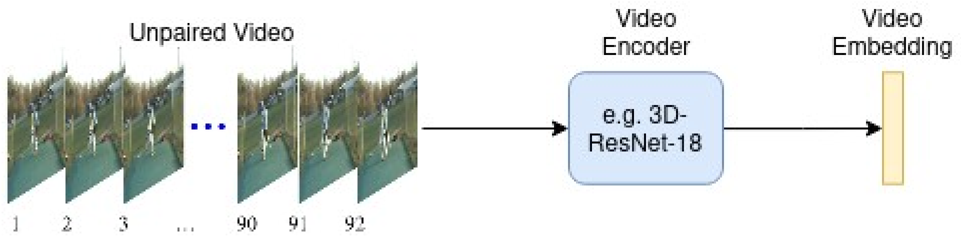
Master Thesis: Enhanced Video Question-Answering on Unpaired Video Data

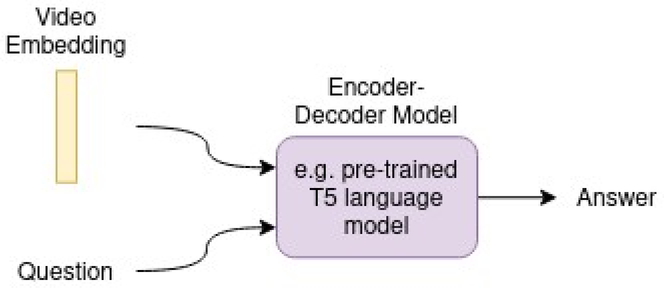
Thesis Idea:

Existing Video Question-Answering systems rely on large amounts of video-text pairs, which are typically scraped from the internet. However, in fields where data can not be scraped from the internet, such as in medical applications, collecting such datasets is expensive and time-consuming. Instead of relying on huge amounts of video-text pairs, the goal of this Master Thesis is to create an Artificial Intelligence (AI) System for Video Question-Answering, that can leverage large amounts of unpaired video data through unsupervised learning. This can be achieved by the following steps:

1) Use Unsupervised Temporal Contrastive Learning on unpaired video data
(model and training code is already available)



2) Train/fine-tune a language-based knowledge interface using a subset of video-text pairs



What we offer:

- GPU infrastructure for model training and evaluation
- Exchange with experts in relevant fields
- Working at the forefront of AI research
- Option to write a scientific research paper about the Master Thesis

What we are looking for:

- **strong background in machine learning (ML)**
- **experience with Python and PyTorch**
- motivation to conduct research at the forefront of AI
- Preferred: experience in ML-based processing of video and/or text data

Contact:

Anthony Richardson
Email: antmen@uni-bremen.de