

Short Term Scientific Mission

Multi-task Learning in Neural Machine Translation

Iacer Calixto.

July, 2017.

COST Action: IC1307

STSM Reference Code: COST-STSM-IC1307-37836

Beneficiary: Iacer Calixto, ADAPT Centre, Dublin City University, Dublin, Ireland.

Host: Prof. Dr. Lucia Specia, Department of Computer Science, The University of Sheffield, Sheffield, UK.

Period: 1st/July/2017 until 16/July/2017.

Introduction

Machine Translation has recently been addressed from a multi-modal perspective, where visual features are incorporated in an Neural Machine Translation (NMT) model, which is trained end-to-end [1, 2]. One of the ways researchers have incorporated visual features into NMT is through multi-task learning, where a network is trained to translate sentences from a source language into a target language (MT) and also to rank sentences given images and vice-versa (image-sentence ranking) [2]. By training a model towards two tasks, not only could overfitting issues be avoided, but also the lack of grounding in a MT model can be addressed (i.e. incorporating visual features).

Aim of the STSM

The aim of the STSM was two-fold: (i) to present some of my previous work to the host research group; (ii) to work collaboratively on a multi-task approach integrating natural language processing (NLP) models using multi-modal corpora.

Description of the work carried out

I gave a talk on the paper “Doubly-Attentive Decoder for Multi-Modal Neural Machine Translation”, which will be presented at ACL 2017. The talk was part of the University of Sheffield's NLP group invited talks and was attended by many researchers from their Computer Science (CS) department. I met with many researchers in the group, mainly to discuss about shared research interests and to organise the work to be done in the 16 days I spent in the NLP group. We have started working collaboratively on a multi-task approach integrating machine translation and auto-encoders, which involved me and two researchers

from the CS department, Dr. Pranava Madhyastha and Dr. Josiah Wang. They already had an initial implementation of a basic sequence to sequence framework written in PyTorch, and I have worked on their code base. We shared code in a private repository for this project, which can be open-sourced as we publish our results, although this is a topic open for discussion at that stage.

I have extended the PyTorch code base of the University of Sheffield NLP group and extensively documented and tested all the changes made. These changes include allowing models to be trained using multi-task learning, where an NMT model is trained jointly with (i) an neural auto-encoder over the source language, (ii) an neural auto-encoder over the target language, (iii) or both. One important novelty in the work we have conducted is that the neural auto-encoder is not necessarily used to recover the same sentence used as input, but also to “recover” semantically similar sentences, e.g. sentences that describe one same image. By doing that, we can incorporate comparable corpora into Machine Translation.

Description about how the results contribute to the action's aims

During my visit to the CS department in The University of Sheffield, we have discussed and experimented with interesting ways to train neural machine translation models using multi-task learning. A novelty in our work is that we incorporate semantic information from comparable corpora directly into a machine translation model, where these comparable corpora describe one same image. These are preliminary experiments with the goal of expanding the model into using image features as well, which relates directly to the Action’s Working Group (WG1) Integrated Modelling of Vision and Language for CV and NLP Tasks and (WG2) Applications of Integrated Models. In addition to that, I gave the talk “Doubly-Attentive Decoder for Multi-Modal Neural Machine Translation”, where I discuss a paper I have co-authored and that was recently accepted for publication at ACL 2017. In this paper, I introduce a model that effectively integrates (multi-modal) image and (multilingual) text. We are currently working on experiments in a multi-task setting involving machine translation and auto-encoders. Finally, we also wish to integrate image description generation to our multi-task setting in the future, therefore also addressing the Action’s Working Group (WG3) Automatic Generation of Image & Video Descriptions.

Confirmation by the host institution of the successful execution of the STSM

A letter from the host, Lucia Specia, will be sent together with this report confirming the success of the visit.

Acknowledgements

I would like to thank the NLP group in the University of Sheffield CS Department for hosting me and providing for all I needed. I would also like to thank prof. Lucia Specia, Dr.

Pranava Madhyastha, and Dr. Josiah Wang for the discussions and collaboration.

References

[1] Calixto, Iacer and Liu, Qun and Campbell, Nick (2017). Doubly-Attentive Decoder for Multi-modal Neural Machine Translation. Proceedings of the 55th Conference of the Association for Computational Linguistics: Volume 1, Long Papers. Vancouver, Canada (Paper Accepted).

[2] Elliott, Desmond and Kádár, Ákos (2017). Imagination improves Multimodal Translation. CoRR, abs/1705.04350. <https://arxiv.org/abs/1705.04350>