Predicting video game properties with deep convolutional neural networks using screenshots

Przemysław Buczkowski[1,2*], Antoni Sobkowicz[1]

[1]National Information Processing Institute, Warsaw, Poland
[2]Faculty of Electronics and Information Technology, Warsaw University of
Technology, Warsaw, Poland
pbuczkowski@opi.org.pl

The amount of visual data is enormous nowadays and it constantly increases. On the other hand information in images is still almost not readable for computer algorithms. In most of image classification problems humans still outperform computers. Sometimes it is not obvious who will perform better, because little is known about the classification problem. Good example of such case might be predicting various video game properties based only on in-game screen shots.

The authors have crawled the Steam platform storing information about genre, date of release, PEGI rating, etc. for thousands of video games. It is difficult to pinpoint a model that can infer those properties by looking only on images of gameplay. During research authors have focused on using standard deep convolutional neural networks in the first part of the network and fully connected layers at the end. As well as working on model we are preparing human annotator based baseline benchmark.

Interesting regularities have been found, e.g. approximation of release year is very difficult task on small images, because lot of details are lost. Downscaling an image is similar to anti-aliasing - in-game models look much less edgy. Both humans and CNNs struggle with this task. For now it is hard to tell who performs better because of small amount of human-annotated data. One of the notices made during the research was that people who don't describe themselves as gamers perform poorly on this task, and experienced players recognise the game from the image and then recall release date from memories. In the nearest future authors plan to use bigger images or even 1:1 crops for the task of release year approximation and extend our database of human-annotated images.

Fig. 1. Low resolution examples of screens: The Book of Unwritten Tales 2 (2015) and Counter Strike (2000).