Final defence review report of PhD scientific work with title

„Indoor Navigation for People with Visual Impairment"

of the candidate **Mostafa Abdallah Abbas Atwa Elgendy**

Actual report is in conformity with the internal rules and regulations with No. 16/2017 of the University of Pannonia Senate carried into effect on 23 February 2017. Majority of the suggestions indicated at the home defence has been applied in the dissertation and the thesis booklet, enhancing the scientific level of these documents.

1. General impression

The subject and title of the dissertation are relevant and correspond to the development trends of assisted technologies for helping people with visual impairment (PVI) in their everyday life. Research motivation is based on the obvious requirement of implementing user-friendly end efficient mobile assistive technology in indoor environment to the over 280 million people affected worldwide. Main target of the research work is finding the best usable technology for PVI and build navigation system to avoid physical obstacles improved with convolutional neural network solutions. Results of the research work are usable in practice to enhance mobile assisted technologies and services for PVI.

The overall format of both documents makes easy readable each of them. The list of references includes relevant studies related to the topic of the research work. The quantity and quality of them are sufficient and satisfactory.

2. Dissertation

The 93 pages document is well structured, contains majority of all necessary elements in well-organized order. Six scientific chapters and the final conclusions chapter give clear view of the executed research and development work. The size of each chapter is in conformance with the deepness of the discussed topic.

Figures and tables are suggestive and offer tangible information about the referred subject. Conclusions of each chapter help the reader to be oriented easily in the topic.

Chapter 1 is the introduction of visual impairment problem, research motivation and scope, objectives, research methodology, short explanation of the own scientific contribution to the publications and organization of the document.

Chapter 2 gives overview of assistive technology, computer vision, machine learning and deep learning. It is explained the difference between the supervised and unsupervised learning methods. Advantages and extra computation capacity of Deep Learning (DL) over the Machine Learning (ML) solutions is well mentioned. Importance of the Recurrent Neural Networks and Convolutional Neural Networks is highlighted correctly, proving research skills of the candidate in this field.

Two types of experimentations applied are implementation of DL architectures to detect markers and obstacles, and mobile application for indoor navigation. Far marker detection was made by local hardware, while obstacle detector was running on Google cloud environment.

In ***Chapter 3*** taxonomy and detailed review of the related literature are presented. State-of-the-art concerning outdoor and indoor navigation and obstacle detection is presented. While outdoor navigation is well solved by GPS services, it is highlighted issue of indoor navigation because of higher precision requirements in smaller distances.
Tens of implemented indoor systems mentioned in the literature from 2010 to 2020, based on RFID, NFC, QR codes, Augmented Reality (AR) markers, Computer Vision (CV) techniques are explained, each of them enrolled in three classes: tag-based, CV and hybrid. Sensing the tags is mentioned correctly to be the main visibility issue of such systems. Hybrid solutions with multiple technics used in parallel are considered possible method to enhance the accuracy of marker detection.

Chapter 4 discusses assistive technologies of the indoor navigation systems for PVI, comparing solutions of current location identification, finding shortest path and safely navigation to the destination. It is presented in detail the navigation prototype for PVI using Aruco tags and the own developed mobile application. Testing process of this own developed system was executed in two cases and documented persuasively.
Usage of eight markers at each interest point made easier the detection of these signals quantified by a metric called mean Navigation Efficiency Index processed on path with 12 sub-paths. Detection service of obstacles was tested with five object classes scanned earlier and identified with YOLOv3 and Tiny_YOLOv3 models based on Convolutional Neural Networks (CNN). The models were compared from the aspects of accuracy and execution time, resulting better performance but longer execution time for YOLOv3.

Chapter 5 discusses improvement of marker detection for longer distances. Identification of marker IDs was improved by separating two steps of marker contour detection from the marker content interpretation. Possible noisy conditions are highlighted in the process of Aruco markers identification.
Because of computation time longer than the required for real time classification, it was necessary to simplify the original CNN model from three to two Feature Extraction Units. Performance of the methods affected by occlusive conditions of the markers are proved with graphs in satisfactory explanation manner.

In ***Chapter 6*** are presented marker detection cases in challenging conditions using Tiny-YOLO-v3. Different versions of YOLO model are applied and compared to perform detection and identification of the markers. Version 1 of the models is generated by modifying the depth of the feature extractor network. Version 2 of the models has variable number of extra branches of the detector network. Combination of these two is forming version 3 of the models.
In general, discussion and evaluation of each model in this chapter is confirmed by well documented experiments including plot of metrics like precision, recall, F1 score, mean Average Precision (mAP), inference time and accuracy.

Chapter 7 concludes the research work and lists shortly the thesis assumptions. To improve the navigation of PVI in indoor environment by computer vision, map of markers

path is created and installed. Characteristics like distance, number of steps, direction of the edges between points of interest should be stored into a dedicated database of the cloud. Voice assisted application on smart phone, visibility of markers and continuous Internet connection is suggested to use during the navigated movements of PVI. YOLO CNN based own proposed management solutions of navigation issues are well summarized.
Three subchapters give shortly thesis assumptions concerning requirement specification of marker map creation, improvement of far marker identification with CNN, finding the optimum YOLO CNN variant for enhancing the marker detection, and identification in accuracy and inference time.

3. Thesis booklet

This 14 pages document includes introduction of research subject of the PVI navigation in indoor environment, research objectives of the candidate, summary of the theses, application of the scientific results and list of own publications. The target, the structure and the specification of the document contains all the necessary elements.

Chapters 1 and 2 serve as problem description and objectives of the scientific research in satisfactory level of elaboration.

Chapter 3 is listing the theses summary. Radical modifications has been made in this final version of the document offering right presentation form to the scientific assumptions.

> ***Thesis I*** refers to the preparation requirements of the general indoor navigation system for PVI based on assistive service. Based on the analysis and discussion of the dissertation subchapter 4.2.1, 4.2.2 and 4.2.3 I accept it as published scientific result of the candidate.
>
> ***Thesis II*** refers to the improvement of the navigation system by using YOLO based CNN at the detection efficiency of the far markers. Based on the analysis and discussion of the dissertation subchapters 5.2 and 5.3 I accept it as published scientific result of the candidate.
>
> ***Thesis III*** refers to the finding optimum YOLO CNN type for identification of Aruco markers using image thresholding and rectangle extraction. Based on the analysis and discussion of the dissertation subchapters 6.1, 6.2 and 6.3 I accept it as published scientific result of the candidate.

Application of the results in practice are mentioned shortly in ***Chapter 4*** of the document. List of own publications with detailed reference to each individual thesis is given in ***Chapter 5***, proving accurate research work executed by the candidate.

4. Final conclusions

Issues from technology and user experience aspects of indoor navigation systems and services for person with visual impairment is discussed in proper level in the two prepared documents. YOLO CNN based Computer Vision assisted system is proposed to help easy navigation of PVI. Application development, measurement series of experiments and evaluation of different metrics is performed to determine optimum solution, taking in consideration engineering and human facilities in practice.

Main contribution of the research work should be mentioned in the field of Aruco marker detection and identification and determination of optimum YOLO CNN type for this issue.

Based on the adequate research work published in the nine journal and conference papers and the joint analysis as dissertation it can be concluded that the scientific quantity and quality proved is sufficient for PhD degree. The theses meet the scientific requirements of publication. I support awarding PhD degree to the candidate.

5. Questions to the candidate

Q1) In section 4.1.1 proof of higher performance of Aruco markers than of QR codes is based on Aruco and open-source library (Zxing), respectively. Which considerations guarantee that Zxing method is representative for QR code recognition? What other methods exist for QR code recognition in the literature? What are their performance compared to Zxing method?

Q2) To reduce the uncertainty and movement time easy help is required to PVI. How the routing algorithm described in subsection 4.1 manages situation when the PVI forgets the source marker ID or gets lost? Would nearest marker be detected and announced by audio feedback "This marker is the closest" or similar with the same meaning during the 360° rotation of the uncertain PVI?

Q3) It is stated that having eight identical markers at each point of interest, the detection of signals is more accurate. Does the physical size of merged marker matrix or the number of identical signals improves the navigation efficiency? Why?

Q4) Convolutional neural network type was applied in the deep learning model during the study. Could be used any other type or neural network to enhance the identification efficiency of the markers? Why?

Debrecen, 27 October 2021

Dr. Zoltán Gál
Associate Professor
Faculty of Informatics
University of Debrecen