Demo: Live Video Stream Triggers

Lenin Ravindranath, Matthai Philipose, Peter Bodik, Paramvir Bahl
Microsoft Research
Redmond, WA, USA
lenin@microsoft.com, matthaip@microsoft.com, peterb@microsoft.com,
bahl@microsoft.com

Live streaming is an increasingly popular way to broadcast videos ranging from formal news channels to kitten cams to home security camera feeds. Live streaming marries the rich detail of video with the timeliness of live transmission and the ease of use of consumer cameras, thus promising to vastly increase the amount of detailed, up-to-the minute information available about the real world. The volume of potentially interesting footage brings up the question of how end-users can avoid being glued to one (or worse, many) streams of videos waiting for events of interest. In this demo, we present Lookout, a system that allows users to register standing queries, called *triggers* over live video streams. Lookout then notifies the user when events of interest to them occur in their streams of interest. For example, a user could point to a cat cam and write a trigger that sends a notification when the cat wakes up and starts moving. Users can also write triggers to look for certain news being covered in a live new channel, a gamer moving to a certain level in a Twitch stream, a stranger showing up in a outdoor surveillance camera, etc.

Research: Lookout allows users to specify streams of interest and write standing queries on a webpage or mobile app, processes these queries on a continuous basis on a cloud backend and notifies the user when the queries are satisfied. In doing so, Lookout addresses several research challenges. It proposes an expressive but efficient query language that allows users to specify useful spatial and temporal queries over video. It provides a cloud-based architecture to index and serve these queries. It includes several techniques to make cloud-based processing efficient, including feedback-based frame selection, decoupled visual processing and text-based search, cost-effective division of labor between
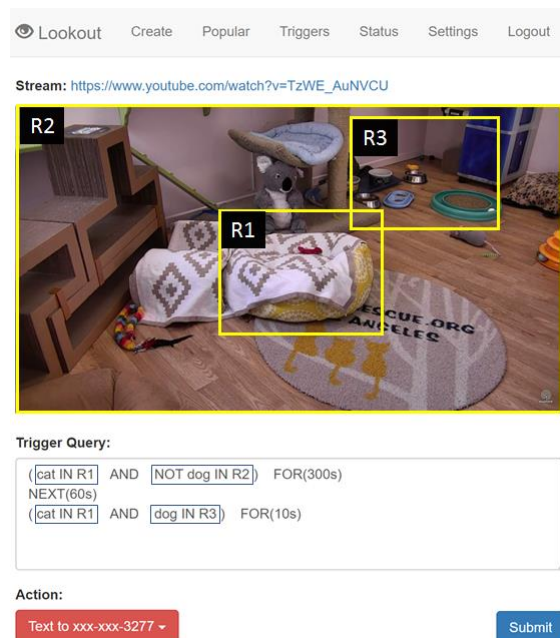
Figure 1: Trigger on a live foster cam with a cat and a dog. The trigger looks for "dog stealing cat's food when the cat is asleep".

CPU and GPU, and techniques for efficient execution of visual processing algorithms, especially Deep Neural Networks (DNNs).

Demo: The demo will provide a web page that allows users to add new live streams of interest to them, including, e.g., those from YouTube Live. We will also have one or more on-site cameras from which we will stream live. For each stream, the user can specify a trigger consisting of *atomic queries*, which are regions of interest with entities (object names, text and faces) of interest, and compositions of these atoms using AND, OR, NOT, FOR (indicating temporal extent of the atom), NEXT (indicating temporal sequencing of atoms), WITHIN (indicating temporal neighborhood next to an atom). The user can associate actions such as text/email-based notification or recording the video snippet. Figure 1 shows the current interface for Lookout with a complex trigger.