Lookapp for Ads – Content-based Advertising by Visual Concept Detection*

Damian Borth
Dept. of Computer Science, University of Kaiserslautern
Kaiserslautern, Germany
d_borth@cs.uni-kl.de

Adrian Ulges
German Research Center for Artificial Intelligence
Kaiserslautern, Germany
adrian.ulges@dfki.de

ABSTRACT

We introduce *Lookapp for Ads*, a system that employs multimedia analysis technology to automatically detect a wide range of semantic concepts (such as locations, objects, or actions) appearing in web videos. Though such automatic tagging remains a challenging problem and accuracy is limited, it can significantly increase the accessibility of poorly tagged videos (which usually form the vast majority on web video portals). Also, by inferring the topics of videos at large scale, we can improve the targeting of ads, a challenge that is of vital interest in web video monetization.

Given the enormous scale of web video, it is vital for our system to operate autonomously. Therefore, we employ cutting-edge machine learning techniques for an scalable learning from web video content, and an automatic selection of the vocabulary of target concepts by mining the web for trending topics of interest.

1. INTRODUCTION

While the web-based video market continues to grow rapidly with respect to user communities and numbers of views [1], monetization is still a challenge: For example, YouTube as the market leader still places advertisements only for a fraction of its video views (according to [2, 4] only 3 Billion out of 21 Billion weekly views on YouTube are monetized). Another key challenge to web video services poses the lack of proper information describing the content of videos as services like search and recommendation require annotations in form of title, tags, and descriptions. Correspondingly, the majority (99%) of video views on YouTube are generated by a few (30%) highly popular videos [3]. Overall, web video portals are currently struggling both to monetize their views and to grant efficient access to vast parts of their content.

Fig. 1 shows a typical YouTube video that is weakly tagged and thus difficult to monetize. This clip has no meaningful title nor tags or a proper description. However, the information that this video shows two dogs – if available – would make it possible to target in-stream advertisements within the segment *Pets & Animals → Pets → Dogs*, each being worth $4.44 - 10.96 \in$ [1].

Figure 1: A YouTube video showing two playing dogs but not providing any meta-data describing the content (a).

Unfortunately, the enormous volume and growth of web video (with 48 hours of new video content being uploaded every minute [2]) renders a careful manual annotation of videos infeasible. One solution to this problem is *concept detection* [9], an automatic computerized tagging of videos with semantic concepts appearing in them, like locations, objects, or actions. For example, a concept detector trained on detecting dogs could be applied on the above video, detect the concept "dog" and write this information into the database making the video available for target advertisement. Though concept detection remains one of the biggest challenges in computer vision with accuracy remaining far from human annotation [8] even a concept detector that is limited in accuracy is a highly interesting approach for both, better targeting ads for poorly tagged videos and also other applications like keyword-based search, video categorization and content-based filtering of violence or pornography.

2. SYSTEM OVERVIEW

Towards this goal we propose the following system and its components as illustrated in **Fig. 2**:

Figure 2: Left: Trending topics are discovered and potential semantic concepts (like "iPhone 5") are selected. **Middle**: For each selected concept, an automatic detector is trained using our *Lookapp* system. **Right**: Poorly tagged and untagged web videos are processed by the concept detectors, and detected concepts are used to improve the accessibility of these videos and link them with targeted ads.

2.1 Discovery of Trending Topics

The creation of a concept vocabulary is usually narrowed down to the domain of the system like "news" or "sports" and fixed during its livespan [7]. However, for a practical system to operate on the hugely diverse and dynamically changing web video space, we have to tailor the concept vocabulary over time, adapting to new concepts of interest rising in media and society. One way to do this is to discover trending topics by analyzing Google search patterns, real-time communication on Twitter and Wikipedia access statistics and continuously select those among them which are suitable for visual learning. This way the system could become aware of the Zeitgeist and keep its vocabulary up to date.

2.2 Visual Learning from the Web

As concept detection requires training material, our system automatically exploits YouTube content for training (so to say, the system learns the appearance of the concept *soccer* by "watching" soccer scenes on YouTube). This way, a time consuming manual acquisition of training videos is omitted. Additionally, by using *Lookapp* [6] – our concept detection system built entirely on distributed cloud computing technology – the time required to build a detector can be reduced from weeks to days. To accomplish this, Lookapp first maps the concept to a YouTube query [5], retrieves videos and extracts visual features from them for a subsequent classifier training using Google's Prediction API as an on-demand supervised machine learning service.

2.3 Advertisement Placement

Once being able to detect concepts in weakly annotated videos, those can be either treated as keywords in keyword auction systems like AdWords[2] or as input for vertical categorization as used in YouTube's Video Targeting Tool[3].

3. CONCLUSION

Summarizing, the presented system allows to automatically annotate videos making them accessible and available for targeted advertising. Further – once initialized – the system can evolve over time by expanding its vocabulary and

building new concept detectors dynamically exploiting the virtually unlimited resources of common cloud computing services.

4. REFERENCES

[1] ComScore August 2011 U.S. Online Video Rankings. available from `http://www.comscore.com/Press_Events/Press_Releases/2011/9/comScore_Releases_August_2011_U.S._Online_Video_Rankings` (retrieved: Sep'11).

[2] YouTube Blog: Thanks YouTube Community. available from `http://youtube-global.blogspot.com/2011/05/thanks-youtube-community-for-two-big.html` (retrieved: Jun'11).

[3] YouTube Blog: YouTube Videos now served in WebM. available from `http://youtube-global.blogspot.com/2011/04/mmm-mmm-good-youtube-videos-now-served.html` (retrieved: Sep'11).

[4] YouTube Press Statistics. available from `http://www.youtube.com/t/press_statistics` (retrieved: Sep'11).

[5] D. Borth, A. Ulges, and T. Breuel. Automatic Concept-to-Query Mapping for Web-based Concept Detector Training. In *Proc. Int. Conf. Multimedia*, November 2011.

[6] D. Borth, A. Ulges, and T. Breuel. Lookapp - Interactive Construction of Web-based Concept Detectors. In *Proc. Int. Conf. Multimedia and Retrieval (demo session)*, April 2011.

[7] M. Naphade, J. Smith, J. Tesic, S. Chang, W. Hsu, L. Kennedy, A. Hauptmann, and J. Curtis. Large-Scale Concept Ontology for Multimedia. *IEEE MultiMedia*, 13(3):86–91, 2006.

[8] A. Smeulders, M. Worring, S. Santini, and A. Gupta R. Jain. Content-Based Image Retrieval at the End of the Early Years. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.

[9] C. Snoek and M. Worring. Concept-based Video Retrieval. *Foundations and Trends in Information Retrieval*, 4(2):215–322, 2009.

---

[2]`http://adwords.google.com/`

[3]`http://www.google.com/videotargeting`