Final Report of Project NKFIH KH-17 #126688 "Multimodal feature fusion for establishing novel 3D saliency models"

Scientific report

In the first year of the project, the main task was to analyze the adaptability of 2D image features among different application areas and to fuse them to create novel saliency models for multimodal sensor data processing. Convolutional neural network models were analyzed and tested, and then, in the second project year, the saliency models were fused with the networks' prediction maps to achieve better segmentation performance in different applications.

As an outcome of the project, **2 journal papers** (summa impact factor (calculated with the 2018 data) = 7.122), **5 international conference** and **3 national conference papers** have been published. Moreover, **2 journal papers** are under review.

The scientific report is divided into 5 sections corresponding to different research tasks related to the work plan.

Task 1. Direction selective contour detection approach for salient objects using feature fusion

A novel approach was proposed for salient object detection, which was published in IEEE Transactions on Circuits and Systems for Video Technology, with the Principal Investigator (PI) of the project as the only author [6]. The principle of the paper is to interpret a salient object as the *fusion of* its components: *the extracted contour and the inner texture*. As a novelty, to feature the contour part, an improved edge map was introduced that incorporated edge direction as a feature. The direction information in the small neighborhoods of image feature points was extracted, and the images' prominent orientations were defined for direction-selective edge extraction. Using such improved edge information, we provided a highly accurate shape contour representation, which we also combined with texture features. Then, the contour part was combined with a textural part in a salient object model, which served as an external force field map for automatic initialization, followed by an active contour segmentation of the image. The proposed model was able to handle complex contours more efficiently, even in the presence of cluttered background.
The introduced model was *tested on multiple widely known and used public databases*, including real world images. Altogether five public databases with more than 11000 images were used to evaluate the detection results. In the first part of the experiments, the proposed technique was compared to other parametric active contour models to confirm the advantages of the automatic initialization and the improved external force. In the second part, the proposed method was tested from the point of view of salient object segmentation, and it was

compared to numerous different state-of-the-art segmentation techniques on different data sets. The compared approaches included both traditional, geometric approaches and learning-based, convolutional neural networks, therefore, very different systems, applying various features were analyzed. With an iterative extension, the introduced algorithm was also able to detect multiple salient objects in the same sample image. The performance of the method was also evaluated quantitatively. The extensive evaluation results confirmed that the proposed method was performing with high accuracy and that - using the presented contributions – it outperformed existing parametric active contour models and achieved higher efficiency than the majority of the compared automatic saliency detection techniques.

Task 2. Image Based Robust Target Classification for Passive ISAR applying saliency information

To analyze the method’s *adaptability*, the salient object model was extended for automatic analysis of *passive radar 2D ISAR images* [1, 5] to evaluate the possibilities and capabilities of image feature based target extraction and classification, applying saliency information.

First, the novel image modality was analyzed and a pre-processing step for Inverse Synthetic Aperture Radar (ISAR) image was introduced, as input images can be of various size and resolution and they can contain targets with different meters/pixel resolution. Therefore, the raw inputs were resized to have a ratio consistent with their resolution. After this, the above mentioned salient object detection model was applied. Passive ISAR images are different from general imagery, however the main rules still hold: more distinct regions have higher statistical texture distinctiveness. Also, in such images the target is usually close to image center, which also attracts higher visual attention, therefore these image areas require higher distinctiveness. In case of ISAR images, the edges can be quite blurry, which makes the object outline detection quite challenging. The proposed salient object model was found to be able to handle these drawbacks efficiently, the 2D saliency features were adaptable for the ISAR images as well. Following the object detection, a classification step was introduced using the extracted features to recognize targets from the same class. For extensive evaluation, we have evaluated the proposed method on a database, which included 294 images of 12 targets. Experimental results confirmed that the presented method is fast, easily embeddable and extendable, works near real-time, and we also showed its viability for classification using real passive 2D ISAR images.

The method was developed in collaboration with colleagues from Department of Information Engineering, University of Pisa, Italy and CNIT Radar and Surveillance System (RaSS) National Laboratory, Pisa, Italy. The preliminary method was introduced in IEEE Radar Conference [1], at KÉPAF Hungarian conference [12], and the improved, final method was published in IEEE Sensors Journal [5]. In each case, the PI of the project was the first author.

Task 3. 3D multimodal medical data processing combining saliency and convolutional network features

Beside the passive ISAR, the project team also concentrated on the *processing of 3D multimodal medical sensor data* (CT and MRI) applying content-based information and saliency models. In this topic talented students from the PPKE ITK were involved, supervised by the PI of the project.

Task 3.1. MRI based Brain Tumor Segmentation

As part of this work, a brain tumor segmentation method based on visual saliency features on MRI image volumes was developed, the preliminary method was introduced in 16th International Conference on Content-Based Multimedia Indexing (CBMI) [2], in the Engineering Section of 4th International Interdisciplinary 3D Conference [4] and at KÉPAF Hungarian conference [10].

The proposed method used a novel combination of multiple MRI modalities as a first step to highlight possible lesions. After extensive analysis, FLAIR and T2 MRI modalities were included in the introduced pseudo-RGB color model, and as an improvement, healthy mean templates, generated from the annotated database slices without tumors, were also integrated in the color model. Therefore, FLAIR was selected as pseudo-R, and the difference between the actual T2 slice and the T2 healthy mean template as pseudo-G, the difference between the actual FLAIR slice and the FLAIR healthy mean template as pseudo-B is proposed. To further enhance local contrast, the pseudo-RGB image was transformed to the CIE L*a*b* color space, which was later used for saliency calculation.
The introduced method proposed a saliency model that included color and spatial features, exploiting the *a priori* knowledge on tumors: in the proposed color model, they had higher intensity than their surroundings and they were smaller connected components in the image. As a novel contribution, we incorporated 3D information as the relation of neighbouring slices in our model. When calculating the color-spatial saliency of the actual slice, the final map was computed as a weighting function, including the previous and next slice’s saliency as well. Tumors are not changing rapidly from slice to slice, therefore, their location and size is similar in these cases. Based on the saliency map, the outline of the tumor was detected by a region-based active contour method. The final results were further refined with post-processing steps as well.

Different state-of-the-art convolutional neural network models, designed for medical data processing, were analyzed and as a proof-of-concept method, the proposed saliency model was combined with the selected convolutional neural networks (2D U-Net and WT-Net) to reduce the networks’ eventual overfitting which may result in weaker predictions for unseen cases. As a result, the fusion of deep learning techniques with saliency-based, handcrafted feature models, the fusion approach had good abstraction skills and yet it was able to handle diverse cases for which the net was less trained.

Task 3.2. Automatic liver segmentation on CT scans

In a similar manner, we have also implemented a technique for liver segmentation in CT scans, the first version of the proposed method was introduced in the Engineering Section of 4th International Interdisciplinary 3D Conference [3] and at KÉPAF Hungarian conference [11]. The improved model was presented in 17th International Conference on Content-Based Multimedia Indexing (CBMI) [9].

A fusion method was developed which combines regional-based techniques and convolutional features. First, a pre-processing was introduced using a bone mask to filter the abdominal region (in case of whole body scans it is important) and a thresholding based on probability density function. Then a combination of region growing and active contour methods was applied for liver region segmentation. This traditional, so called “handcrafted” feature based technique was fused with convolutional neural network models to increase segmentation accuracy. 3D U-Net prediction results were first filtered, to keep only the largest detected blob in the abdominal region, and the prediction masks were further improved by removing most of the false positive voxels using the GrowCut method. This was followed by a shallow convolution of traditional and learning-based result masks.

The quantitative evaluation for the fusion method was performed on three different databases, including 2 publicly available ones, the VISCERAL and the SLIVER databases, and also on a private database. The results show that the proposed fusion method is uniformly efficient on different databases and achieves similar performance results when compared to the state-of-the-art on the publicly available SLIVER database.

Task 4. Fusion of salient and convolutional features on brain MRI scans

Based on the achieved results in Task 3.1., we continued our work to fuse traditional, “handcrafted” features (such as a saliency model) and learning-based, convolutional features on a deeper level. The proposed fusion model was introduced in a journal paper, which was submitted to Springer Multimedia Tools and Applications [7] and it is under revision.

In this work, we further improved the pseudo-color model introduced in [2], calculating the pseudo-RGB channels as difference images between a specific image patch and a healthy image template built using the healthy slices of the database for *FLAIR*, *T2* and also the *T1c* sequences.

This contribution was novel from multiple aspects, as state-of-the-art discriminative methods were usually applying only slices with malignant areas for training; therefore, a large part of the database was unused. Moreover, these healthy slices can help in highlighting tumor regions as differences from the healthy ones. Thus, these constructed pseudo-color difference images should be included in the training process.

The proposed algorithm followed the same workflow as [2], however, skipped the RGB to CIE L*a*b* color conversion of the pseudo color image, and instead, the color-based saliency is calculated on the RGB channels. According to our experiments, by applying the pseudo-

RGB difference image in the saliency calculation model, more information was exploited, therefore the segmentation performance was higher than with the converted Lab color space.
The proof-of-concept step of the fusion of the proposed saliency map and the prediction map of the trained convolutional neural networks (U-Net and WT-Net) was further analyzed and an extensive experimental evaluation was performed. Moreover, the idea of the healthy template based pseudo-RGB difference image was also integrated in a deeper level, and a retraining process of the traditional U-Net network was tested.
The proposed methods were tested on the BRATS 2015 and BRATS 2018 public databases, and the quantitative results showed that hybrid models (including both trained and handcrafted features) can be promising alternatives to reach higher segmentation performance. Moreover, healthy templates can provide additional information for the training process, enhancing the prediction performance of neural network models.

Task 5. Saliency-fused convolutional network for brain tumor segmentation

The achieved results of Task 3.1. and Task 4. showed that a saliency map and a CNN's prediction map can be efficiently fused, and by applying a shallow convolution of these maps, the segmentation accuracy can be increased. As a next step, we proposed an embedded model to incorporate additional saliency feature information in the convolutional neural network's training process. This work [8] was submitted to IEEE Transactions on Medical Imaging journal and it is under review.
As the main contribution of this work, we presented a convolutional network - SFUnet: Saliency-FUsed convolutional network - that incorporated additional feature information in the training process besides brain scan volumes and the associated tumor labels. The additional information was in the form of the aforementioned pseudo-color brain/tumor model images associated to the brain scan slices, and brain tumor region saliency maps enhancing the regions of interest of the slices.
These additional feature maps were not used as pre- or postfiltering steps or late-fused to the network predictions, but formed an inherent part of the training process of the network model. The architecture of the presented solution was inspired by the encoder-decoder/convolutional-upscaler design of the original U-Net convolutional network.
Thus, the proposed network used 3 main sources of information during the training process: the brain scans, the pseudo-colored brain slice images, and the generated saliency maps.
We extensively evaluated the presented model on the BRATS 2015 and 2018 brain tumor datasets, which produced results that support the viability of the approach and showed that in its current form the SFUnet model outperforms others over the 2015 dataset and can produces close results to the state-of-the-art over the 2018 dataset.
Our future work will concentrate on further improving the accuracy of the predictions by investigating deeper versions of the network and by including further data augmentation steps. Our goal is produce a generic architecture for saliency-fused training for aberration and change detection in medical visual data.

Application of the results

The analysed and developed saliency models were used in multiple application areas. A saliency-based target extraction algorithm was proposed for passive radar (ISAR) data, and the result of the extraction was then applied for also recognizing the target based on a classification step. The target objects – extracted as salient objects -, can be efficiently emphasized and classified, even within the noisy image characteristics. The algorithm was applied in the European Defence Agency (EDA) "**MAPIS** Multichannel passive ISAR imaging for military applications" project.

The saliency based segmentation algorithms for 3D medical image data included an automatic lesion segmentation model for brain MRI data and an automatic liver segmentation approach for abdominal CT data. These methods were successfully applied in the NGM GINOP-2.2.1-15-2017-00083 **zMed** project, where the aim was to integrate novel 3D sensors and augmented reality based visualization technology for innovative healthcare application to improve the diagnostic process; strengthen the doctor-patient relationship and also open up new horizons in the medical education.

Further scientific activities and events related to the project

I was a session chair in IEEE International Conference on Content-Based Multimedia Indexing (CBMI) 2018.
I was selected as steering committee member of NJSZT Hungarian Association for Image Processing and Pattern Recognition (KÉPAF) in January 2019.
Vanda Czipczer and Petra Takács (supervised MSc students at Pázmány Péter Catholic University, Faculty of Information Technology and Bionics, both working on the project) won 1st prize in 2019 National level Scientific Student Conference (OTDK), moreover they also won 4 national special awards for their work. Petra Takács was awarded with the "Novofer alapítvány Gábor Dénes Országos Diákköri ösztöndíj". Vanda Czipczer was awarded with the 1st prize of the IEEE HS Student Paper Contest in MSc category, with the "Doktoranduszok Országos Szövetsége különdíj" and with the "Nokia Bell Labs különdíj".

Our publications with project funding:

2018

[1] **A. Manno-Kovacs**, E. Giusti, F. Berizzi, L. Kovács, "Automatic Target Classification in Passive ISAR Range-Crossrange Images", 2018 IEEE Radar Conference (RadarConf'18), Oklahoma City, USA, April 23-27, 2018.

[2] P. Takacs, **A. Manno-Kovacs**, "MRI Brain Tumor Segmentation Combining Saliency and Convolutional Network Features", 16th International Conference on Content-Based Multimedia Indexing (CBMI), La Rochelle, France, September 4-6, 2018.

[3] Kriston, A., Czipczer, V., Manno-Kovács, A., Kovács, L., Benedek, Cs., Szirányi, T., „Segmentation of Multiple Organs in Computed Tomography and Magnetic Resonance Imaging Measurements", Proceedings of the 4th International Interdisciplinary 3D Conference: Engineering Section, pp. 51-56, Pécs, Hungary, October 5-6, 2018.

[4] Takács, P., Manno-Kovacs, A., „Brain Tumor Segmentation in MRI Data", Proceedings of the 4th International Interdisciplinary 3D Conference : Engineering Section, pp. 68-74, Pécs, Hungary, October 5-6, 2018.

2019

[5] A. Manno-Kovacs, E. Giusti, F. Berizzi, L. Kovács, "Image Based Robust Target Classification for Passive ISAR", IEEE Sensors Journal, vol. 19, no. 1, pp. 268-276, 2019, IF*: 3.076.

[6] A. Manno-Kovacs, "Direction Selective Contour Detection for Salient Objects", IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 2, pp. 375-389, 2019, IF*: 4.046.

[7] P. Takacs, L. Kovacs, **A. Manno-Kovacs, „**A fusion of salient and convolutional features applying healthy templates for MRI brain tumor segmentation**", Multimedia Tools and Applications, Springer, 2019, IF*: 2.101, bírálat alatt.**

[8] A. Manno-Kovacs, L. Kovács, "Saliency-fused convolutional network for brain tumor segmentation", Transactions on Medical Imaging, IEEE, 2019, IF*: 7.816, bírálat alatt.

[9] V. Czipczer, A. Manno-Kovacs, „Automatic liver segmentation on CT images combining region-based techniques and convolutional features", 17th International Conference on Content-Based Multimedia Indexing (CBMI), Dublin, Ireland, September 4-6, 2019.

[10] Takács, P., Manno-Kovács, A., „Agytumor Szegmentálás MRI Képeken, Szaliencia alapú algoritmussal és Neurális Hálózatokkal", KÉPAF 2019. Képfeldolgozók és Alakfelismerők

Társaságának 12. országos konferenciája, Debrecen, Magyarország, pp. 1-11. Paper: 18, 2019.

[11] Czipczer, V., Manno-Kovács, A., „Májszegmentálás orvosi képadatok tartalom alapú elemzésével", KÉPAF 2019. Képfeldolgozók és Alakfelismerők Társaságának 12. országos konferenciája, Debrecen, Magyarország, pp. 1-16. Paper: 19, 2019.

[12] Manno-Kovács A., Kovács L., "Automatikus célpont klasszifikáció passzív ISAR képeken", KÉPAF 2019. Képfeldolgozók és Alakfelismerők Társaságának 12. országos konferenciája, Debrecen, Magyarország, pp. 1-13. Paper: 17, 2019.