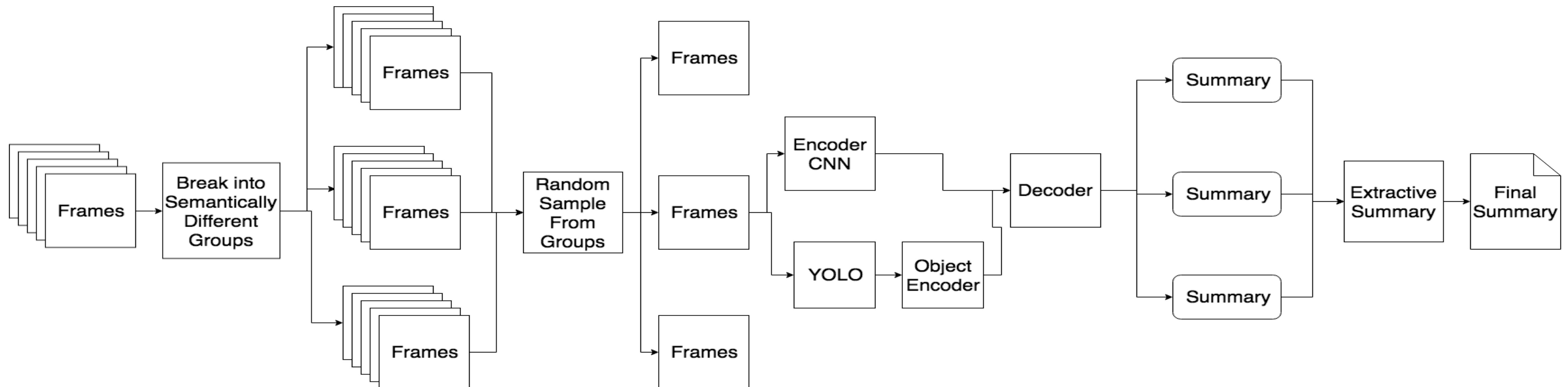


“What’s that Video Again?”: Video Search Engine

Abby Gray, Akshat Shrivastava, Kevin Bi, Sarah Yu
CSE 455: Computer Vision | Spring 2018 | University of Washington



Abstract

Our project offers video search through video summarization. Our system breaks videos down into semantically different groups of frames, recognizes objects in those frames to create image captions for each frame, and ‘extractively summarizes’ the aggregated frame captions for a final video summary. These video summaries are housed in a database searchable through our simple search engine.

System Design

Video Segmentation

- Distributor extracts all frames from video using **OpenCV**
 - Distributor groups every 450 consecutive frames, sends to each Worker
 - Workers extract semantically similar frames using structural similarity index (SSIM)

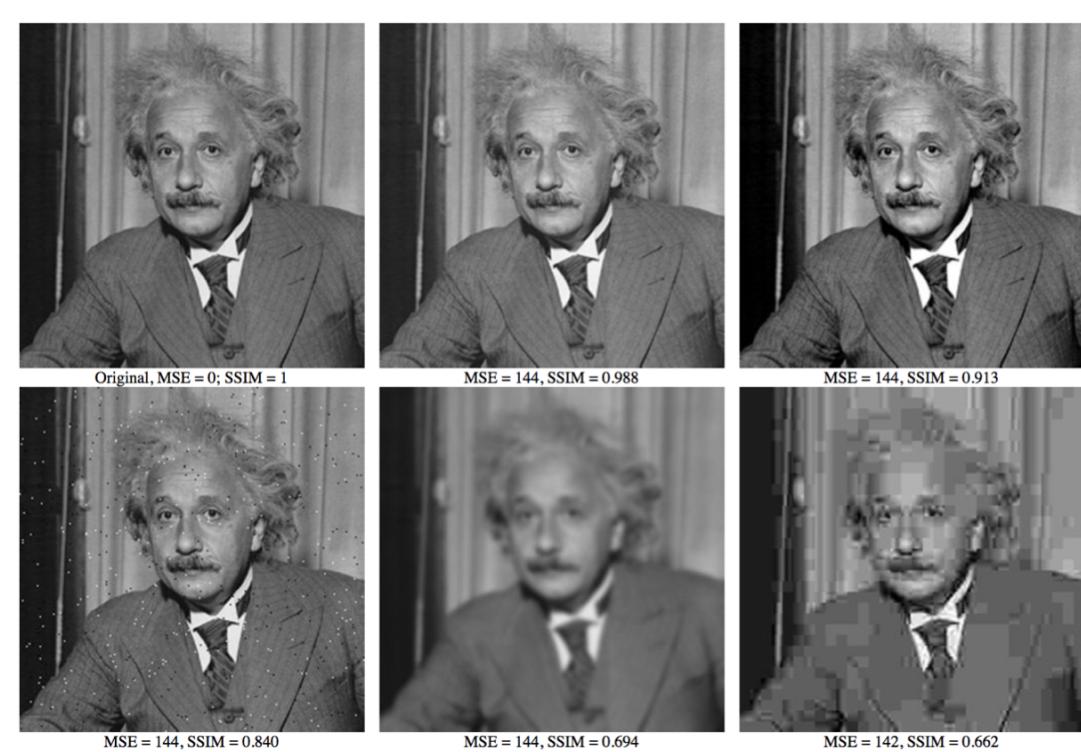
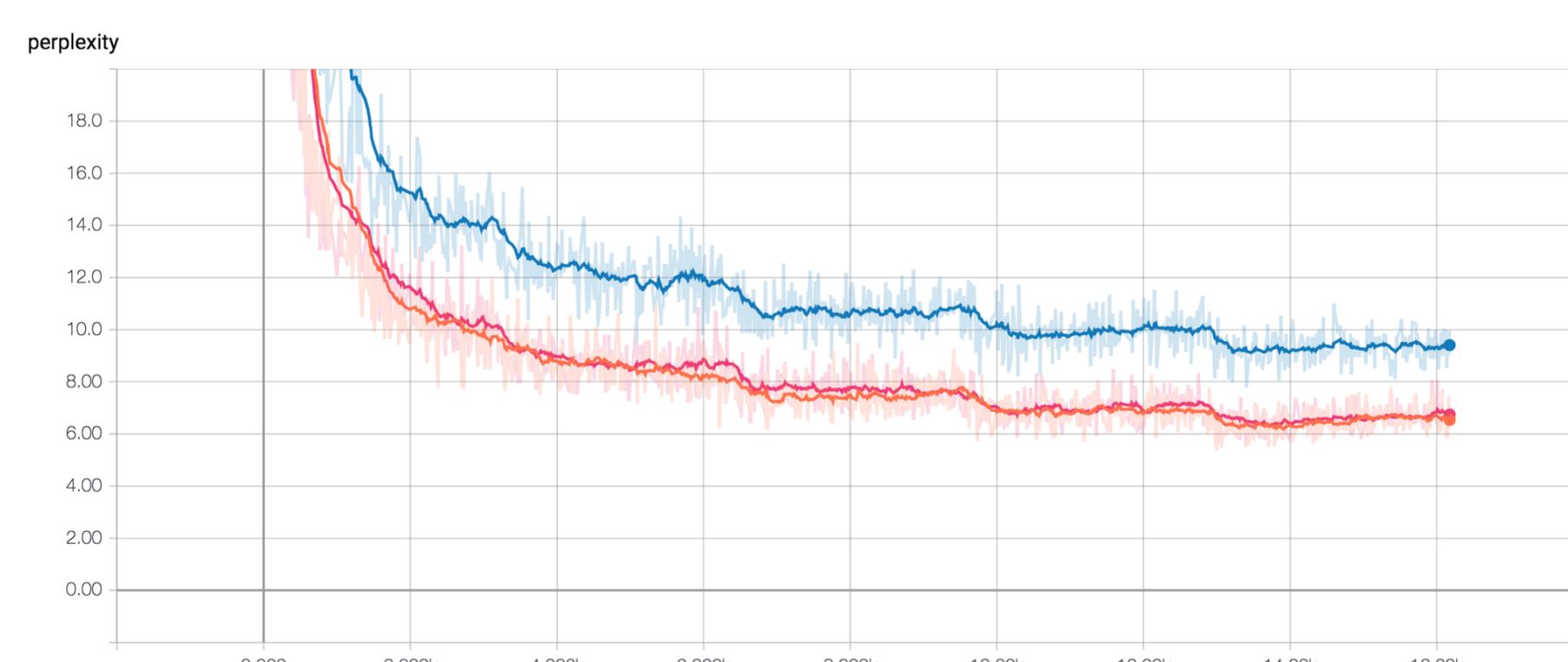


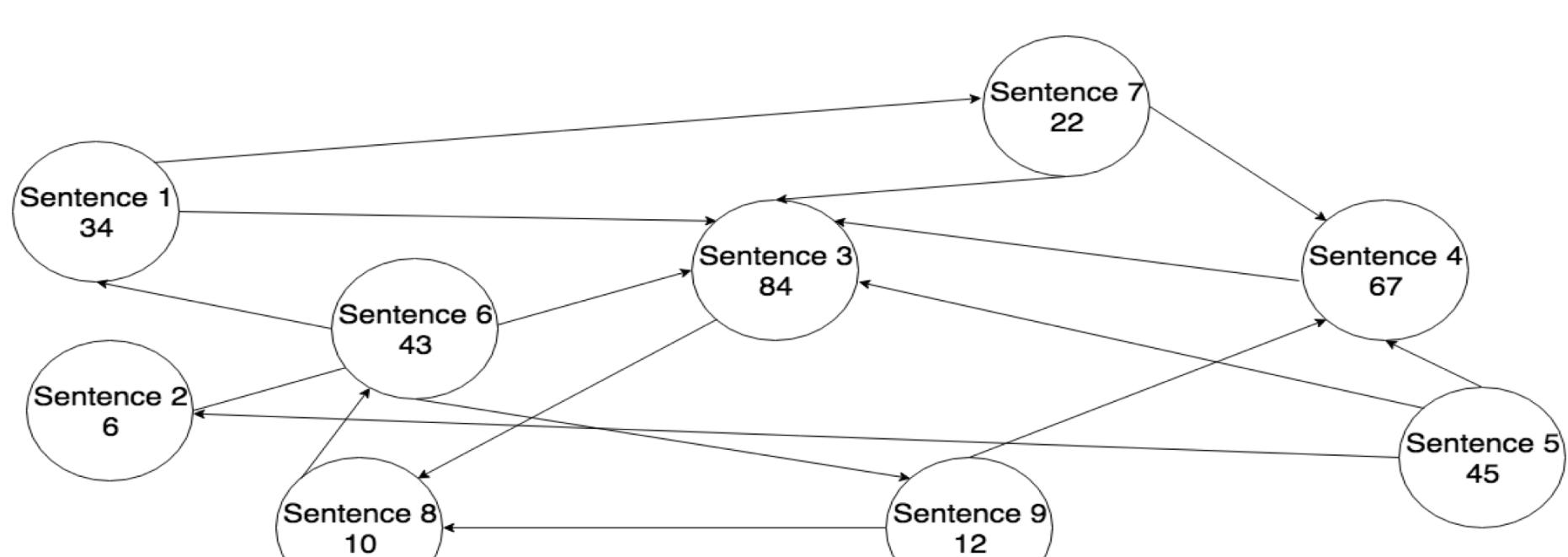
Image Captioning

- 1st encoder: encodes image using ResNet features
 - 2nd encoder: encodes **YOLO** extracted object bounding boxes & object names
 - Decoder: combines features from encoders → generates image caption
 - Addressing Noise: We use **Tiny YOLO** for training and **YOLO** for testing

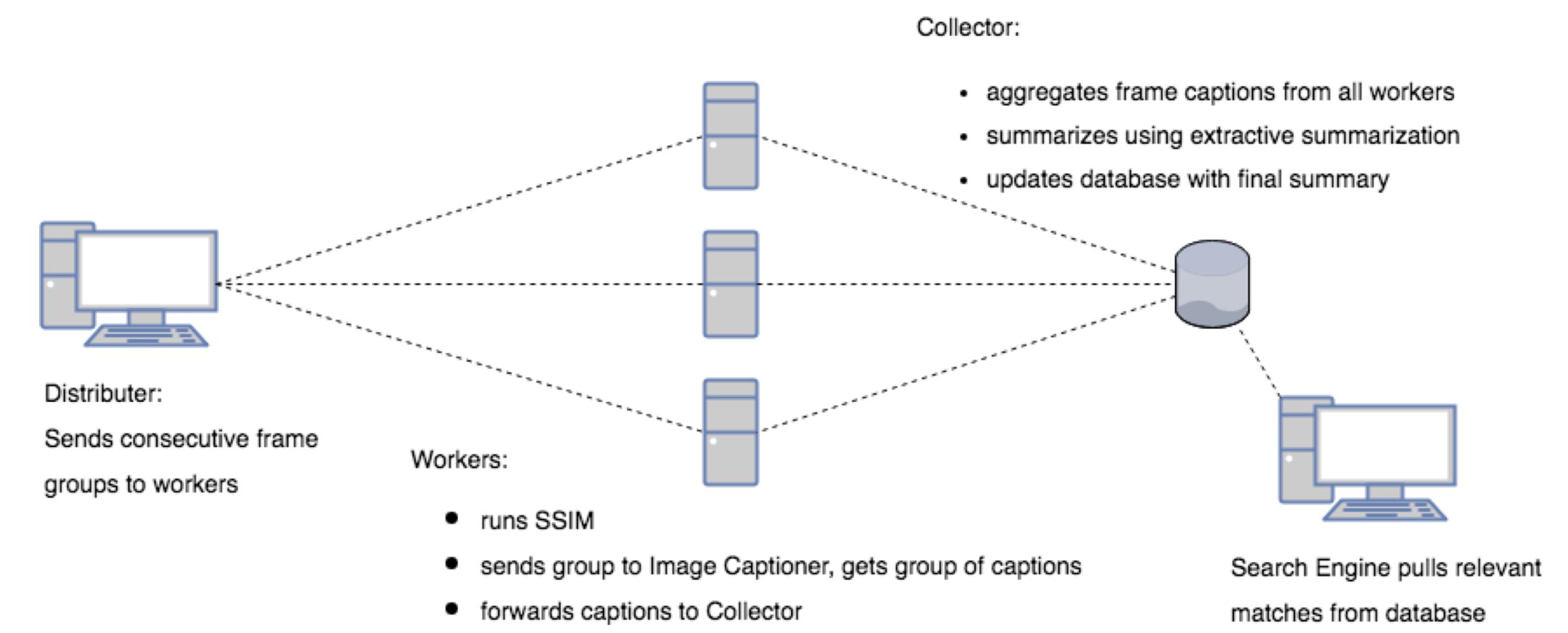


Extractive Summarization

- **TextRank** using **Cosine Similarity** – unsupervised method for summary
 - List of image captions per frame ➡ Top 5 ranked sentences

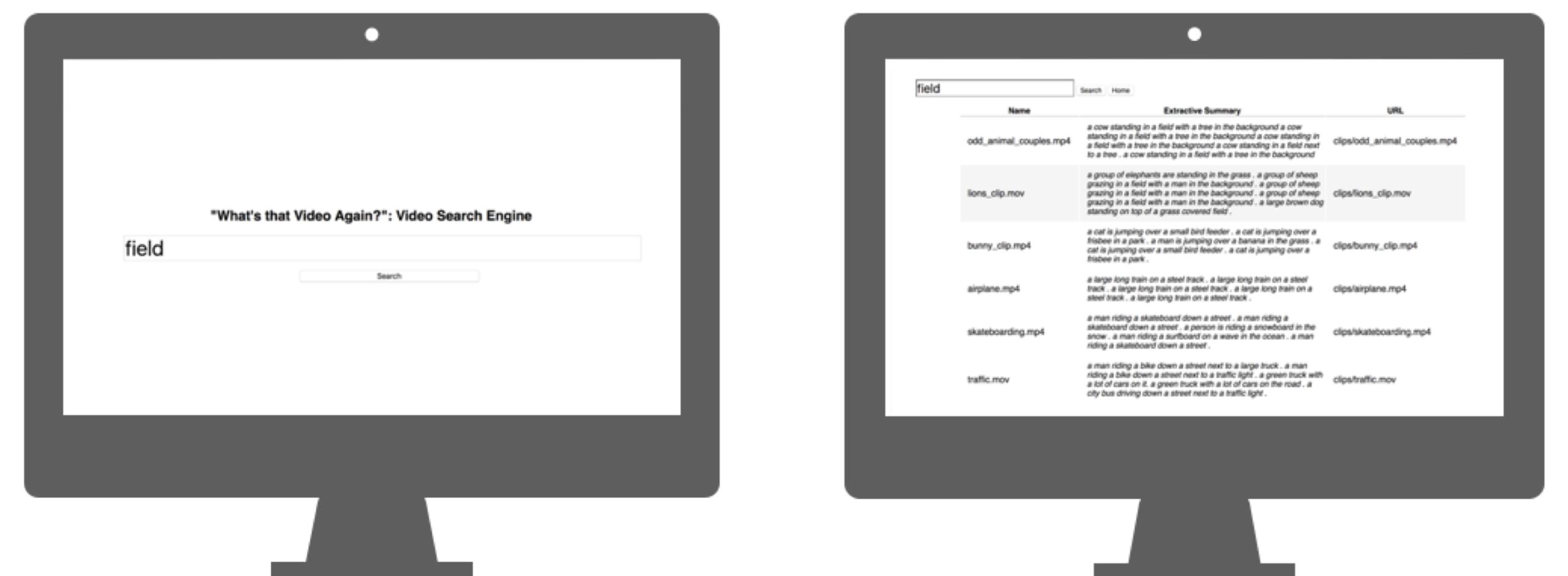


System Distribution



Search & Rank

- **Flask** web application to connect the database to a simple search engine
 - Compare search query - generated video summaries using TF-IDF (**scikit learn**)
 - Shows relevant results ordered by TF-IDF cosine similarity score rankings



Results



Summary:

- a man riding a bike down a street next to a large truck
 - a man riding a bike down a street next to a traffic light
 - a green truck with a lot of cars on it
 - a green truck with a lot of cars on the road
 - a city bus driving down a street next to a traffic light

References

- <http://www.cns.nyu.edu/pub/eero/wang03-reprint.pdf>
 - <https://arxiv.org/pdf/1707.07102.pdf>
 - <https://pjreddie.com/media/files/papers/YOLO9000.pdf>