First lightning talk session (10:25AM)

Alexander Dietmüller (ETH Zurich - Remote)

Title: Chasing the Tail in Video Streaming
Abstract:
Because of complex traffic dynamics, machine learning is becoming more and more important for video streaming, and with that comes the challenge of keeping the training data and model up-to-date. We present Memento, a method to update a training set by estimating the sample space density, which allows to: (i) focus on rare patterns important for tail performance; and (ii) detect whether the training set has changed enough to warrant retraining.

Oscar Moll (MIT)

Title: Seesaw: interactive ad-hoc image searches over image databases
Abstract:
Many video data tasks require first constructing datasets for training and testing. Locating useful training examples on large databases of unlabeled data can be challenging due to lack of initial models. Seesaw is a system to help users bootstrap their searches with natural language using a semantic embedding (CLIP), and then improve search results within seconds through image-region based feedback.

Jenya Pergament (Stanford University)

Title: An Interactive Annotation Tool for Perceptual Video Compression
Abstract:
Perceptual quality metrics (PQMs) are at the core of lossy video compression and are the focus of active research spanning from methods driven by the human visual system, to using GANs or features of pre-trained DNNs. In PQM development, human feedback is typically incorporated as a single scalar quality score indicating preference of one distorted video over another. In reality, some videos may be better in some parts but not in others. We propose an approach to collecting finer-grained feedback in the form of human spatio-temporal importance maps in a video, which we can then leverage toward better perceptual compression. To this end, we built a novel web-tool which allows users to paint these spatio-temporal importance maps over videos. The tool allows for interactive successive refinement: we iteratively re-encode the original video according to the painted importance maps, while maintaining the same bitrate, thus allowing the user to visually see the trade-off of assigning higher importance to one spatio-temporal part of the video at the cost of others. We use this tool to collect data in-the-wild and utilize the obtained importance maps in the context of x264 coding to demonstrate that the tool can indeed be used to generate videos which, at the same bitrate, look perceptually better through a subjective study — and are 1.9 times as likely to be preferred by viewers.

Wenjia He (University of Michigan)

Title: Opaque filter query optimization for video analytics
Abstract:
The opaque filter query, a query with a selection predicate that is implemented with a user-defined function (UDF), constitutes an important workload when processing unstructured video data. Typically, the UDF includes a CNN-style trained image classifier with relatively long runtimes. Because the predicate's semantics cannot be understood by the query optimizer, many traditional optimizations are not possible. We propose a novel index mechanism, voodoo indexing, to optimize this kind of query.

Gaurav Tarlok Kakkar (Georgia Tech)

Title: EVA: A Symbolic Approach to Accelerating Exploratory Video Analytics with Materialized Views
Abstract:
Exploratory video analytics tend to have overlapping computation and often differ in their predicates. However, these predicates are computationally expensive to evaluate since they contain user-defined functions (UDFs) that wrap around deep learning models. In this talk, I present EVA, a video database management system that automatically materializes and reuses the results of expensive UDFs to facilitate faster exploratory analysis.

Second lightning talk session (2:30PM)

Vishakha Gupta (ApertureData)

Title: Can We Simplify Video Management and Access for Analytics?
Abstract:
Visual data (images/videos) is rich in valuable insights. Data science and ML techniques can help understand visual content and enable better customer experience across application domains, in turn driving its exponential growth. A key factor for success in visual ML is to get the foundational data infrastructure right, which can be extremely challenging and time consuming due to a lack of data management solutions designed with visual data or data science in mind. During my lightning talk, I will start by briefly highlighting why visual data today needs special treatment and how it can be challenging to achieve that. I will then show how ApertureDB, a purpose-built database for analytics on visual data can address some of these challenges. During the poster session, I can dive deeper into certain architecture and design decisions that have worked well so far as we build, deploy, and scale ApertureDB.

Pulkit Tandon (Stanford University)

Title: Txt2Vid: Ultra-Low Bitrate Compression of Talking-Head Videos via Text
Abstract:

We ask the question: “Can we compress audio-video content generated via webcams to just text and recover videos with similar Quality-of-Experience compared to standard codecs in a low bitrate regime?” and answer it in affirmative using state-of-the-art deep learning models.

Favyen Bastani (AllenAI)

Title: Data Science Platforms for Image and Video Analytics
Abstract:
In this talk, I will discuss the development of tools that can enable domain experts, who may not have much ML experience, to build robust image and video analytics pipelines.

Yao Lu (Microsoft Research)

Title: Serving and optimizing ML workflows on Heterogeneous Infrastructures
Abstract:
I will talk about our recent system for serving and optimizing ML workflows on heterogeneous infrastructures (e.g., IoT or hybrid cloud).

Francisco Romero (Stanford University)

Title: VIVA: An End-to-End System for Interactive Video Analytic
Abstract:
The growth of video volumes and increased DNN capabilities has led to a growing desire for video analytics. In response, the data analytics community has proposed multiple systems that optimize specific query types (e.g., selection queries) or a particular step in query execution (e.g., video retrieval from storage). However, none of these systems provide end-to-end, practical video analytics for users to iteratively and interactively engage with queries, as is the case with analytics systems for structured data. In response, we are building VIVA: an end-to-end system for interactive video analytics. In this talk, we will describe the challenges of designing VIVA, and outline ongoing and future work for realizing VIVA.