Center for Visual & Decision Informatics

LEARN-TO-SEGMENT by OPERATIONAL OBJECT SEGMENTATION NETWORKS

PROJECT ID | **TYPE |** [X] New [ ] Continuing **START DATE |** 1st August 2015

PROJECT LEAD/PARTICIPANTS | Serkan Kiranyaz, Moncef Gabbouj and Caglar Aytekin

DESCRIPTION | Object extraction in an image is a challenging problem since segmenting an object needs a combination of various visual cues such as texture information, distribution of the pixel intensities, object boundaries and dissimilarity from other objects/regions. Furthermore, the visual scenery may bear more than one object, or partially occluded objects. Unfortunately, image segmentation that can be defined as extracting meaningful regions from an image is an ill-posed problem since one has to first define and fix a priori what a "meaningful region" is. As a result, current automatic object segmentation algorithms suffer from various problems that make them unreliable and deficient for the purpose of (salient) object extraction. Learn-to-segment (L2S) proposes a novel operational neural network (ONN) in order to improve the traditional deep learning neural networks, in which "convolution" is the *only* operator used. We believe that extending the set of operations, both linear and non-linear, can lead to improving the generalization capability of the neural network. However ONNs, like their predecessor CNNs, will be classifiers that are proven to localize the object(s) in an image and can (learn to) extract the features for the purpose of image (object) classification. Therefore, we shall then modify the ONN topology, its training and outcome so as to serve to the L2S purpose to yield Operational (Object) Segmentation Networks (OSNs) where the output of a OSN will be the saliency map of the image from which salient object(s) can be segmented.

EXPERIMENTAL PLAN | During the project period the focus will be particularly drawn to the following key research topics:

- Develop the new "Operational" Neural Network systems to "deep" "learn-to-segment" objects.
- Develop Operational (Object) Segmentation Networks (OSNs) that can be trained by Back Propagation method.
- Perform deep vs. "Divide and Conquer" learning by OSNs to "learn" and "mimic" the human visual system for object segmentation.
- Test and explore the optimal OSN topologies and operators utmost segmentation accuracy.

RELATED WORK | A young and skilled team of doctoral students and recent doctorates will carry out the proposed work under the guidance and supervision of Academy Professor Moncef Gabbouj. The Team has won several paper awards, including the IBM Best Student Paper in ICPR 2014, whose topic represents a core subject in this proposal, and several international challenges, e.g. 2nd and 4th place in MSR-Bing Challenge, 3rd place in Face Antispoofing Contest. The MUVIS system built by the Team and former colleagues has been selected as a testbed for EU COST Action 292 and the national SHOK D2I project. The PI, a Fellow of IEEE, is a well-established world expert in the field of multimedia communication, currently holding the prestigious post of Academy of Finland Professor since 2011. He has been leading the Multimedia Research Group for nearly 25 years and managed successfully a large number of projects in excess of 12 Meuro.

HOW OURS IS DIFFERENT | In the L2S approach we propose an unprecedented intelligent system for object segmentation using novel Operational Segmentation Networks. The input to the proposed OSNs consists of raw image data and possibly supplementary information. The main supplementary information considered in the proposal are: the saliency maps, superpixels, edge fields and affinity matrices in multiple scales. A rich set of operations such as linear and non-linear filtering, correlation, and edge operations acting on the input image and possibly additional representations of the raw data will be considered.

MILESTONES FOR YEAR |

3 months: Development for ONNs and testing against CNNs

6 months: Design and development of OSNs and training using back-propagation method.

9 months: Test and evaluation of OSNs with different operators. Development of heterogeneous OSNs with multiple operator sets.

12 months: Final system deployment and speed optimization using the GPU programming to minimize the training/computation time.

DELIVERABLES |

1. Implementation of ONNs and OSNs.
2. Realization of potential capability of the "Divide and conquer" paradigm and the corresponding topologies of OSNs.
3. GPU programming and efficient utilization of system resources.
4. Realization and integration into the final system.

BUDGET FOR YEAR |

Student Salary	€45,900
Salary Overhead	€44,000
Travel and services	€24,100
Overhead	€15,000
Total	**€129,000**

ECONOMICS | As the foremost application, since the camera is one of the key selling points of smartphones, and increasingly important also in other types of consumer devices, the mentioned improvements in the image and video quality as well as in usability will benefit industry concerned with digital camera algorithms. The research results will be readily exploited by industry in several fields, including media, health, manufacturing, machine vision, intelligent homes and smart cities. Automatic object extraction and tracking in videos can also be performed with a great accuracy especially using an embedded L2S system. Other potential applications include augmented reality, adaptive image compression, world scan, depth photography, just to name a few.

POTENTIAL MEMBER COMPANY BENEFITS | A solid theoretical foundation from novel machine learning and computational paradigms is the basis of the proposed L2S approach. Thus, the CVDI projects can benefits from the proposed technique in both sense of segmentation and classification.

PROGRESS TO DATE | Initial implementation of Generalized Operational Perceptrons (GOPs) which can generalize traditional MLPs with different operators. Theoretical foundations for ONNs based on GOPs were also already accomplished.

KNOWLEDGE TRANSFER TARGET DATE | 12 months