POSTER **P5178** | CONTACT NAME **Denis Timoshenko: Denis.Timoshenko@kuznech.com**

GPU TECHNOLOGY CONFERENCE

www.kuznech.com

Adult-Content Detection in Video with the Use of NVIDIA GPU

Michael Pogrebnyak, Denis Timoshenko, Ivan Burcev, Anton Kulinkin

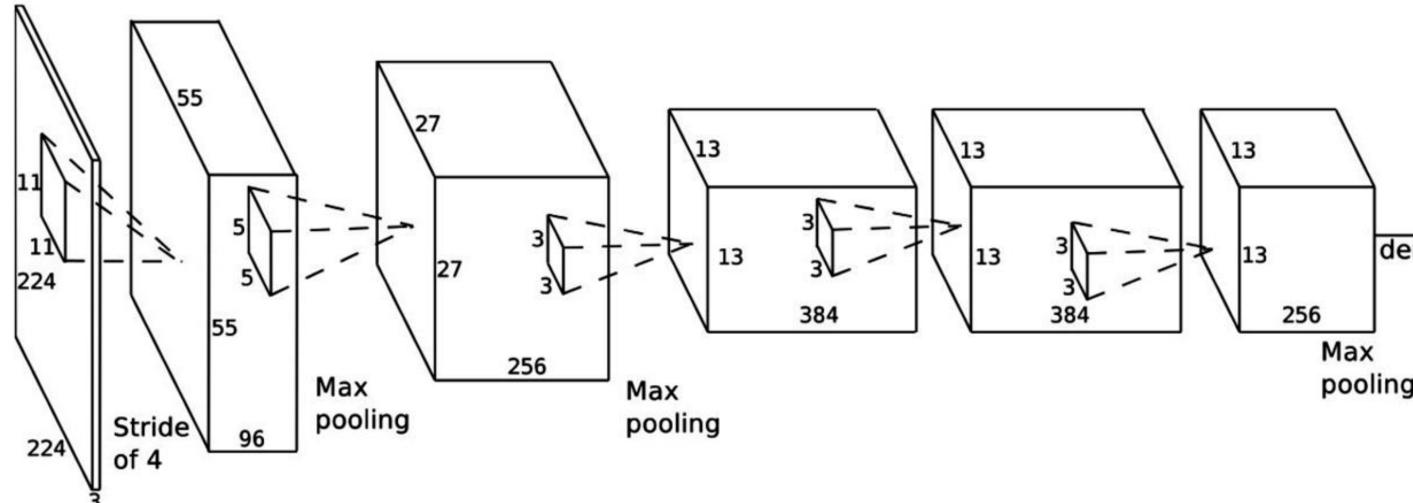Visual object recognition is one of the most important challenges in computer vision, and automatic adult-content detection in video, a subtask of visual object recognition, is a problem of interest for many internet companies.

The pornography industry has become a big business, and the Web has made it available to a large audience: nowadays even an underage person can easily gain access to "adult" materials. Thus, one of the biggest challenges for most social networks and video sharing websites is to prevent the uncontrolled uploading and sharing of adult content.

Methodology:

1. First, we used **logos of pornography film studios** as "signals" that the video may contain sexual (pornographic) scenes.
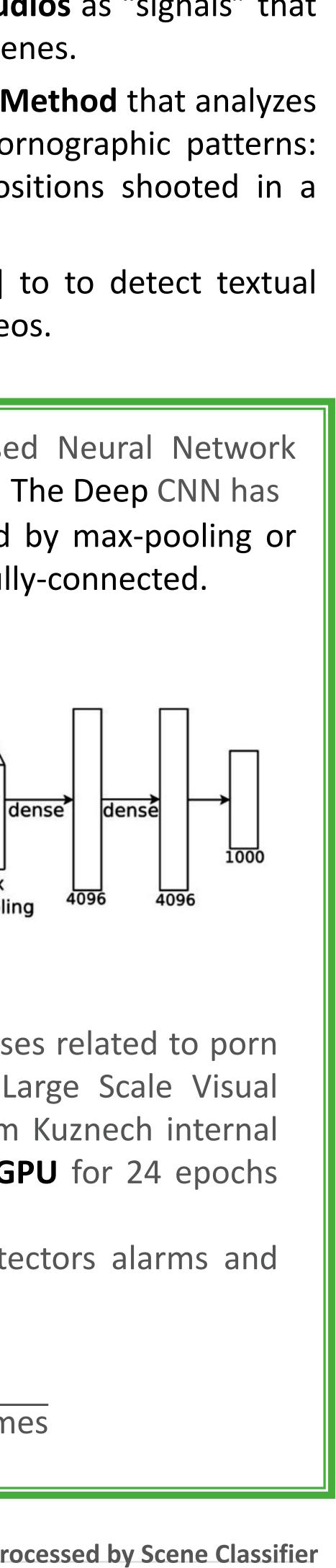
2. Secondly, we developed a **Scene Classifying Method** that analyzes individual frames and finds most common pornographic patterns: typical copulation scenes, different sexual positions shooted in a variety of camera angles.
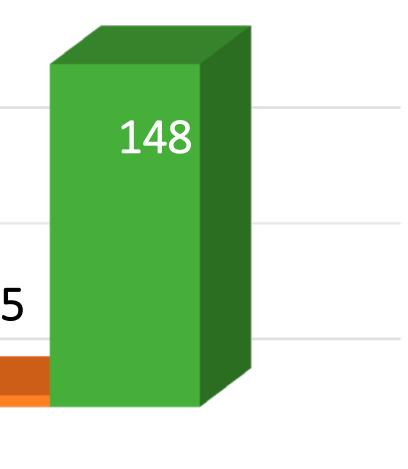
3. Additionally, we applied **Tesseract OCR** [2] to to detect textual warnings on the frames in the beginning of videos.

Logo Detector

For the purpose of **logo detection** we used a combined logo detector [2] based on decision trees (cascades) and Convolutional Neural Network (CNN). We trained one cascade per each company logo using Gentle AdaBoost algorithm and one common CNN for all cascades. We trained the network to filter the noises on single **NVIDIA Titan GPU** for 360 epochs during 3 hours.

The network has 6 trainable layers:
- 1st and 3rd layers are convolutional,
- 2nd is max-pooling,
- 4th is average-pooling,
- last two layers are fully-connected.

The output layer has two neurons with softmax activation functions, labelled as "porn" (positive class) and "non-porn" (negative class).

Scene Classifier

For a purpose of **scene classification** we used Neural Network architecture, similar to the one described in [3]. The Deep CNN has eight trainable layers: the first five layers are convolutional (followed by max-pooling or local response normalization layers), while the other three layers are fully-connected.

The network has been trained to recognize 1,000 categories: 20 classes related to porn scenes, 950 categories taken from URLs provided by ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [4], and 30 categories collected from Kuznech internal datasets. The network has been trained on a single **NVIDIA Titan GPU** for 24 epochs during 95 hours.

For each video from test database we gathered a statistics of detectors alarms and calculated the simple score:

$$\text{Probability of "porn"} = \frac{\text{number of alarmed frames}}{\text{total number of considered frames}}$$

References

[1] https://code.google.com/p/tesseract-ocr/
[2] Timoshenko D., Grishkin V., Smirnov E. Effective false positive reduction in multilevel face detection system using convolutional neural networks // ICCTPEA. – 2014. – P. 187.
[3] A. Krizhevsky, I. Sutskever, G.E. Hinton. ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems 25 (NIPS'2012), 2012.
[4] http://www.image-net.org/

Method	Detection rate
Tesseract OCR	21,4 %
Logo det. + OCR	58,8 %
Scene classifier	91,8 %

Percentage of correct porn detections performed by the described methods (tested on HD video 720p)

Video quality	Detection rate
240p	90,3 %
360p	89,8 %
720p	91,8 %

*Percentage of correct porn detections performed by **Scene Classifier** on videos of different quality*

Videos per minute, processed by Scene Classifier

- 2 x Intel Xeon 2630L v2: 5
- NVIDIA Tesla K40: 148