MODERN METHODS AND TECHNOLOGIES OF DEEP LEARNING IN COMPUTER VISION

DESCRIPTION

The course examines the practical application of deep learning for solving actual problems of computer vision during developing video surveillance systems.

The following topics are covered.

1. Goals and tasks, course structure. The general scheme of solving computer vision problems using deep learning (from preparing data to assessing the quality). Overview of software tools used at each step.
2. Image classification with a large number of categories using deep learning.
3. Overview of the Intel Distribution of OpenVINO toolkit (Inference Engine, OpenCV, Open Model Zoo).
4. Object detection in images using deep neural networks.
5. Deep models for tracking objects in videos.
6. Semantic segmentation of images using deep learning.
7. Preparing synthetic data based on generative adversarial networks.

The course is practice-oriented. It includes 14 hours of lectures (7 lectures of 2 academic hours each) and 18 hours of practical training with the team performing of 4 tasks. Lectures are held in the classical form and are accompanied by examples of solving practical tasks using the Intel Distribution of OpenVINO toolkit. The implementation of practical tasks assumes that students are divided into groups of 2-3 people, choose a task for subsequent implementation. Further, individual consultations are conducted for groups during the implementation of practical tasks. The mode of collective development is simulated, when students demonstrate skills in using team development tools. The final control of knowledge assumes presentation of the developed project.

The course is aimed at engineers, teachers and researchers, as well as postgraduate students and students of higher educational institutions.

PRELIMINARY REQUIREMENTS

The course is aimed at students who have basic programming skills in the C++ and Python programming languages. Along with this, the course requires theoretical knowledge in the area of image processing, computer vision, machine learning and deep learning.

THE COURSE STRUCTURE

The course schedule is as follows.

Topic	Content	Hours	
		Lecture	Practice
Lecture 1. Goals and tasks, course structure. The general scheme of solving computer vision problems using deep learning (from preparing data to assessing the quality). Overview of software tools used at each step.	Statement of computer vision tasks considered within the course (classification/recognition, detection, segmentation). The general scheme of solving computer vision problems using deep learning: 1. Preparing and labeling data. 2. Study of existing models and quality measurements. Preparing (search for existing or development) tools for assessing the quality of models. 3. Transfer learning of deep models for solving the problem. 4. Training and testing the model.	2	

	5. Modification of the model, training and testing. 6. Model compression. Overview of software tools used at different stages to solve problems considered in the course.		
Lecture 2. Image classification with a large number of categories using deep learning.	Overview of modern deep neural networks for classifying images with a large number of categories (LeNet, AlexNet, VGG, GoogLeNet, ResNet). Key features of model architectures; problems solved in these architectures; the model complexity (the number of parameters, the amount of memory, the complexity of calculations).	2	
Lecture 3. Overview of the Intel Distribution of OpenVINO toolkit (Inference Engine, OpenCV, Open Model Zoo).	The main components of the Intel Distribution of OpenVINO toolkit. Use the Inference Engine to implement inference of deep models. Overview of OpenCV and its module dnn. Using the Inference Engine as a backend for inference of deep neural networks. Open Model Zoo.	2	
Practice 1. Image classification with a large number of categories using deep learning.	Transfer learning to image classification. Direct application of trained models using the Inference Engine component of the OpenVINO toolkit.		5
Lecture 4. Object detection in images using deep neural networks.	Problems of object detection in images (different scale, angle, many objects). Overview of modern deep neural networks for detecting objects of different classes on images (R-CNN, Faster R-CNN, Mask R-CNN, SSD, YOLO, MobileNet-SSD, etc.).	2	
Practice 2. Object detection in images using deep neural networks.	Direct application of trained object detection models from the Open Model Zoo, using the Inference Engine or the dnn module of the OpenCV library.		5
Lecture 5. Deep models for tracking objects in videos.	Deep topologies for the feature extraction, embedding deep models to the algorithm of object tracking.	2	
Practice 3. Video object tracking.	Implementation of the Hungarian algorithm for matching bounding boxes obtained on a pair of consecutive video frames.		4
Lecture 6. Semantic segmentation of images using deep learning.	Semantic segmentation problems (object scale, output image resolution). An overview of modern deep neural networks for semantic segmentation (encoder-decoder architecture, application of CRF, dilated convolutions, UNet-like architectures).	2	
Lecture 7. Preparing synthetic data based on generative adversarial networks.	Generation of synthetic data using generative adversarial networks, increasing the amount of the training set in tasks with a small amount of labeled data.	2	
Practice 4. Solving the problem of video analytics, including detection, recognition and tracking of objects in the video.	The solution of the problem involves the integration of modules developed in the previous practices.		4
EXAM: project presentation, exam/credit			
Total:		**14**	**18**

COURSE MATERIALS

Course materials:

1. *Syllabus*.

2. *Lectures*: 7 lectures. There are a presentation and an annotation for each lecture (2-3 pages with the list of recommended literature).

3. *Practical classes:* instructions for implementing practical tasks; a case study in source code that demonstrates the implementation of practical tasks, and it is supported by a step-by-step description; list of possible practical assignments. Step-by-step tutorials will be published on GitHub.

LANGUAGE

Russian

English

LICENSE

Free access. We plan to publish all materials on the UNN web site under Apache 2.0 license [http://www.apache.org/licenses/LICENSE-2.0.txt].

PROJECT TERMS AND STAGES

Term of the project: 16.10.2019 – 29.02.2020.

One stage of the project with intermediate control of readiness of lecture and practical materials approximately 15.01.2020.

AUTHORS

Turlapov Vadim Evgenievich, Dr., Prof., department of computer software and supercomputer technologies, Institute of Information Technologies, Mathematics and Mechanics, Nizhny Novgorod State University. Lead.

Vasiliev Engeny Pavlovich, lecturer, department of computer software and supercomputer technologies, Institute of Information Technologies, Mathematics and Mechanics, Nizhny Novgorod State University. Developer.

Getmanskaya Alexandra Alexandrovna, lecturer, department of computer software and supercomputer technologies, Institute of Information Technologies, Mathematics and Mechanics, Nizhny Novgorod State University. Developer.

Kustikova Valentina Dmitrievna, Phd, assistant professor, department of computer software and supercomputer technologies, Institute of Information Technologies, Mathematics and Mechanics, Nizhny Novgorod State University. Developer.