1 slide

Good afternoon, dear colleagues! I want to tell you about “Application of Satellite Image Segmentation for Urban Planning Optimization”

2 slide

Here is presented an aerial photo of the district in Shanghai from Google map. For all of us it is not problematic to observe contours of buildings. However, the problem of automatic segmentation of satellite images is challenging.

3 slide

4 slide

5 slide

6 slide

7 slide

Nowadays this problem is in the focus of research community. In machine learning, the segmentation is usually formulated as a pixel labeling task. So the goal of our research was development of effective algorithm for building detection on satellite images based on deep learning methods.

8 slide

Nowadays, in solving modern problems of computer vision, CNNs which are usually developed using Keras open framework with Tensorflow as a backend, exceed traditional methods from sklearn-library, such as k-means clustering or EM-algorithm, and even work of experts.

9 slide

However, the usage of such methods instead of traditional approaches is non-trivial, because they should:

Take into account the small size of objects.

Be invariant to rotation.

Have enough training examples.

Handle huge images.

Cope with noise.

10 slide

Now I’m going to present developed deep learning algorithm. U-Net has about 8 mln trainable parameters. It consists of two parts: an encoder and a decoder. The encoder has the typical architecture of CNN and contains four blocks. The decoder contains the same number of blocks as an encoder. The last layer of CNN uses a 1x1 convolution to match each component vector to 1 of 2 classes (buildings or not-buildings).

11 slide

A database of images is the important part for learning and efficiency evaluation of different machine learning algorithms. To research the developed network there was used Planet database.

This dataset contains 14 RGB images in JPG format

The aerial photos was marked up manually, using tools of supervise.ly

Each sample includes more than 65 mln pixels.

On these photos there were shoted 3 Russian cities: Moscow, Yaroslavl and Rybinsk.

12 slide

In spite of little amount of images, extracting methods, such as data windowing and data augmentation, allow to crop smaller images with resolution 512x512 px. As a result the training set contains about 20000 samples and the test set contains about 5000 photos. Train and test samples did not have same pictures.

13 slide

To estimate developed model there was used Jaccard index of similarity. The launch of the CNN was carried out on the supercomputer NVIDIA DGX-1 of Artificial Intelligence Center of P.G. Demidov Yaroslavl State University. As a numerical optimization algorithm, Adaptive Moment Estimation optimizer (Adam) was chosen. As a loss function, binary cross entropy function was used. On every training iteration the model updated its weights using the batch of 36 samples. The model ended its training after completing 100 epochs. In conclusion, the training process lasted about 4 hours.

14 slide

The dependency of Jaccard index on training epochs is presented on this slide. The highest value of index was 0.65.

15-16 slides

Here, some test results are presented on these slides.

17 slide

This model can put into practice and applied to real data from satellites to solve problems of planning new urban districts or building control. Moreover, it might be applied to search for best locations for outlets, depending on the population of the block.

18 slide

To conclude it is claimed that CNNs can be efficiently used for building detection on aerial photos. Therefore, U-Net was developed and used for satellite images segmentation. To evaluate the quality of developed algorithm Jaccard index was choosen. After training and testing the model on supercomputer NVIDIA DGX-1, the highest value of Jaccard index was equal to 0.65.

19 slide

Regarding to future plans we intend to try algorithms of seeking building corners and mask search models. In addition, it might be useful to try more complicated loss function or shift color channels.

20 slide

(Read or skip)