https://blog.floydhub.com/attention-mechanism/

https://www.quora.com/What-is-the-best-research-paper-about-recurrent-neural-networks-to-start-with

https://arxiv.org/pdf/1506.00019.pdf

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.55.5709&rep=rep1&type=pdf

https://machinelearningmastery.com/how-does-attention-work-in-encoder-decoder-recurrent-neural-networks/

https://blog.floydhub.com/attention-mechanism/

https://github.com/spro/practical-pytorch/blob/master/seq2seq-translation/seq2seq-translation.ipynb

https://www.quora.com/What-is-the-best-research-paper-about-recurrent-neural-networks-to-start-with

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.55.5709&rep=rep1&type=pdf

https://machinelearningmastery.com/how-does-attention-work-in-encoder-decoder-recurrent-neural-networks/

https://towardsdatascience.com/understanding-encoder-decoder-sequence-to-sequence-model-679e04af4346

https://papers.nips.cc/paper/5346-sequence-to-sequence-learning-with-neural-networks.pdf

https://arxiv.org/pdf/1908.08597.pdf

https://pdfs.semanticscholar.org/70eb/31c05b4eddaa81acee12343667dadeb71123.pdf

https://arxiv.org/pdf/1409.0473.pdf

https://ai.stackexchange.com/questions/8190/where-can-i-find-the-original-paper-that-introduced-rnns

https://www.bioinf.jku.at/publications/older/2604.pdf

https://arxiv.org/pdf/1508.04025.pdf

https://blog.floydhub.com/attention-mechanism/

https://github.com/spro/practical-pytorch/blob/master/seq2seq-translation/seq2seq-translation.ipynb

http://colah.github.io/posts/2015-08-Understanding-LSTMs/

https://www.quora.com/How-is-the-hidden-state-h-different-from-the-memory-c-in-an-LSTM-cell

https://arxiv.org/pdf/1412.3555.pdf

https://towardsdatascience.com/day-1-2-attention-seq2seq-models-65df3f49e263

https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21

https://towardsdatascience.com/illustrated-guide-to-recurrent-neural-networks-79e5eb8049c9

https://towardsdatascience.com/understanding-bidirectional-rnn-in-pytorch-5bd25a5dd66 https://deeplearning4j.konduit.ai/language-processing/tokenization

https://stackoverflow.com/questions/44238154/what-is-the-difference-between-luong-attention-and-bahdanau-attention

https://towardsdatascience.com/day-1-2-attention-seq2seq-models-65df3f49e263

https://arxiv.org/pdf/2002.00479.pdf

-->enj003.pdf