XXVI Congresso de Iniciação Científica Unicamp

17 a 19 de outubro   Campinas | Brasil

Pornographic Cartoon Video Detection through Deep Neural Networks

Akari Ishikawa*, Mauricio Perez, Sandra Avila

Abstract

Despite the efficient solutions in pornography detection literature, specific solutions for sensitive content in cartoons have not been developed yet. In this work, we evaluate how state-of-the-art solutions for natural videos (with humans) perform in cartoons. Also, we propose a new method with higher accuracy, showing that treating cartoons independently can improve sensitive content filtering.

Key words:
Deep Learning, Pornography Classification, Cartoons

Introduction

Children that have grown up with technology, the so-called "digital natives", spend most of their recreation time on the internet, commonly watching cartoons. According to the United Nations Children's Fund (UNICEF), children and teenagers represent (in 2016) one in three users on the internet [1]. Also, another study showed that 30% of all content on the web is pornographic [2], making it easy for kids to watch porn cartoons accidentally. Although we have vast literature in pornography detection, the solutions are all focused on videos with humans (natural videos). As far as we know, there is no research related to cartoons and sensitive content (inappropriate for children). In this project, we first evaluate a convolutional neural network (CNN) trained by a recent work that achieved 97.9% of accuracy in detecting porn in natural videos [3], and we compare its performance with a CNN explicitly trained for cartoons.

Results and Discussion

As a first to explore the sensitive cartoons problem, we built a database with 544 non-porn cartoons and 195 porn cartoons. Inspired by the methodology proposed by Perez et al. [3], we extracted static information (i.e., raw frames) and motion information (i.e., optical flows) from the dataset. As shown in Figure 1, we fed these information to a CNN that works as a feature extractor. The extracted features are then pooled, forming a description for each video. A Support Vector Machine classifier outputs the probability that a video is porn (i.e., 1 for porn and 0 for non-porn) for each type of information (raw frames and optical flows). Also, that information can be combined in a process called Late Fusion.
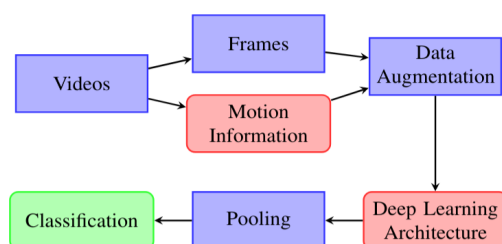

Figure 1: Overview of the performed methodology.

As a preliminary experiment, we first extracted the features using the CNN of Perez et al. [3], specialized in detecting pornography in natural videos. The results are showed in Table 1.

Table 1. Classification using a feature extractor trained to recognize porn in natural videos.

	Frames	Optical Flows	Late Fusion
F2-measure	80.9%	90.1%	91.0%
Accuracy	86.6%	93.1%	93.7%
Total	330/370	346/370	347/370

Next, we trained the model with our porn cartoons, initializating the CNN with the porn classifier weights. The final classification with the new feature extractor achieved scores even higher than the ones present in pornography classification literature (Table 2).

Table 2. Classification using a feature extractor specialized in porn cartoons.

	Frames	Optical Flows	Late Fusion
F2-measure	96.1%	93.9%	98.1%
Accuracy	97.5%	96.0%	98.7%
Total	363/370	364/370	365/370

Conclusions

Our results with frames are already competitive with state-of-the-art solutions. Besides, the combined use of static and motion information proved to be superior to state-of-the-art solutions in pornography classification. As future work, we intend to make a real-time filtering application for mobile platforms.

Acknowledgement

[1]Livingstone, Sonia; Byrne, Jasmina; Carr, John (2016). "One in Three: Internet Governance and Children's Rights," Innocenti Discussion Papers no. 2016-01, UNICEF Office of Research - Innocenti, Florence.
[2]Short, Mary B., et al. "A review of Internet pornography use research: Methodology and content from the past 10 years." Cyberpsychology, Behavior, and Social Networking 15.1 (2012): 13-23.
[3]Perez, Mauricio, et al. "Video pornography detection through deep learning techniques and motion information." Neurocomputing 230 (2017): 279-293.