

June 22, 2021

MASTER'S THESIS

for

Sheng Jiang

Student ID 03724842, Degree EI

Generative Answering System based on Image Captioning

Problem description:

In this work, the main objective is to design a smart reply system, which can answer to human language as well as image inputs. While the traditional chatbot focuses on responding to natural language inputs, the proposed system would also aim to respond to given image inputs as well. For example, when a human sends a picture of dog to the system, the proposed system would respond as 'That's a very cute dog!'. In order to accomplish this goal, the proposed system would consist of image captioning network as well as Generative Question Answering (GQA) network. When the language is given as an input, the input would be sent to the GQA network, such that the answer can be generated. When the image is given as an input, the language caption describing the image will be generated first, and the caption will be sent to the GQA in order to generate the right answers. Therefore, the objective of this work can be divided into two folds: (1) image captioning system (2) generative question answering system. To be more specific, a transformer-based architecture will be used for image captioning system [2], and GPT-2 [1] will be used for generative question answering system.

Tasks:

- Literature review on image captioning and generative question answering (GQA) networks.
- Apply transformer-based model for image captioning task, GPT-2 for generative question answering.
- Implement and evaluate the online demo of the proposed smart reply system.

Bibliography:

- [1] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9, 2019.
- [2] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *arXiv preprint arXiv:1706.03762*, 2017.

Supervisor: Dr. Hyemin Ahn
Start: 24.05.2021
Intermediate Report: xx.xx.2021
Delivery: xx.11.2021

(D. Lee)
Univ.-Professor