ADULT CONTENT DETECTION
IN IMAGES & VIDEOS

kuznech.com
The working principle of the Internet is such that anyone who sets up a computer and connects it to the web becomes able to share with the world any type of information. Generally, some of the information dispatched is not appropriate for viewing by children and teens, and some steps should be taken to classify and control access to it.

Detection of pornography is a significant subtask of content filtering problem on the internet. Many studies have estimated that of the world's 42 million websites, 12% contain adult content. Another statistical result estimated 70% of teens have inadvertently come across pornography on the web. This is, of course, alarming.

The problem of detecting an adult-oriented content in images still remains unsolved and is widely addressed by many researchers. Over the past years there have been many approaches to detect and filter pornography. Authors had tried to detect adult images by considering color, texture and geometrical features of an image, proposing novel approaches, algorithms and methods.

To contribute to solving the problem, Kuznech proposes an innovative approach based on logotype detection, text recognition and scene classification by convolutional neural networks (CNN).
II: APPROACH

The task of Kuznech Adult Content Detection System is to detect pornographic content using all available information, both textual and visual. To interpret the results, we used a three-color scale graduated by the detector’s level of “confidence”:

- **Red**: definitely contains unwanted content
- **Yellow**: may contain prohibited materials
- **Green**: adult content free

During the R&D phase, we prepared a database of porn videos, which was afterwards scientifically analyzed. The results showed that:

i) Most professional pornographic videos contain special “markers”, as:
- studio logos
- studio names
- warning text screens.

ii) Almost 30 most typical and frequent scenes and images of specific objects come up in the overwhelming majority of porn videos.
III: MULTILEVEL ADULT-CONTENT DETECTION SYSTEM

To detect adult content within images and videos, Kuznech developed a complex system based on the company’s own software solutions. Below we briefly describe some of the methods that are used in our content filtering system:

a) Logo Detection

To mark their videos, porn studios use special markers: a studio logo or the studio name (see Fig. 1). Kuznech took advantage of these markers as “signals” that the video may contain sexual (pornographic) scenes.

![Figure 1. Examples of porn studio logotypes](image)

b) Warning Text Detection

According to the law, legal porn studio are required to put a textual warning at the beginning of their video (as in Fig. 2) that indicates the presence of sexual content in the video.

In these warning texts Kuznech revealed keywords (or key phrases) that are definite markers of the prohibited content (such as “sexually explicit conduct”, “sexually explicit material”, “adults only”, “title 18 united states code section 2257” etc.). Should at least one of the keywords be present on a frame, or the file contains a porn studio logo, the video is marked “red” as definitely containing adult content.
Scene Classification

Although the methods mentioned above are effective in detecting adult content, not all pornographic videos have warning screens or logotypes interposed. The Web is flooded by a huge amount of low quality “home videos” coming from unknown sources.

By analyzing more than 500 hours of “adult” videos, Kuznech identified 30 most typical and frequently showing explicit scenes, like:

- images of the genitals and their combinations,
- different poses and angles of typical scenes of copulation, or preparation for this process.

Should at least one object or scene be detected, the video is marked as prohibited.
We trained our neural network to recognize over 1000 (one thousand) categories: 20 classes related to porn scenes, 950 categories were taken from URLs provided by ImageNet Large Scale Visual Recognition Challenge (ILSVRC), and 30 categories were collected from Kuznech internal datasets (fashion photography, goods images etc.). **Total training dataset of Kuznech contains around 1.3 million images.**

Figure 3. Example of typical sexual poses
Figure 4. Example of video analysis with different methods
Kuznech classified almost 700 test videos using detectors separately and together. The test base consisted of 48 porn videos (“positive”) and 95 non-porn files (“negative”) which have high quality (720p); 100 positive and 97 negative examples for middle quality (360p); 98 positive and 85 negative examples for low quality (240p). Additionally, Kuznech collected 100 porn videos with warning screens and logotypes of studios from Fig. 1 and 65 videos with non-porn logotypes and text. The results were obtained by exponential approximation of probabilities taken from detectors on each video from testing base. ROC curves of detection results on each testing subset are shown in Fig. 3. According to Kuznech results, the proposed pornography detector found most of the videos with adult content and did a sufficiently small percent of “false positive” errors.

### IV: RESULTS

<table>
<thead>
<tr>
<th>Video quality</th>
<th>Error, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>240</td>
<td>8,0</td>
</tr>
<tr>
<td>360</td>
<td>8,0</td>
</tr>
<tr>
<td>720</td>
<td>8,1</td>
</tr>
</tbody>
</table>
V: PERFORMANCE DATA

The GPU realization of Large CNN processes 50-150 videos/minute (depending on video quality), while the CPU-based feed-forward routine processes 5 videos/minute.

<table>
<thead>
<tr>
<th>Requirements</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intel® Xeon® Processor E5-2630L v2 (15M Cache, 2.40 GHz) 64 Gb RAM NVidia Tesla K40</td>
<td>GPU: up to 150 videos/minute</td>
</tr>
<tr>
<td></td>
<td>CPU: 5 videos/minute</td>
</tr>
<tr>
<td>x 2 units</td>
<td></td>
</tr>
</tbody>
</table>

VI: LICENSE POLICY

<table>
<thead>
<tr>
<th>License lease</th>
<th>Non-Exclusive License</th>
<th>Exclusive License</th>
<th>Exclusive License with a training set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Porn detecting option</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Deployment on customer’s server</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Source codes</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Training set</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>
KUZNECH.COM
Demo: PORNDETECTION.KUZNECH.COM
E-mail: WELCOME@KUZNECH.COM