FINAL REPORT OF PROJECT
NRDI #K-120233

Instant environment perception from a mobile platform with a new generation geospatial database background

Principal Investigator: Dr. Csaba Benedek
Host Institute: ELKH SZTAKI
Duration: 01.10.2016-31.03.2021

FUNDED BY THE NATIONAL RESEARCH, DEVELOPMENT AND INNOVATION (NRDI) OFFICE

Table of contents

Summary of project achievements 3
Project results 4
Task 1 Large scale GIS data management and visualization 4
Task 2 Point cloud registration from different sensors 5
2.1 Registration of sparse point clouds captured for real-time scene analysis to dense reference point cloud maps 5
2.2 Fusion of high resolution optical images with low resolution depth maps 6
2.3 SFM based automatic targetless camera - Lidar calibration 7
Task 3. Scene segmentation and classification 8
3.1 3D CNN based semantic segmentation of large scale MLS point clouds 8
3.2 CNN-based Watershed Marker Extraction for Brick Segm. in Masonry Walls 9
Task 4. Object extraction and recognition 10
4.1 Instant Object Detection in 3D Point Clouds and Voxel Models 10
4.2 Automatic tumuli detection in LiDAR based digital elevation maps 11
4.3 HierarchyNet: Hierarchical CNN-Based Urban Building Classification 12
Task 5. Dynamic event recognition 13
5.1 Lidar-based Gait Analysis and Activity Recognition in a 4D Surveil. System 13
5.2 Multi-object Detection Utilizing 3D Background Maps and Tracking 14
Task 6. Scene understanding and object context modelling 14
6.1 An Embedded Marked Point Process for Multi-Level Obj. Population Analysis 14
6.2 Exploitation of Dense MLS City Maps for 3D Object Detection 15
Task 7. Change detection and change classification 16
7.1 Change Detection by a Real Time Lidar Scanner and MLS Reference data 16
7.2 A GAN-based Blind Inpainting Method for Masonry Wall Images 17
7.3 ChangeGAN: A deep network for change detection in unregistered point clouds 18
Publications of the project results 19
Books and theses 19
Journal papers 19
International Conference papers 20
National Conference papers 21
Technical reports 22
Exhibitions, technical demonstrations 22
Contributions of project participants 23
Scientific activities, awards, and events related to the project 25
Modified workplan of project (modifications accepted in October 2020) 27

Final Report of Project NKFIH K-16 #120233 "Instant environment perception from a mobile platform with a new generation geospatial database background"

Call: National Research, Development and Innovation Office, K-16 Researcher initiated projects
Duration: 01.10.2016-31.03.2021.
Principal Investigator (PI): Dr. Csaba Benedek

Summary of the project achievements

The project team has accomplished the targeted research work in the addressed area of 3D machine perception and large scale Geo-Information System (GIS) data processing. While the majority of tasks listed in the workplan of the proposal have been implemented during the originally planned 48 months time frame (2016-2020), due to the COVID Situation we requested a half year project extension until 31. March 2021. In addition, in a contract amendment request approved on 16. Oct. 2020, we reported some changes in the scopes of a few project tasks, which we justified due the appearance of new scientific and industrial trends (new sensors, new software solutions in the market), since the original proposal was prepared in early 2016. In particularly, in the last part of the project we have put more emphasis on sensor fusion algorithms, and on the generalization of our developed spatial and visual data processing methodologies to further application fields such as *archeology*, *urban area monitoring* and *medical image analysis*. A detailed summary of workplan modification is also included in this report (pages 27-28).

As outcome of the project *8 journal papers* have been published with a summarized impact factor (IF) 28.6, and *2 further journal papers* (with a summarized IF 9.4) are under review. In addition, 13 international conference papers (including a conference *keynote lecture* paper in the Springers CCIS series, 6 IEEE Xplore, 4 LNCS, 1 ISPRS Archives papers), 7 national conference papers and 3 technical reports have been published.

During the project the PI, Csaba Benedek has defended his DSc dissertation [B3], habilitation thesis [B1], and he prepared a book (to be published by Springer) covering various results of the project [B5*]. Two PhD dissertations have also been prepared in connections with the project.

In the following section of the report, we summarize the project's main contributions, connected to the 7 main workpackages (tasks) introduced in the workplan of the project. Thereafter, we provide a complete list of publications prepared with the support of the project, we summarize the personal contributions of the project participants, and also report further scientific activities, awards, and events related to the project.

Project results

Note: Tasks 1-7 have been defined in the original workplan of the project in 2016, with minor changes approved by the NRDI Office on 16. October 2020 (the modified workplan is available in the "Events with proposal, project" section of the NKFI-ERP system).

Task 1 Large scale GIS data management and visualization

By the evolution of 3D scanning techniques, creating 3D models of real world objects is getting much easier. Beyond the human-sized objects one can easily scan complete buildings, roads, squares, or even towns and countries as well. The raw data that scanning technologies, such that Light Detection and Ranging (Lidar) or photogrammetry based applications can provide are point clouds. The size of such a point cloud can be enormous, with billions of points, and processing and converting it to another for-mat is very costly. Due to this progress it is important to efficiently visualize large point clouds and to make it possible to modify them. Large-scale point cloud scene annotation is a crucial step in deep learning based Mobile Laser Scanning (MLS) data classification and object detection algorithms.

In the project, we have developed an "out-of-core" approach on rapid access of any requested segment in large point clouds, which cannot be read in full neither into the GPU memory, nor to the RAM of the computer. New data structures and algorithms have been elaborated for point cloud storage, reading, and manipulation, which enable quick data access, *without* downscaling the spatial resolution, and *with* keeping the genuine accuracy of the data. We have implemented *a software system for point cloud visualisation* that also provides tools to *manipulate point clouds* by selecting, annotating, deleting, cleaning necessary parts of them [R3] (see Fig. 1). Using our above described software engine, we have developed an advanced 3D point cloud annotation tool which can be used for efficient training and test data generation for machine learning based semantic segmentation and object detection methods. The user friendly 3D tool allows operators to label arbitrary shaped 3D volumes quickly, by assigning unique labels to occupied voxels of the scene, where the voxel size (in our case 10cm) determines the spatial accuracy of annotation. In one step, the operator can mark a rectangle area on the screen, which defines with the actual viewpoint a 3D pyramid volume in scene's 3D coordinate system. Then, the annotated volume can be created through a combination of union and intersection operations on several pyramids. With this method we manually labeled around 327M points over a 30.000 m^2 area of the city, with more than 50m elevation differences, using nine classes. As a result of annotation, we created a new publicly available benchmark set called SZTAKI CityMLS which was used in our point cloud segmentation research work (see later Task 3).

The proposed system was used as a key element in several publications of the project team [J3], [J5], [C3], [C9] and [C10], where the efficient data management was a crucial initial step in processing and semantic interpretation of large scale point clouds obtained by mobile laser scanning, for example in generating 3D city maps in autonomous driving applications. In addition, Zsolt Jankó and Csaba Benedek presented a film-making related application of the proposed system as exhibitors in the International Conference on Animation, Effects, VR, Games and Transmedia (FMX) [E1], where the technical and practical contributions of the research work were also highlighted. The details of the developed tool are described in a technical report [R3].

Fig 1. Demonstration of the large point cloud visualization and manipulation tool developed in the project. Left: annotated windows in the 1.5 billion sized point cloud of the Hungarian Parliament. Right: annotated pedestrians by Deák Square Budapest. MLS test data has been provided by Buda pest Közút Zrt.

Task 2 Point cloud registration from different sensors

2.1 Registration of sparse point clouds captured for real-time scene analysis to dense reference point cloud maps

We have proposed in [C1], [H1] a general approach for registration of point clouds obtained by different mobile laser scanning technologies. Our method is able to robustly match measurements with significantly different density characteristic including the sparse and inhomogeneous instant 3D (I3D) data taken be self-driving cars, and the dense and regular point clouds captured by mobile mapping systems (MMS) for virtual city generation (Fig. 2).

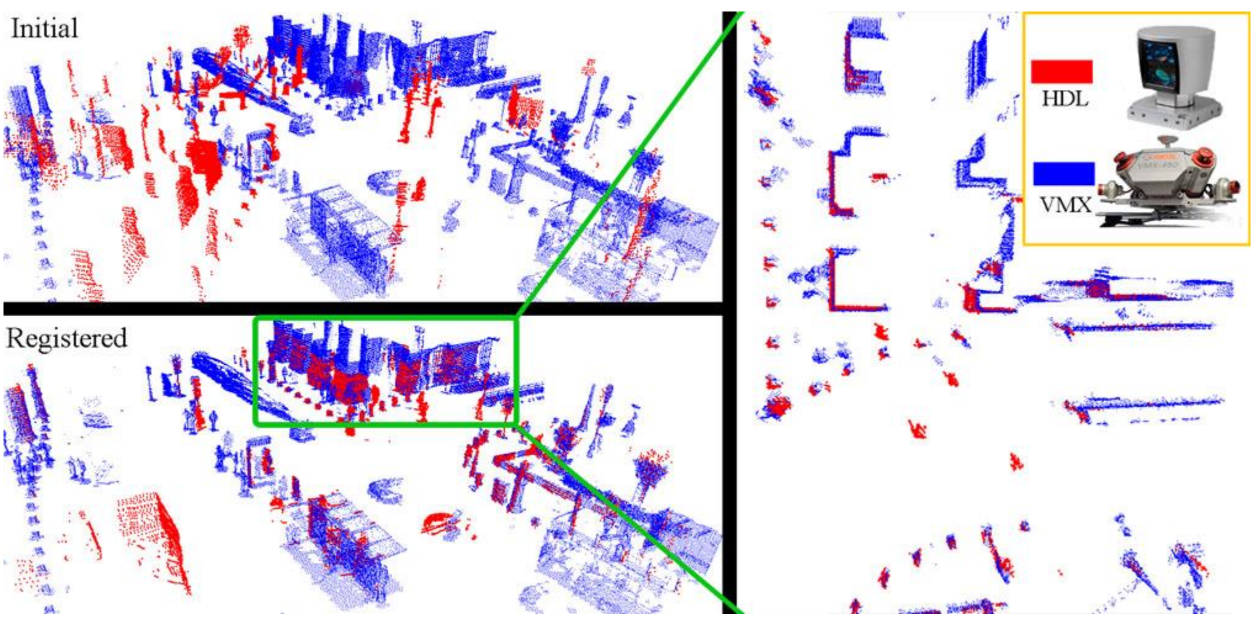

Fig 2. Result of the proposed point cloud registration technique applied for a low resolution real-time Lidar frame (red) and a high resolution MLS point cloud map, by Fővám Square, Budapest.

The core steps of the point cloud registration algorithm are robust scan segmentation, abstract street object extraction, object based coarse transformation estimation in the Hough accumulator space, and point-level registration refinement. Experimental results were provided using three different sensors: Velodyne HDL64 and VLP16 I3D scanners, and a Riegl VMX450 MMS. Application examples are shown regarding self-localization of autonomous cars through crossmodal I3D and MMS frame registration, IMU less SLAM and change detection based on I3D data.

In a next work phase, we have extended the above technique in [C5], obtaining an improved solution which is significantly quicker than the original model [C1], but it is still able to robustly register the sparse point clouds of the self driving vehicles to the High Definition maps based on dense MLS point cloud data, starting from a GPS based initial position estimation of the vehicle. The main steps of the method are robust object extraction and transformation estimation based on multiple keypoints extracted from the objects, and additional semantic information derived from the MLS based map. We tested our approach on roads with heavy traffic in the downtown of a large city with large GPS positioning errors, and showed that the proposed method enhances the matching accuracy with an order of magnitude. Comparative tests were provided with various keypoint selection strategies, and we showed the superiority of the new model against state-of-the-art techniques, including our earlier approach [C1].

2.2 Fusion of high resolution optical images with low resolution depth maps

In [J1], we presented a review of the approaches that couple Time-of-Flight (ToF) depth images with high-resolution optical images (Fig. 3). Recently, there has been remarkable growth of interest in the development and applications of ToF depth cameras. However, despite the permanent improvement of their characteristics, the practical applicability of ToF cameras is still limited by low resolution and quality of depth measurements. This observation has motivated many researchers to combine ToF cameras with other sensors in order to enhance and upsample depth images. Other classes of upsampling methods were also briefly discussed. Finally, we provided an overview of performance evaluation tests presented in the related studies.

Fig 3. Result of an upsampled depth map (center), to a high resolution optical image (left), with available ground truth depth information (right)

2.3 SFM based automatic targetless camera - Lidar calibration

We introduced a new approach for automatic targetless Lidar-camera calibration [J6, C6, C8]. Using our technique one can robustly project the 3D points obtained by a Lidar system onto the image domain (Fig 4 images on the right). The main steps of the algorithm are structure from motion (SFM) based point cloud generation, object level course alignment between the generated and the Lidar point cloud and registration refinement in the image domain (Fig 4).

The main challenge in this context is that state-of-the-art Rotating Multi-Beam (RMB) Lidar sensors used for autonomous driving (such as Velodyne HDL64, or VLP16) provide inhomogeneous, colorless, sparse point cloud streams with typical ring patterns, therefore it is difficult to extract specific structures for feature matching - such as corresponding corners, lines, or planar segments - between the 3D point clouds and the 2D image data. To avoid feature detection we turn to a structure from motion (SfM) based technique to generate point clouds from the image sequences recorded by the moving vehicle, and we perform an alignment between the Lidar and the generated point clouds. In this way, our main task can be interpreted as a point cloud registration problem, where we can adopt our earlier discussed multi model point cloud registration approach [C5]. In the present implementation, our technique extracts object blob centers in both point cloud frames, which are matched in the Hough domain, based on the idea of a fingerprint minutiae matching algorithm. Although that approach is able to find a robust transformation between two point sets even if the number of points are different, it becomes sensitive to several false or inaccurate hits of the object detector, which are present in our case since both the RMB Lidar and the SfM point clouds are quite sparse and noisy. In particularly, we observed that vehicles in the SfM clouds often fall into several pieces due to their homogeneous surfaces, causing false matches to the Hough- based estimator [C6]

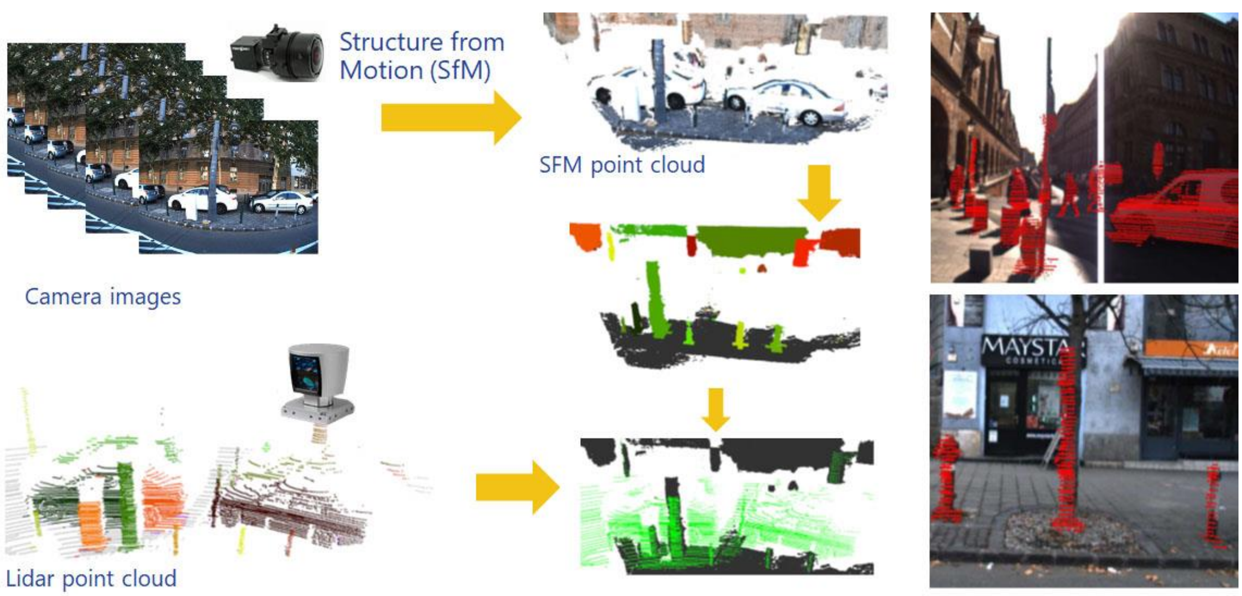

Fig 4. Workflow demonstration of the proposed fully automatic camera-Lidar calibration process. Right: demonstration of the results by projecting the object extracted from the Lidar point clouds to the camera

The next key step is to use *semantic information* for eliminating many of the false object candidates [C8]. While object segment classification in sparse point clouds is often unreliable due to occlusion, we can robustly detect vehicle instances in the original camera images with deep neural networks such as *Mask R-CNN*. Even in deficient SfM clouds, by projecting the 2D class labels into 3D the vehicle points can be efficiently identified) and removed helping registration enhancement.

As a further development of the above method, we applied a control point based nonrigid transformation refinement step to register the point clouds more precisely. Finally, we calculate the correspondences between the 3D Lidar points and the pixels in the 2D camera domain.

To evaluate our proposed target-less and fully automatic self-calibration method we compared it to two state-of-the-art target-less techniques, and an offline target-based method, and measured the superiority of our new approach. We quantitatively evaluated the considered methods by measuring the magnitude and standard deviation of the pixel level projection error values both in the x and y directions along the image axes. We collected test data from overall 10 km long road segments in different urban scenes such as boulevards, main roads, narrow streets and large crossroads. To take into account the daily traffic changes, we recorded measurements both in rush hours in heavy traffic, and outside the peak hours as well.

The results show that our proposed extrinsic calibration approach is able to provide accurate and robust parameter settings on-the-fly, and it surpasses the reference methods.

We have published the technique in an Open Access (Scimago D1) journal paper [J6] in 2020, while earlier stages of the approach have been presented in international conferences [C6, C8].
.

Task 3. Scene segmentation and classification

3.1 3D CNN based semantic segmentation of large scale MLS point clouds

In [C3] we introduced a new deep learning based approach to detect and remove phantom objects from point clouds produced by mobile laser scanning (MLS) systems. The phantoms were caused by the presence of scene objects moving concurrently with the MLS platform, and appeared as long, sparse but irregular point cloud segments in the measurements. We proposed a new 3D CNN framework working on a voxelized column-grid to identify the phantom regions. We quantitatively evaluated the proposed model on real MLS test data, and compared it to two different reference approaches.

In the next work phase [J5], we extended the model [C3], specifically developed for phantom detection and removal, for recognizing nine different semantic classes required for High Definition (HD) map generation: phantom, tram/bus, pedestrian, car, vegetation, column, street furniture, ground and façade (Fig. 5).

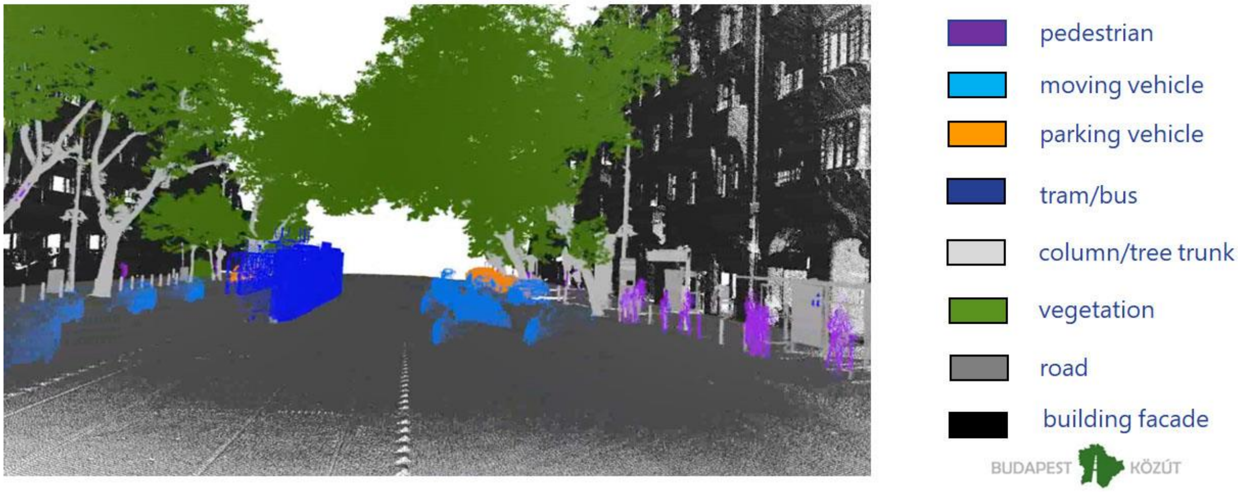

Fig 5. Result of semantic MLS point cloud segmentation by the proposed 3D CNN model. Test data has been provided by Budapest Közút Zrt.

As main methodological differences from [C3], we used in [J5] a two channel data input featuring local point density and elevation; and a voxel based space representation, which can handle the separation of tree crowns or other hanging structures from ground objects more efficiently than the earlier pillar based model.

To keep the computational requirements low, we implemented a sparse voxel structure avoiding unnecessary operations on empty space segments. We evaluated our proposed method against three reference techniques in qualitative and quantitative ways. First, we tested a single channel 3D CNN, which uses a 3D voxel occupancy grid (OG) as input (OG-CNN). Second, we implemented a multi-view method which projects the point cloud onto different planes, and achieves CNN classification in 2D. Third, we tested the PointNet++ deep learning framework, using their publicly available source code. We evaluated our proposed method against five reference techniques in qualitative and quantitative ways. The overall results of the reference techniques fall behind our proposed method, for example with margins of 14.8% (OGCNN), 17.4% (multi-view), and 3.8% (PointNet++), respectively [J5, H5, H6].

3.2 CNN-based Watershed Marker Extraction for Brick Segmentation in Masonry Walls

Nowadays there is an increasing need for using artificial intelligence techniques in image-based documentation and survey in archeology, architecture or civil engineering applications. Brick segmentation is an important initial step in the documentation and analysis of masonry wall images. However, due to the heterogeneous material, size, shape and arrangement of the bricks, it is highly challenging to develop a widely adoptable solution for the problem via conventional geometric and radiometry based approaches.

In our work published in [C7], we proposed a new technique which combines the strength of deep learning for brick seed localization, and the Watershed algorithm for accurate instance segmentation. More specifically, we adopted a U-Net-based delineation algorithm for robust marker generation in the Watershed process, which provides as output the accurate contours of the individual bricks, and also separates them from the mortar regions (Fig. 6).

Fig 6. Automated brick segmentation results for different (regular and irregular) masonry wall images, using the proposed CNN and Watershed based model

For training the network and evaluating our results, we created a new test dataset which consist of 162 hand-labeled images of various wall categories. Quantitative evaluation has been provided both at instance and at pixel level, and the results have been compared to two reference methods proposed for wall delineation, and to a morphology based brick segmentation approach. The experimental results showed the advantages of the proposed U-Net markered Watershed method, providing evaluation rates (recall, and precision, F-measure) between 70% to 97% in every test category.

Task 4. Object extraction and recognition

4.1 Instant Object Detection in 3D Point Clouds and Voxel Models

In [J2] we presented a new approach for object classification in continuously streamed Lidar point clouds collected from urban areas. The input of our framework are raw 3-D point cloud sequences captured by a Velodyne HDL-64 Lidar, and we aim to extract all vehicles and pedestrians in the neighborhood of the moving sensor. We propose a complete pipeline developed especially for distinguishing outdoor 3-D urban objects. Firstly, we segment the point cloud into regions of ground, short objects (i.e. low foreground) and tall objects (high foreground). Then using our novel two-layer grid structure, we perform efficient connected component analysis on the foreground regions, for producing distinct groups of points which represent different urban objects. Next, we create depth-images from the object candidates, and apply an appearance based preliminary classification by a Convolutional Neural Network (CNN). Finally we refined the classification with contextual features considering the possible expected scene topologies. We tested our algorithm in real Lidar measurements, containing 1159 objects from different scenarios (Fig 7.).

Apart from focusing on the analysis of urban laser scanning data, we also explored the usability of the developed algorithms for medical [H4, H7] applications: we developed methods for the segmentation of different organs (liver, spleen, kidney) from 3D Computed tomography (CT) and magnetic resonance (MR) volumes, with using in part similar approaches to our Lidar based object classifiers.

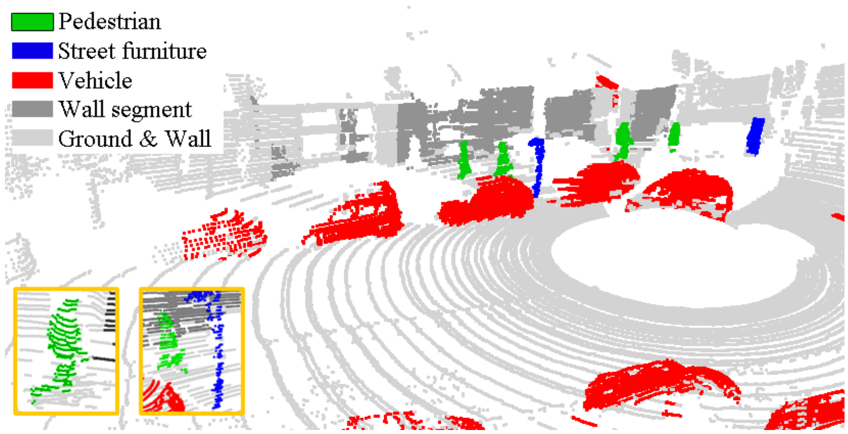

Fig 7. Result of object detection and classification in real-time Lidar frames by the proposed CNN model

On the other hand, we also adopted our earlier point cloud and image analysis methodologies for automatic extraction and segmentation of stonewalls in archeological terrestrial laser scanning measurements, soil radar measurements [H3].

4.2 Automatic tumuli detection in LiDAR based digital elevation maps

Extending our multilevel Marked Point Process (MPP) model presented in [J3], we proposed a new archeological application of the approach in [C11]. We developed an automatic MPP method which aims to identify Iron Age tumuli in digital terrain models extracted from airborne LiDAR measurements (Fig 8). Our task was particularly challenging due to the strong and uneven elevation of the ground where tumuli are located. In addition, the area is densely covered by vegetation, mainly by forests, which makes it difficult to create a surface model. The difficulty was exacerbated by the fact that the height and width of tumuli is very diverse but incomparably smaller than the extension of the area. In contrast to the various visualization techniques (openness, local relief model) used since decades for identifying tumuli, our method recognises these features automatically. We quantitatively evaluated the method on six test data sets and we demonstrated its advantages compared to a traditional Hough transform-based method.

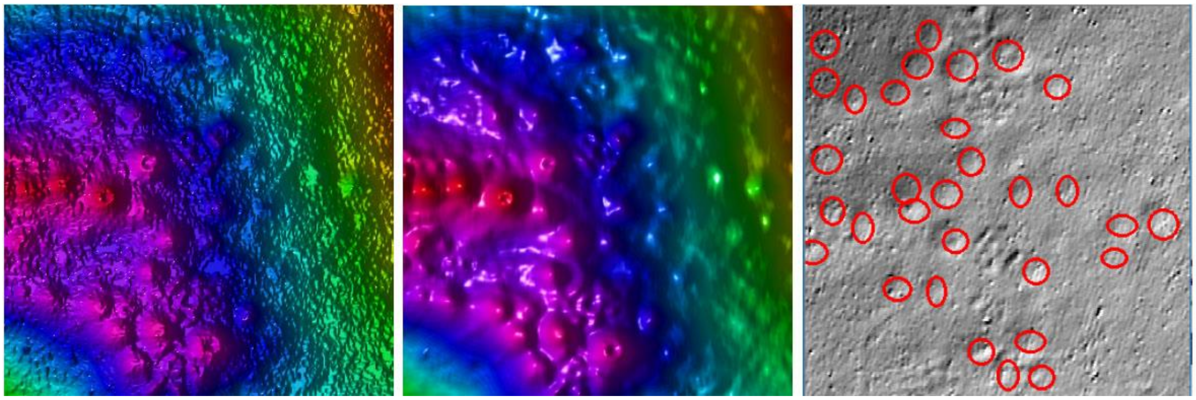

Fig 8. Results of Iron Age tumuli detection in Lidar based Digital Elevation Models (DEM). Left: input DEM, center: filtered DEM, right: detected tumuli

4.3 HierarchyNet: Hierarchical CNN-Based Urban Building Classification

Automatic building categorization and analysis are particularly relevant for smart city applications and cultural heritage programs (Fig. 9). Taking a picture of the facade of a building and instantly obtaining information about it can enable the automation of processes in urban planning, virtual city tours, and digital archiving of cultural artifacts. In [J7], we have gone beyond traditional convolutional neural networks (CNNs) for image classification and proposed the HierarchyNet: a new hierarchical network for the classification of urban buildings from all across the globe into different main and subcategories from images of their facades. We introduced a coarse-to-fine hierarchy on the dataset and the model learns to simultaneously extract features and classify across both levels of hierarchy. We proposed a new multiplicative layer, which is able to improve the accuracy of the finer prediction by considering the feedback signal of the coarse layers (Fig. 10). We have quantitatively evaluated the proposed approach both on our proposed building datasets, as well as on various benchmark databases to demonstrate that the model is able to efficiently learn hierarchical information. The HierarchyNet model is able to outperform the state-of-the-art convolutional neural networks in urban building classification as well as in other multi-label classification tasks while using significantly fewer parameters.

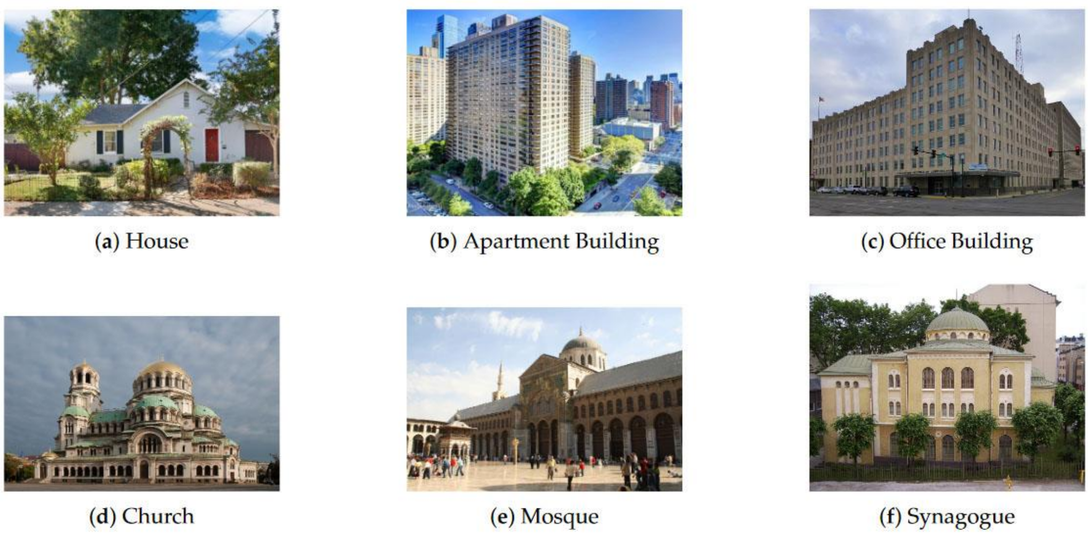

Fig 9. Examples of different building categories distinguished by the fine branch of the proposed *HierarchyNet* model

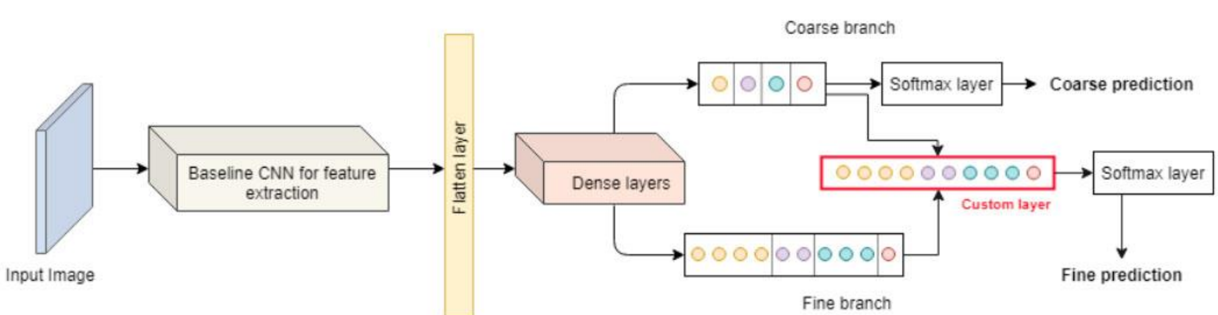

Fig 10. Structure of the proposed HierarchyNet hierarchical deep neural network

Task 5. Dynamic event recognition

5.1 Lidar-based Gait Analysis and Activity Recognition in a 4D Surveillance System

In [J4], [C2], [H2] we presented new methods exploiting LiDAR scanners with different resolutions, which can be used in video surveillance application. We showed in [J4] that gait recognition methods based on the point clouds of a Velodyne HDL-64E Rotating Multi-Beam LiDAR can be used for people re-identification in outdoor surveillance scenarios, and they can also be adopted for the recognition of specific activity patterns such as bending, waving, making phone calls and checking the time looking at wristwatches. The descriptors for training and recognition (Fig. 11, Fig. 12) were observed and extracted from realistic outdoor surveillance scenarios, where multiple pedestrians are walking in the field of interest following possibly intersecting trajectories, thus the observations might often be affected by occlusions or background noise. Since there is no public database available for such scenarios, we created and published a new Lidar-based outdoors gait and activity dataset on our website, that contains point cloud sequences of 28 different persons extracted and aggregated from 35 minutes-long measurements. The presented results confirmed that both efficient gait-based identification and activity recognition is achievable in the sparse point clouds of a single RMB Lidar sensor.

Since the initial high cost and the weight of the 64-beam scanner meant a bottleneck for its wide application in surveillance systems, we have shown in [C2] that the proposed Lidar-based Gait Energy Image descriptor can be efficiently adopted to the measurements of the compact and significantly cheaper Velodyne VLP-16 LiDAR scanner, which produces point clouds with a nearly four times lower vertical resolution than HDL-64. On the other hand, due to the sparsity of the data, the VLP-16 sensor proveed to be less efficient for the purpose of activity recognition, if the events were mainly characterized by fine hand movements. The evaluation was performed on five tests scenarios with multiple walking pedestrians, which have been recorded by both sensors in parallel.

The working method was also demonstrated in the Frankfurt Motorshow [E2], where we participated as exhibitor in cooperation with the sensor manufacturer company Velodyne.

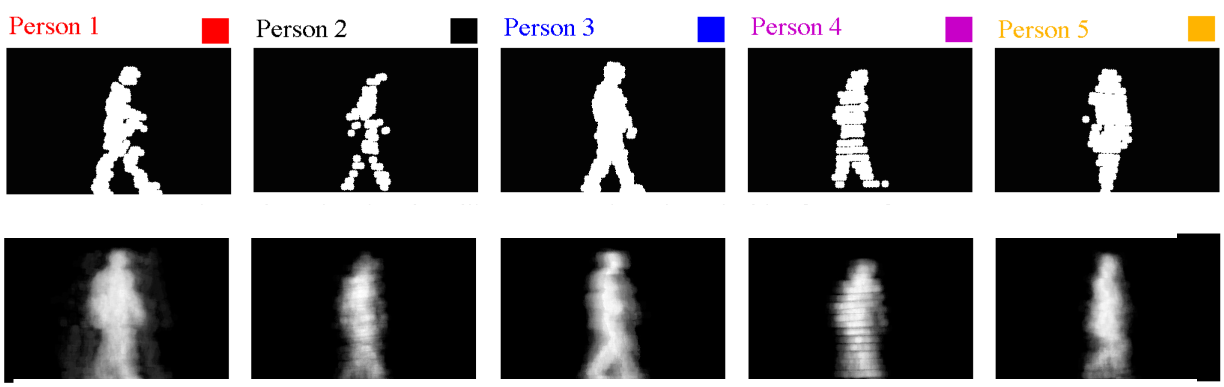

Fig 11. Biometric person re-identification from real-time Lidar sequences. Top: human silhouettes extracted from Lidar frames. Bottom: generated Gait Energy Images used for person recognition

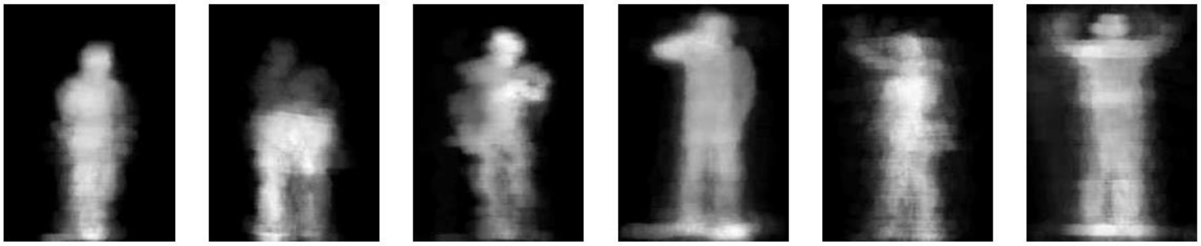

Fig 12. Activity recognition from Lidar sequences: walk, bend, checking watches, waving, and two-handed waving

5.2 Multi-object Detection in Urban Scenes Utilizing 3D Background Maps and Tracking

We have extended the dynamic object detector discussed in Task 2, with the utilization of temporal tracking information (published in [C9], further technical details provided in [R1], [R2]). To find additional objects missed by the object detector on the current measurement, we apply a Kalman-filter based object tracking. Hereby we first predicted the current state of the previously detected and tracked objects. Next, we applied a Hungarian matcher based assignment between the tracked and the current objects and update the object list according to the result. For better accuracy, we keep all predictions through a couple of frames.

We evaluated our method qualitatively and quantitatively in crowded urban scenes of Budapest, , and the results showed via tracking, a 12,84% improvement for vehicles and 14,34% for pedestrians in recall against the state-of-the-art object detection method relying purely on a single Lidar time frames.

Task 6. Scene understanding and object context modeling

6.1 An Embedded Marked Point Process Framework for Multi-Level Object Population Analysis

In [J3] we introduced a probabilistic approach for extracting complex hierarchical object structures from digital images used by various vision applications. The proposed framework extends conventional Marked Point Process (MPP) models by (i) admitting object-subobject ensembles in parent-child relationships and (ii) allowing corresponding objects to form coherent object groups, by a Bayesian segmentation of the population. Different from earlier, highly domain specific attempts on MPP generalization, the proposed model is defined at an abstract level, providing clear interfaces for applications in various domains. We also introduced a global optimization process for the multi-layer framework for finding optimal entity configurations, considering the observed data, prior knowledge, and interactions between the neighboring and the hierarchically related objects. The proposed method was demonstrated in three different application areas: built in area analysis in remotely sensed images (Fig. 13), traffic monitoring on airborne and mobile laser scanning (Lidar) data and optical circuit inspection. A new benchmark database was published for the three test cases, and the model's performance is quantitatively evaluated

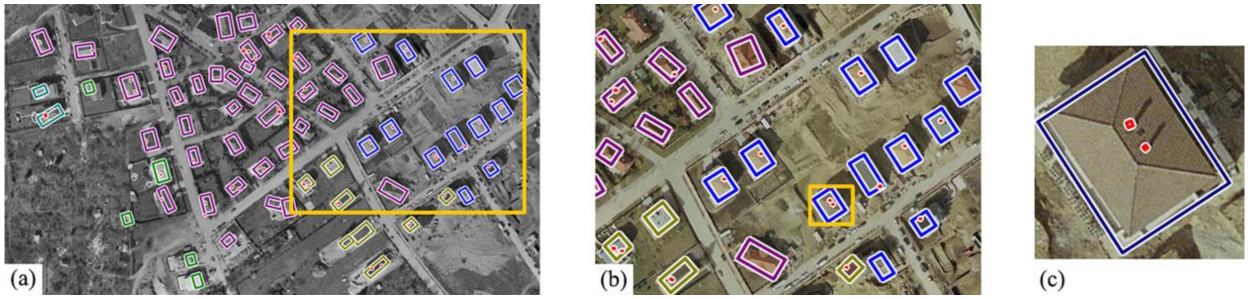

Fig 13. Built in area analysis application example of the proposed Embedded Marked Point Process (EMPP) framework. Corresponding (a) residential regions, (b) building rooftops and (c) chimneys are simultaneously extracted

6.2 Exploitation of Dense MLS City Maps for 3D Object Detection

We proposed a novel method for the exploitation of High Density Localization (HDL) maps obtained by Mobile Laser Scanning in order to increase the performance of *state-of-the-art* real time dynamic object detection (RTDOD) methods utilizing Rotating Multi-Beam (RMB) Lidar measurements (Fig. 14). First, we align the onboard measurements to the 3D HDL map with a multimodal point cloud registration algorithm operating in the Hough space (proposed earlier in Task 2 of the project). Next we apply a grid based probabilistic step to filter out the object regions on the RMB Lidar data which were *falsely* predicted as dynamic objects by RTDOD, although they are part of the static background scene.

For decreasing the false negative detections we use a change detection approach as introduced by Task 7 in details. Finally, to analyse the changed but previously unclassified segments of the RMB Lidar clouds, we apply a geometric blob separation and a Support Vector Machine based classification to distinguish the different object types. Comparative tests are provided in high traffic road sections of Budapest, Hungary, and we show an improvement of 5,96% in precision, 9,21% in recall and 7,93% in *F*-score metrics against the *state-of-the-art* RTDOD algorithm. Results have been published in [J10*], [C10].

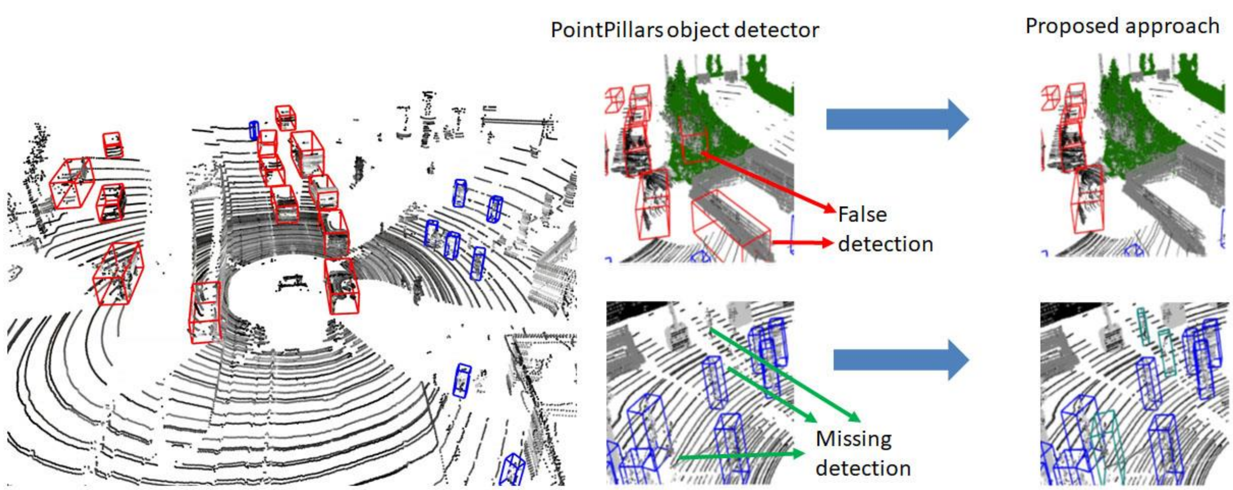

Fig 14. Lidar object detection enhancement by the proposed MLS map-based approach

Task 7. Change detection and change classification

7.1 Change Detection in Urban Streets by a Real Time Lidar Scanner and MLS Reference data

In [C4], we introduced a new technique for change detection in urban environment based on the comparison of 3D point clouds with significantly different density characteristics. Our proposed approach extracts moving objects and environmental changes from sparse and inhomogeneous instant 3D (i3D) measurements, using as reference background model dense and regular point clouds captured by mobile laser scanning (MLS) systems. The introduced workflow consists of consecutive steps of point cloud classification, crossmodal measurement registration, Markov Random Field (MRF) based change extraction in the range image domain and label back projection to 3D (Fig. 15). Experimental evaluation is conducted in four different urban scenes, and the advantage of the proposed change detection step is demonstrated against a reference voxel based approach (Fig. 16). [C13] summarizes the complete workflow of multi-sensorial environment perception proposed at various steps of the project.

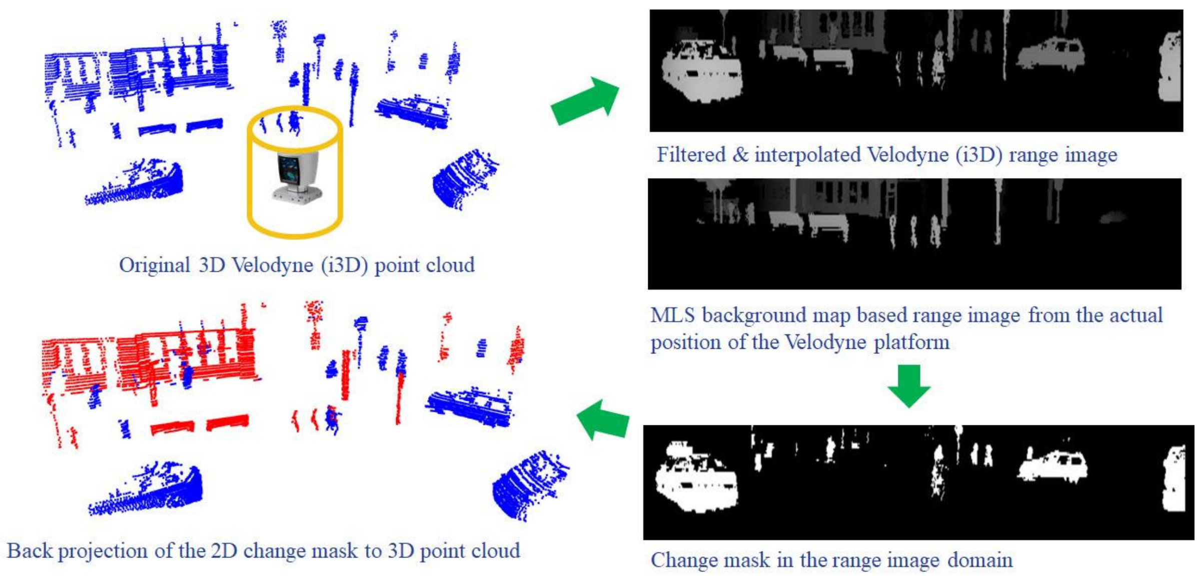

Fig 15. Demonstration of the change detection workflow between a real time Lidar frame, and an MLS reference map

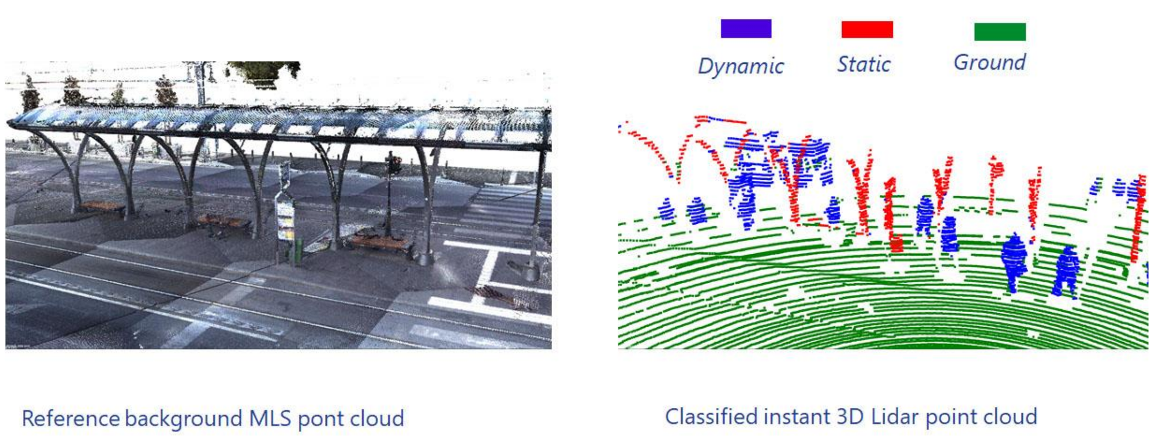

Fig 16. MLS map based change detection results in a tram station by Kálvin Square, Budapest

We have adopted our Lidar based change detection algorithm [C4], to improve the accuracy of the scene understanding module of Task 6. To find objects erroneously *missed* by the RTDOD predictions, we exploited the MRF based point level change detection approach between the map and the current onboard measurement frame (published in [J10*], [C10]).

7.2 A GAN-based Blind Inpainting Method for Masonry Wall Images

For detecting and virtually repairing damaged parts of masonry walls for archeological architectural applications we introduced a novel end-to-end blind inpainting algorithm (published in [J8], [C12]). For this purpose, we proposed a three-stage deep neural network that comprises a U-Net-based sub-network for wall segmentation into brick, mortar and occluded regions, which is followed by a two-stage adversarial inpainting model (Fig. 17).

The first adversarial network predicts the schematic mortar-brick pattern of the occluded areas based on the observed wall structure, providing in itself valuable structural information for archeological and architectural applications.
Finally, the second adversarial network predicts the RGB pixel values yielding a realistic visual experience for the observer. While the three stages implement a sequential pipeline, they interact through dependencies of their loss functions admitting the consideration of hidden feature dependencies between the different network components. For training and testing the network a new dataset has been created, and an extensive qualitative and quantitative evaluation versus the state-of-the-art is given.

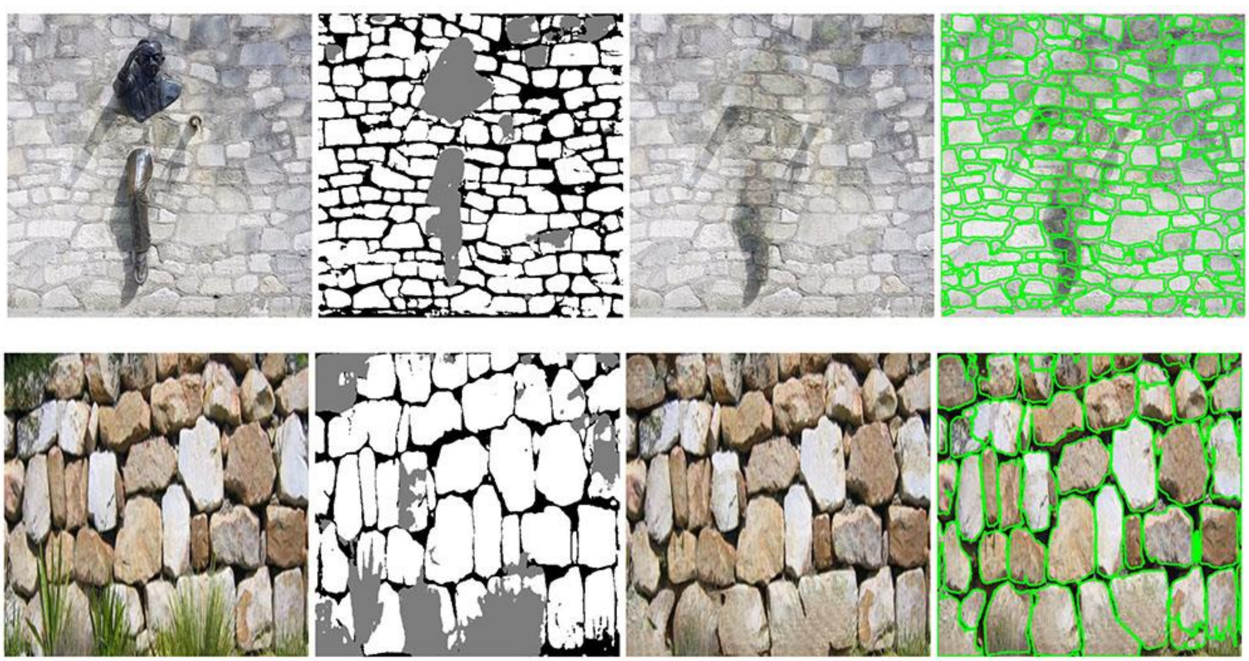

Fig 17. Occlusion removal and inpainting in real masonry wall images by the proposed end-to-end, fully automated GAN-based approach

7.3 ChangeGAN: A deep network for change detection in unregistered point clouds

We introduced a novel change detection approach called ChangeGAN *for unregistered point clouds* in complex street-level urban environment in [J9*]. Our a generative adversarial network (GAN) architecture compounds Siamese-style feature extraction, U-Net-like use of multiscale features, and Spatial Trans-formation Network (STN) blocks for optimal transformation estimation. The input point clouds are represented by range images, which enables using of 2D convolutional neural networks. The result is a pair of binary masks showing the change regions on each input range image, which can be backprojected to the input point clouds without loss of information (Fig. 18). We have evaluated the proposed method on various challenging scenarios and we have shown its superiority against state-of-the-art change detection methods.

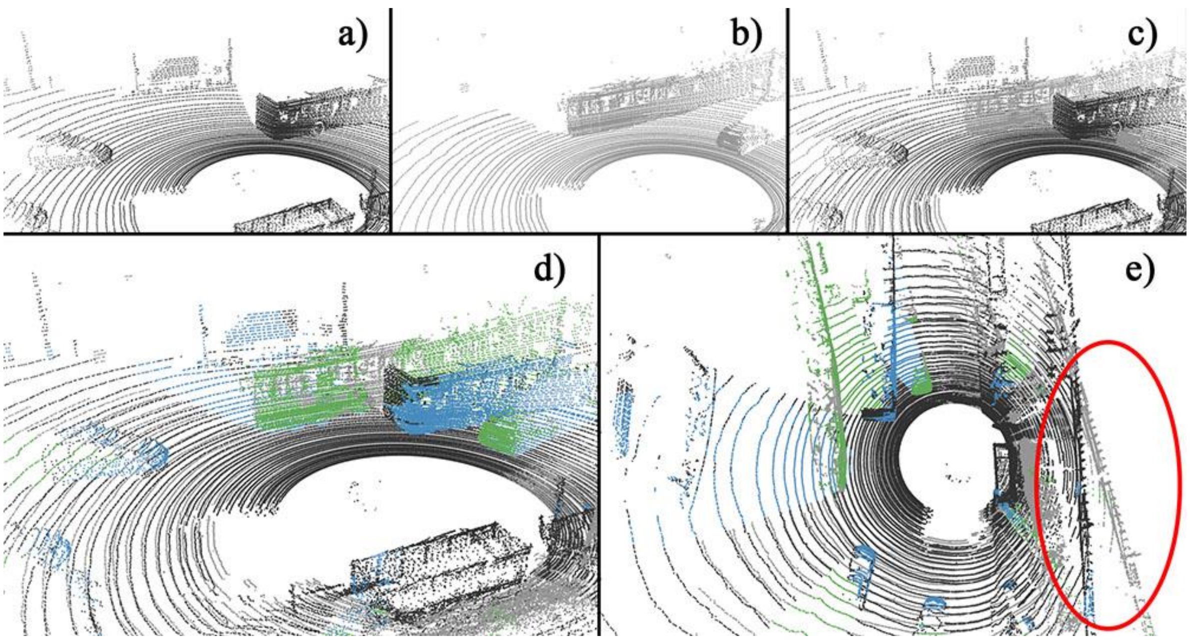

Fig 18. Detection changes between unregistered sparse point clouds with the proposed ChangeGAN approach (green and blue color mark changed objects in the first and second frames, respectively, red ellipse highlight the global registration error)

Publications of the project results

Notes: * marks publications under review, authors participated in the project are underlined

Book and theses

[B5*] Cs. Benedek, Multi-level Bayesian Models for Environment Perception, Research Monograph, Springer Nature (under review)

[B4] B. Nagy, Three-dimensional Scene Understanding in Mobile Laser Scanning data, *PhD. dissertation*, Pázmány Péter Catholic University, 2020 (supervisor: C. Benedek)

[B3] Cs. Benedek, Képi alapú többszintű környezetelemzés (Image based multi-level environment analysis), *MTA Doktori Disszertáció (DSc dissertation)*, 2020

[B2] A. Börcs, Four-dimensional Analysis of Dynamic UrbanEnvironments in Terrestrial and AirborneLiDAR Data, *PhD. dissertation*, Budapest University of Technology and Economics, 2018 (supervisor: C. Benedek)

[B1] Cs. Benedek: Probabilistic Approaches on Environment Perception with 2D-3D Imaging Sensors, *Habilitation thesis*, Pázmány Péter Catholic University, Budapest, Hungary, 2017

Journal papers

[J10*] Ö. Zováthi, B. Nagy and Cs. Benedek, „Change Detection in Urban Environment Using an Onboard Lidar Sensor and MLS Reference Data", IEEE Trans Geoscience and Remote Sensing, (under review), 2021

[J9*] B. Nagy, L. Kovács and Cs. Benedek „ChangeGAN: A deep network for change detection in unregistered point clouds", IEEE Robotics and Automation Letters, (under review), 2021

[J8] Y. Ibrahim, B. Nagy and Cs. Benedek: "Deep Learning-based Masonry Wall Image analysis," Remote Sensing, vol. 12, no. 23, article 3918, 2020, IF: 4.509*, Open Access, https://www.mdpi.com/2072-4292/12/23/3918

[J7] S. Taoufiq, B. Nagy and Cs. Benedek: "HierarchyNet: Hierarchical CNN-based Urban Building Classification," Remote Sensing, vol. 12, no. 22, article 3794, 2020, IF: 4.509*, Open Access, https://www.mdpi.com/2072-4292/12/22/3794

[J6] B. Nagy and Cs. Benedek: "On-the-Fly Camera and Lidar Calibration," Remote Sensing, vol 12, no. 7, article 1137, 2020, IF: 4.509*, Open Access, https://www.mdpi.com/2072-4292/12/7/1137

[J5] B. Nagy and Cs. Benedek: "3D CNN Based Semantic Labeling Approach for Mobile Laser Scanning Data," *IEEE Sensors Journal*, vol. 19, no. 21, pp 10034 – 10045, 2019, IF: 3.073 Open Access, https://ieeexplore.ieee.org/document/8756228

[J4] Cs. Benedek, B. Gálai, B. Nagy and Z. Jankó: ”Lidar-based Gait Analysis and Activity Recognition in a 4D Surveillance System,” IEEE Trans. on Circuits and Systems for Video Technology, vol. 28, no. 1, pp. 101-113, 2018, IF: 4.046

[J3] Cs. Benedek: "An Embedded Marked Point Process Framework for Three-Level Object Population Analysis", IEEE Trans. on Image Processing, vol. 26, no. 9, pp. 4430-4445, 2017, IF: 4.828 http://real.mtak.hu/id/eprint/63625

[J2] A. Börcs, B. Nagy and Cs. Benedek: "Instant Object Detection in Lidar Point Clouds", IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 7, pp. 992 - 996, 2017, IF: 2.761 http://real.mtak.hu/id/eprint/63631

[J1] Eichhardt, I., Chetverikov, D. and Z. Jankó, "Image-guided ToF depth upsampling: a survey", *Machine Vision and Applications*, vol 28, no. 3-4, pp. 267–282, 2017, https://eprints.sztaki.hu/9116/

International Conference papers

[C13] Cs. Benedek „Multi-sensorial Environment Perception in Urban Environment”, invited paper of a **keynote lecture,** International Conference on Robotics, Computer Vision and Intelligent Systems (ROBOVIS), Nov. 2020, to appear in the Springer Book Series *Communications in Computer and Information Science, 2021*

[C12] Y. Ibrahim, B. Nagy and Cs. Benedek: "A GAN-based Blind Inpainting Method for Masonry Wall Images", International Conference on Pattern Recognition (ICPR), 2021

[C11] Zs. Németh and Cs. Benedek: "Automatic tumuli detection in LiDAR based digital elevation maps," Internat’l Arc. Photogrammetry, Rem. Sens. and Spatial Inf. Sci, XXIV ISPRS Congress, volume XLIII-B2-2020, pp. 879–884,

[C10] Ö. Zováthi, B. Nagy and Cs. Benedek: "Exploitation of Dense MLS City Maps for 3D Object Detection", International Conference on Image Analysis and Recognition, LNCS, pp. 393-403, 2020

[C9] Ö. Zováthi, L. Kovács, B. Nagy and Cs. Benedek: "Multi-object Detection in Urban Scenes Utilizing 3D Background Maps and Tracking", International Conference on Control, Artificial Intelligence, Robotics and Optimization, IEEE, 2019

[C8] B. Nagy, L.A. Kovács and Cs. Benedek: "SFM and Semantic Information Based Online Targetless Camera-Lidar Self-Calibration", *IEEE International Conference on Image Processing (ICIP)*, Taipei, Taiwan, 22-25 September, 2019

[C7] Y. Ibrahim, B. Nagy and Cs. Benedek: "CNN-based Watershed Marker Extraction for Brick Segmentation in Masonry Walls", *International Conference on Image Analysis and Recognition* (ICIAR), Waterloo, Canada, August 27-29, 2019, Lecture Notes in Computer Science, Springer,

[C6] B. Nagy, L.A. Kovács and Cs. Benedek: "Online Targetless End-to-End Camera-LIDAR Self-calibration", *International Conference on Machine Vision Applications (MVA)*, Tokyo, Japan, 27-31 May, 2019

[C5] B. Nagy, and Cs. Benedek: "Real-time point cloud alignment for vehicle localization in a high resolution 3D map", *Workshop on Computer Vision for Road Scene Understanding and Autonomous Driving at ECCV*, Munich, Germany, Sept 14, 2018, to appear in Lecture Notes in Computer Science, Springer, 2018, http://real.mtak.hu/id/eprint/85680

[C4] B. Gálai, and Cs. Benedek: "Change Detection in Urban Streets by a Real Time Lidar Scanner and MLS Reference data", International Conference on Image Analysis and Recognition (ICIAR), Montreal, Canada, July 5-7, 2017, vol. 10317 of Lecture Notes in Computer Science, pp. 210-220, Springer, 2017 http://real.mtak.hu/id/eprint/63633

[C3] B. Nagy, and Cs. Benedek: "3D CNN Based Phantom Object Removing from Mobile Laser Scanning Data", International Joint Conference on Neural Networks (IJCNN), pp. 4429-4435, Anchorage, Alaska, USA, 14-19 May, 2017 http://real.mtak.hu/id/eprint/63635

[C2] B. Gálai, and Cs. Benedek: "Gait Recognition with Compact Lidar Sensors", International Conference on Computer Vision Theory and Applications (VISAPP), pp. 426-432, Porto, Portugal, 27 February - 1 March, 2017 http://real.mtak.hu/id/eprint/63647

[C1] B. Gálai, B. Nagy and C. Benedek: ”Crossmodal Point Cloud Registration in the Hough Space for Mobile Laser Scanning Data", *International Conference on Pattern Recognition (ICPR)*, Cancun, Mexico, 2016

National Conference papers

[H7*] O. Benis, Cs. Benedek “Hepatic vessel extraction using a 3D marked point process model and region growing”, Képfeldolgozók és Alakfelismerők Társaságának 13. konferenciája, 2021

[H6] Ö. Zováthi, B. Nagy and Cs. Benedek: Valós idejű pontfelhőillesztés és járműlokalizáció nagy felbontású 3D térképen, Képfeldolgozók és Alakfelismerők Társaságának 12. konferenciája, 2019

[H5] B. Nagy and Cs. Benedek: 3D CNN alapú MLS pontfelhőszegmentáció, Képfeldolgozók és Alakfelismerők Társaságának 12. konferenciája, 2019

[H4] Cs. Benedek, V. Czipczer, L. Kovács, A. Kriston, A. Manno-Kovács and T. Szirányi, “Segmentation of multiple organs in Computed Tomography and Magnetic Resonance Imaging measurements”, IV. Nemzetközi Interdiszciplináris 3D Konferencia, Pécs, 2018. október 4-5.

[H3] B. Nagy, Y. Ibrahim, P. Vámos, G. Bertók and Cs. Benedek, Detection of archaeological objects from radar and Lidar measurements, Workshop on Intelligent Data Analysis, Esztergom, June 2018, abstract and lecture

[H2] B. Gálai and Cs. Benedek: "Járás alapú személyazonosítás és cselekvésfelismerés LiDAR szenzorokkal" NJSZT Képfeldolgozók és Alakfelismerők Társaságának Konferenciája (Képaf) , Szováta, Románia, 2017 január http://real.mtak.hu/id/eprint/63650

[H1] B. Nagy, B. Gálai and Cs. Benedek: "Multimodális pontfelhőregisztráció Hough tér alapú előillesztéssel" NJSZT Képfeldolgozók és Alakfelismerők Társaságának Konferenciája (Képaf) , Szováta, Románia, 2017 január http://real.mtak.hu/id/eprint/63652

Technical reports

[R3] Zs. Jankó, A tool for manipulating huge point clouds, SZTAKI Technical Report, 2020 http://real.mtak.hu/116129/

[R2] Gy. Horváth, Cs. Benedek, Szenzorfúziós és objektumkövető eljárások Lidar méréssorozatok felhasználásával, SZTAKI Technical Report, 2020, in Hungarian, https://eprints.sztaki.hu/9956/

[R1] Gy. Horváth and Cs. Benedek Környezetelemzés kompakt lézerszkennerekkel SZTAKI Technical Report, 2018, in Hungarian, https://eprints.sztaki.hu/9955/

Exhibitions, technical demonstrations

[E4] Cs. Benedek and B. Nagy, "Lidar based environment perception and mapping for autonomous vehicles", AutoSens, Brussels, Belgium, 18-20 September, 2018, Poster presentation

[E3] Cs. Benedek, B. Nagy and Z. Jankó, "Instant environment perception from a mobile platform with a new generation geospatial database background", EU-US Frontiers of Engineering Symposium, Davis, CA, 2017, poster

[E2] Cs. Benedek, B. Nagy, Multiple pedestrian tracking demonstration in the Velodyne exhibitor area, Frankfurt Motor Show (iaa.de), Frankfurt am Main, Germany, 2017

[E1] Zs. Jankó, Cs. Benedek, Integrated 4D Film Preproduction System, International Conference on Animation, Effects, VR, Games and Transmedia (FMX), Exhibitor, Stuttgart, Germany, 2017

Contributions of project participants

This section briefly summarizes the personal contributions of the project participants.

Dr. Csaba Benedek, PhD, DSc, dr. habil, principal investigator of the project, senior researcher: he coordinated the project and participated in the scientific work of all work packages. He has been the sole author of a book and two dissertations, and a co-author of 30 publications or research reports related to the project. He supervised several MSc and PhD students working on the project.

Dr. Zsolt Jankó, PhD, senior researcher and developer: he participated in both research and development work during the project. He coordinated several development tasks related to large point cloud management visualization, ground truth generation, scene reconstruction, 3D/4D model generation. He has been co-author of two journal papers supported by the project, and the sole author of a detailed research report covering various scientific and technical issues MLS data processing. He also contributed to the development to the hardware-software systems demonstrating the project developments, and participated at various international and domestic exhibitions.

Attila Börcs, PhD candidate (obtained PhD degree in 2018) working as a PhD student of the PI, he contributed to Lidar based instant object detection task until his PhD defense (without financial support of the project).

Balázs Nagy, PhD student (obtained PhD degree in Dec. 2020): he has been a PhD student of the PI during almost the whole project period, in a PhD topic strongly related to project (Mobile 3D environment perception with a geospatial database background). Besides preparing his PhD thesis, he has been co-author in 8 journal articles, and in 12 conference publications, and he significantly contributed in demonstrator development and also in co-supervision of young BSc and MSc student colleagues. He won several SZTAKI Institute awards during the project period.

Örkény H. Zováthi, MSc/PhD student: he joined the project in late 2019 as a MSc student of the PI, and started PhD studies in 2020 (also supervised by the PI) in a project related research topic. In his early researcher period he has been the first author of a journal paper (currently under review), and 2 conference papers. He also participated in the National Student Conference (OTDK) 2021, where he received a third prize.

Bence Gálai, MSc student: he worked in the early phase of the project (2016-June 2017) as a student researcher, and contributed to the registration (Task 2.1) and to the Lidar based people surveillance (Task 5.1) by algorithm research and development activities. He has been co-author of a journal article and 5 conference papers. He won first prize at the National Student Conference (OTDK) 2017, and also received Denes Gabor special award.

Olivér Benis, MSc student: he worked in the late phase of the project (2019-2021) as a student researcher. Studying in info-bionics at the Péter Pázmány Catholic University he mainly focused on the possible medical image processing applications of the point cloud registration and object recognition methods developed in the project for various sorts of spatial data. He submitted his first conference paper in 2021.

Dávid Leichner, MSc student: he worked as a scientific programmer in the project in the early phase (2017-2018), and contributed to several development tasks related to large point cloud management, segmentation and visualization

Győző Horváth, MSc student: he worked as a scientific programmer in the project in the later phase (2019-2020), and contributed to several various development tasks related to real time point cloud processing, object detection and tracking and events analysis. He has been forst author of technical reports in the above topics.

Scientific activities, awards, and events related to the project and project participants

Thesis Defenses during the project

- Csaba Benedek, Habilitation, Pázmány Péter Catholic University, 2017
- Attila Börcs, PhD, Budapest University of Technology and Economics (supervisor: Csaba Benedek), 2018
- Csaba Benedek, Doctor of the Hungarian Academy of Sciences, 2020
- Balázs Nagy, PhD, Pázmány Péter Catholic University Economics (supervisor: Csaba Benedek), 2020

Scientific awards of the PI (Csaba Benedek) during the project

1. Michelberger Master Prize from the Hungarian Academy of Engineering, 2020
    - Based on competition, awarded one candidate and his/her mentor who obtained outstanding result in engineering (prize won together with Prof. Tamás Szirányi)
2. Bolyai plaquette from the Board of Trustees HAS Bolyai Research Fellowship, 2019
3. Imreh Csanád Plaquette from the Hungarian National Scientific Student Conference (OTDK)
    - Awarded bi-annually to one supervisor in information technology
4. Hungarian Academy of Sciences (HAS) Secretary-General's acknowledgement, 2018

Participation in scientific juries, reviewer boards of grants, scholarships by the PI

1. National Research, Development and Innovation Office (NKFIH) board of thematic research projects based on Hungarian–Austrian or Hungarian–Slovenian international cooperation (ANN_20, SNN_20), 2020 - member of evaluation committee
2. NKFIH postdoctoral excellence programme committee (two times) 2017, 2018 - committee and jury member, reviewer
3. John von Neumann Society Képaf Attila Kuba Prize committee member 2017, 2019

Membership of editorial boards of scientific journals

1. Elsevier Digital Image Processing, editorial board member (2019-)
2. MDPI Remote Sensing, guest editor of special issue (2019-2021)

International and domestic organizational memberships, positions:

1. Csaba Benedek became the president of the Hungarian Association for Image Processing and Pattern Recognition (Képaf), member society of the John von Neumann Computer Society (NJSZT), in 2019
2. Csaba Benedek became the Hungarian Governing Board member of the International Association for Pattern Recognition (IAPR)

3. Csaba Benedek continued to work as the *Hungarian National Representative in the* European Association for Biometrics (EAB) and in the International Cartographic Association (ICA) Commission on Sensor-Driven Mapping
4. Csaba Benedek became a *Senior Member of IEEE in* 2018

Featured presentations, invited lectures

1. Csaba Benedek was a **Keynote speaker at the ROBOVIS 2020 international conference** (*International Conference on Robotics, Computer Vision and Intelligent Systems),* title of talk: "Multi-sensorial Environment Perception in Urban Environment", November 4-6, 2020. Video of the talk: https://player.vimeo.com/video/480457538?title=0&portrait=0
2. *Csaba Benedek was the* invited presenter at the ***Jubilee General Assembly of the Hungarian Academy of Engineering*** in connection with receiving the Michelberger Master prize (title of talk: „Procedures for comprehensive perception and analysis of our dynamic environment" – in Hungarian), Hotel Gellért, February 6, 2020
3. Csaba Benedek was **invited presenter** at the *Hungarian Academy of Sciences* **Bolyai Day 2019**, in connection with receiving the Bolyai plaquette (title of talk: „ Environment perception with laser scanners"), Ceremonial Hall of the Hungarian Academy of Sciences, June 2019
4. *Csaba Benedek* was an invited presenter at a scientific meeting focused on medical application of computer science, in connection to the **191th General Assembly of the Hungarian Academy of Sciences** (Title of talk: „ Innovative medical solutions in an augmented reality based visualization technology of 3D medical scans and reality", in Hungarian), Hungarian Academy of Sciences, May 2019
5. *Csaba Benedek* was an invited presenter at the scientific meeting "*People-focused science"* connected to the **Hungarian Scientific Festival (MTÜ)** (title of talk: „Walking in laser-scanned cities", in Hungarian), Péter Pázmány Catholic Univeristy, November 2017
6. *Csaba Benedek* was at the **EU-US Frontiers of Engineering Symposium,** participant and presenter delegated by the Hungarian Academy of Engineering, Davis CA, USA, 2017

Awards, selected scholarships of student participants of the project

1. Örkény H. Zováthi won Cooperative Doctoral Program, Doctoral Student Scholarship in 2020
2. Balázs Nagy won ÚNKP Research Scholarship for Doctoral Candidates (NKFIH) in 2020
3. Örkény H. Zováthi won the SZTAKI Student Prize in 2020
4. Örkény H. Zováthi won the BME Pro Progressio Foiundation VIK special prize (for excellent student conference thesis) in 2019
5. Bence Gálai won the Dénes Gabor scholarship for distinguished 1st prize at the national Scientific Student Conference, 2017
6. Balázs Nagy won the SZTAKI Institute Prize for Best Ph.D. Student three times: in 2017, 2018, 2019, and the SZTAKI Youth Prize in 2017

Modified Workplan of project 120233

This section summarizes the differences between of the final project workplan, and the initial workplan from the original project proposal. Titles of Tasks 1-7 and their main goals are the same as in the original workplan, only minor modifications of the scope have been implemented, as detailed in the following. All changes have been accepted by the NRDI office on 16. October 2020. The timing of the tasks are given in half years H1-H8 for the **originally** 4 years project**, which has been completed with the H9 half year that belongs to the fifth work section**.

Tasks affected by the contract modification compared to the above work plan

The project follows the goals of the original workplan, whose fulfilment is the main goal of the fitfth work section starting from 1. Oct. 2020. The majority of tasks listed in the workplan of the proposal have been implemented during the originally planned 48 months time frame (2016-2020). Because of the appearance of new scientific and industrial trends (new sensors, new software solutions in the market), since the original proposal was prepared in early 2016, the original workplan needed some modification and re-scheduling, which will be detailed in the following.

Task 2 Point cloud registration from different sensors (H2-H3-H9)

Beyond implementing the originally planned research tasks, we have also obtained significant new results in spatial upscaling depth sensor measurements based on optical images, and fully automatic calibration of camera and Lidar sensors. For the Velodyne-MLS registration task also mentioned in the project proposal various results have been obtained during the project, including in the last H9 work section.

Task 3. Scene segmentation and classification (H3-H4-H9)

Beyond implementing the originally planned research tasks, we have proposed a new method for classification of building photos, in a hierarchical manner based on function and style (H9).

Task 4. Object extraction and recognition (H4-H6-H9)

The methods mentioned in the original project plan have been implemented, but the application are has been modified a bit due to availability issues of various sorts of test data.
Instead of detailed analysis of point clouds scenes of city streets, we could implement new object recognition and analysis methods on archaeological spatial (GIS) data samples, partially in H9.

Task 5. Dynamic event recognition (H6-H8)

The planned *person detection, tracking, biometric identification* and *activity recognition* methods have been implemented and published. Our originally planned solutions for the *traffic monitoring* and *event recognition* subtasks have lost their actuality during the project period, since several industrial companies proposed in similar area efficient solutions based on complex computer graphics simulation and machine learning. With these products, we could not compete due to the necessity of strongly multi-disciplinary staff, and huge amount of human resources (e.g. for system development and data annotation).

Task 6. Scene understanding and object context modeling (H6-H8)

The methods detailed in the original project plan have been implemented, while the application focus has been changed a bit: apart from the analysis of built-in areas and city surveillance, we have also demonstrated the efficiency of the proposed methodology for printed circuit board analysis.

Task 7. Change detection and change classification (H7-H8-H9)

The implementation of the task followed the original project plan with minor modifications: instead of Markovian techniques, we mainly approached the problem by artificial intelligence methods, and instead of CAD models we used as reference models dense point clouds obtained by mobile and terrestrial laser scanning.