Mobile Cognitive Indoor Assistive Navigation for the Visually Impaired

Bing Li[1], Manjekar Budhai[2], Bowen Xiao[3], Liang Yang[1], Jizhong Xiao[1] (✉)

[1]Department of Electrical Engineering, The City College, City University of New York, New York, USA

[2]Department of Computer Engineering, The City College, City University of New York, New York, USA

[3]Bergen County Academies, New Jersey, USA

bli@ccny.cuny.edu, mbudhai000@citymail.cuny.edu, bowxia@bergen.org,

lyang1@ccny.cuny.edu, jxiao@ccny.cuny.edu

Abstract

Indoor assistive navigation system plays an essential role for the independently mobility in unfamiliar environment for the Blind & Visually Impaired (BVI). It has been researched extensively in recent years with the fast evolution of mobile technologies, from applying robotics simultaneous localization and mapping (SLAM) approach, deploying infrastructure sensors to integrating GIS indoor map databases. Although these researches provide useful prototypes to help blind people with independently traveling, cognitive assistant is still in early stage in the era of deep learning for computer vision. In this paper, we propose a novel cognitive assistive indoor navigation system to provide blind people cognitive perception for the surroundings during the navigation journey based on deep learning computing. First indoor semantic map database is built to model the environment spatial context-aware information based on Google Tango visual positioning service (VPS), then a TinyYOLO (You-only-look-once) convolutional neural network (CNN) model is applied on the Tango Android phone for real-time moving person recognition and tracking, and finally scene understanding using CNN and long short-term memory (LSTM) network is performed on the Cloud Server. Experiments by blind people and blind people-folded sighted subjects shows the system works effectively to guide the user to a destination and to provide ambient cognitive perception.

Keywords Indoor Assistive Navigation, Cognitive Computing, Semantic Maps, Real-time People Tracking, Scene Tell, Blind People, Google Tango

1. Introduction

To travel and understand unfamiliar environments is always a challenge for BVI individuals. For the indoor scenarios, the maps information in the medium of braille and audio materials provides conventional assistant for BVI. In the past decades researchers have developed various assistive navigation and electronic aids devices to help BVI with mobility. The WiFi access points and Bluetooth low energy iBeacons are the most common infrastructures devices can be deployed in the environment to provide localization and approximation for the system, then the spatial-context information stored in geographic information system can be delivered to the user according to his/her current location.

The computer vision approach becomes an alternative way to assist BVI as the vision substitution, since it is greatly augmented by more and more powerful hardwares (like GPU and embedded ARM chips) and softwares (like visual odometry and deep learning) in recent years. It has been applied by the researchers from both robotics and computer vision communities. The vision cameras provide real-time context-rich information for the agent (either BVI or robot) at low cost and with minimal weight. These low-latency perception in the forms of 2D RGB or 3D RGB-Depth images makes the visual odometry available to be used for the indoor localization without any extra infrastructure sensors setup in the environment. Further, with more data in the indoor venue maps and the potential of leveraging the deep learning for visual processing, cognitive assistive navigation and scene perception becomes promising to help BVI with high-level understanding for his/her traveling and surroundings.

In this paper, we propose a vision cognitive assitive navigation system based on our previous researches on indoor assistive navigation (Li et al.), we augment our intelligent situation awareness and navigation aid (ISANA) App with cognitive perception functionalities to help BVI with indoor traveling. There are mainly three components designed and evaluated for the system. (1) Semantic maps, which is constructed based on building architecture CAD drawing and is modeled for cognitive spatial-context navigation, and way-point navigation is performed based on Google Tango visual positioning service (VPS). (2) Scene dynamics understanding. The moving people are selected to verify the deep learning CNN vision-based cognitive detection and tracking, running real-time with ISANA on the mobile phone. (3) Scene tell. We perform CNN+LSTM network to perform query-based cognitive scene recognition, which requires high-performance computing and runs on the remote Server.

2. Discussion

2.1 Related Work

Portable mobile phone has became an indispensable tool in our daily life. To help BVI with indoor assistive navigation using mobile device, various research prototypes have been developed in the recent years using the on-board sensors. WiFi fingerprinting based on received signal strength indication (RSSI) classification or triangulation approaches provides a classical localization solution to augment mobile assistive navigation (Au et al.Zhuang et al.). Bluetooth low energy iBeacons is also applied as a supplementary approach for approximation (Ahmetovic et al.van der Bie et al.). Magnetic fingerprinting also has been applied by some research for the indoor positioning (Zhang et al.). More recently, computer vision based approach provides an alternative and promising mobile indoor assistive navigation solution (Möller et al.Li et al.Hu et al.), since the mobile devices becomes more and more computationally powerful. In addition it doesn't require any infrastructures sensors and devices deployed in the environment, which makes this solution scalable.

To move towards context-aware assistive indoor navigation so that the user can be aware of the surroundings, the context information has been firstly added into the system. A context-aware indoor navigation system (CoINS) is presented for user-centric path adaptation to help people find their destination according to their preferences and contextual information (Lyardet, Szeto, and Aitenbichler). To support indoor navigation, CoINS applies a hybrid models of symbolic graph-based models and geometric models for semantic spatial-context awareness and path planning correspondingly. A indoor semantic maps generation framework has also been research to parser the AutoCAD floor plan drawing for the visually impaired pre-journey training (Tang et al.).

For cognitive visual recognition to assist blind navigation, object and text detection has been later applied using image recognition and classification approaches. A comprehensive review for assistive text reading was summarized for natural scene and hand-held object recognition for blind persons (Yi and Tian). A text spotting and signage recognition aid using optical character recognition and template-matching approaches was implemented and evaluated to help the visually impaired with contextual information understanding (Rong et al.).

Towards an intelligent and cognitive blind navigation, mobile-cloud computing has became a popular framework to assist the user. A mobile-cloud based integrated approach was early presented for blind navigation with context-awareness, person recognition and object recognition (Angin, Bhargava, et al.), however it only shows the

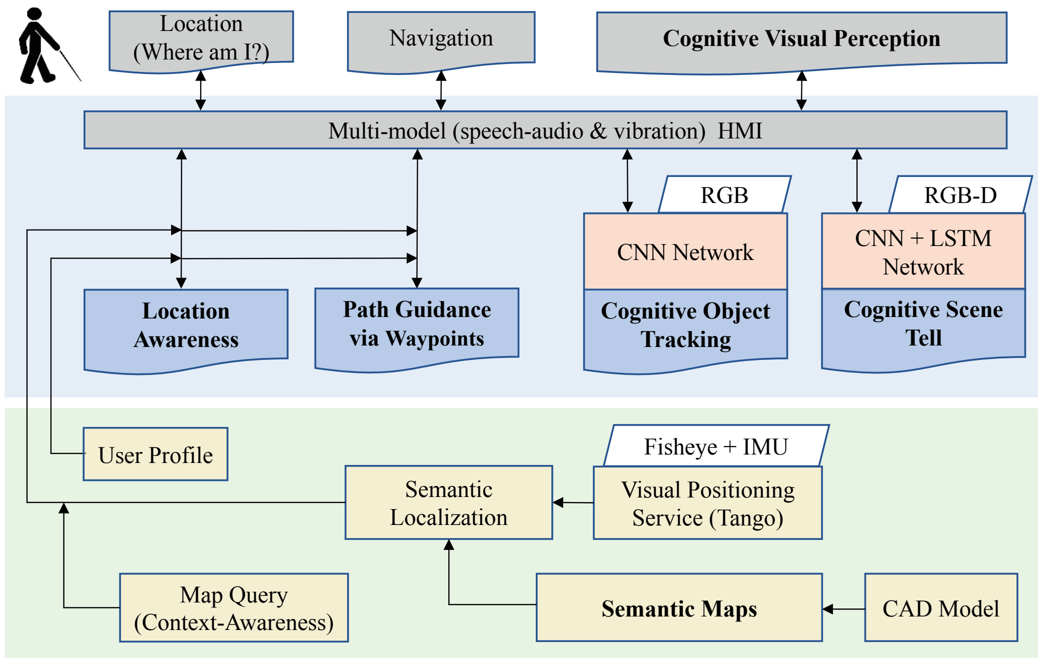

Figure 1: System architecture

concepts of these functionalities, rather than implementation and experimental evaluation. Researchers from IBM and Carnegie Mellon University presented the pioneer work of NavCog App as a navigational cognitive assistant (Ahmetovic et al.Ahmetovic et al.) powered by IBM Watson system. NavCog shows state-of-the-art research for cognitive computing to assist blind navigation.

2.2 System design and implementation

In this project, we aim to develop an intelligent assistive navigation App with cognitive perception solution to help BVI in indoor environment based on the cutting-edge deep learning computing approaches. Using our ISANA App, The user will be able to navigate his/her-self to a specified destination, to query the spatial-context information of the environment, to be aware of moving object in front him/her (moving people is evaluated in this paper), and to understand "what's going on there?" or "what's the scene in front of me?' (we called Scene Tell in this paper). In addition, a friendly multi-model (speech-audio and vibration) human machine interface (HMI), and metric grid-map based path planning have been integrated in the system. The system runs on the Google Tango mobile phone and is based on VPS of vision-based visual odometry and localization, without any extra sensors deployed in the environment. The architecture of the system is illustrated in Fig. 1.

Figure 2: CCNY Testbed geographic layers from semantic maps showing on Google Maps

2.2.1 Semantic Maps

The indoor maps modeling provides essential prior knowledge to support location-based services (LBS) and spatial-context awareness for the system. Building information modeling (BIM) is the process to generate and manage the model of the physical building, and describe the usage of building spaces. We developed an Indoor Maps Editor to parse the architectural CAD drawings and extract the spatial geometric information of the environment. A screenshot of the extracted geographic map layers for our testbed is as shown in Fig. 2.

Then the indoor geographic information is encoded in SQLite database for ISANA, such as walls, room text labels and doors as illustrated in Fig. 3.

Based on the parsed geographic layers, we further retrieve the occupancy grip map and topological connections between room labels and doors (Li et al.), as shown in Fig. 4. We call it Semantic Map in our research, and it includes necessary information to support BIV navigation and location-based services functionalities.

Using IndoorGML as a reference, we simplify the spatial modeling with limited implementations to support the context-aware indoor assistive navigation and other LBS. Based on our previous research on the indoor maps modeling (Albuquerque and Xiao), we supplement our database by extending our Semantic Map SQLite tables with

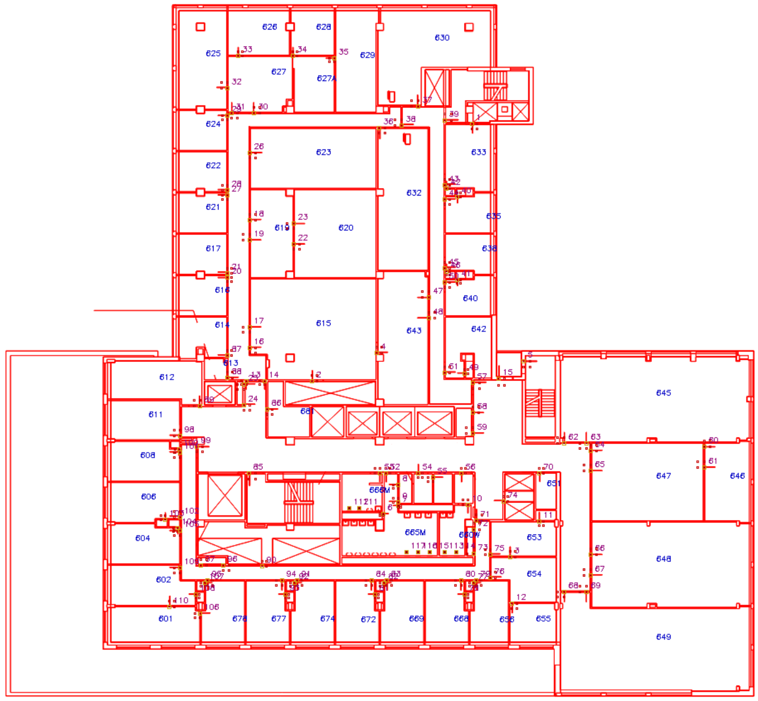

Figure 3: Retrieved geographic layers from architectural CAD drawings

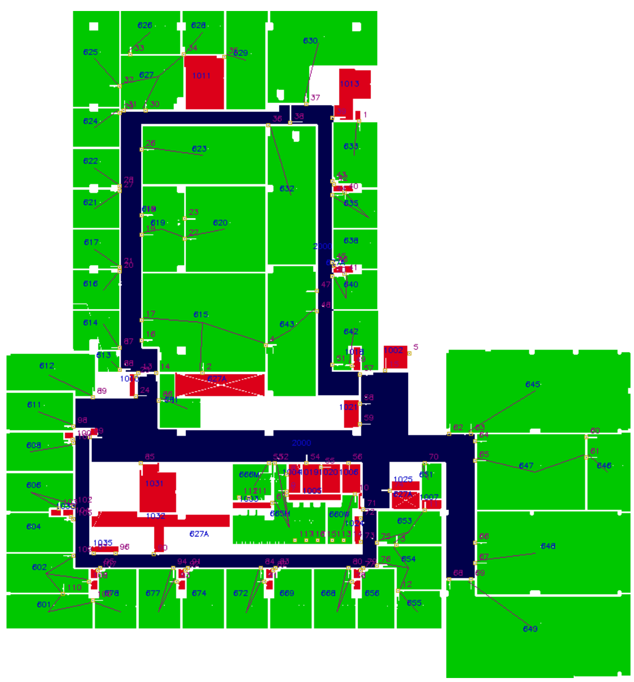

Figure 4: Retrieved semantic map (Li et al.) to build raster and vector models

Table 1: Indoor maps database table for grid map

Field Name	Field Type	Comment
id	*INTEGER* primary key	Auto increment, not null
addressID	*INTEGER*	Index to map address
submapID	*INTEGER*	Sub-map id within the floor
rotation	*REAL*	Alignment rotation angle
trans_x	*REAL*	Alignment translation x
trans_y	*REAL*	Alignment translation y
img	*BLOB*	Occupancy raster grid map

hybrid raster model and vector symbolic model. The raster model is used to support metric-level path planning, sensor perception updating and semantic alignment between raster map with VPS; and the vector symbolic model provides high-level semantic topological connections between semantic landmarks and multiple floors within a building.

1) Raster model

The 2D grid map retrieved from the architectural model is modeled as raster data, which is stored in ISANA indoor maps SQLite database as shown in Tab. 1. The attribute of *addressID* points to the index of map address with floor number and *submapID* specifies the sub-map id within the floor in case multiple sub-maps exists in one floor. Attributes of *rotation*, *trans_x* and *trans_y* hold the semantic alignment between raster map with VPS, and *img* contains the *BLOB* for the occupancy map for metric path planning.

ISANA provides high-level semantic localization based on the VPS of area description file (ADF) within Google Tango device. A Singular Value Decomposition (SVD) approach is utilized to perform the alignment between semantic map and VPS (Li et al.).

2) Vector symbolic models

POI: point of interest, the room texts and doors are modeled as points for the map model. Attribute of Accessible or inaccessible can be applied.

AOI: area of interest, which is represented as polygons for the map model. An accessors attribute is applicable to certain sides to denote the accessible directions of the area. The Accessors are defined as the geometry features where the AOI can be accessed from and it is part of the AOI. Accessible AOI space may not need to have an accessors since it can not be access into.

An AOI can have multiple child AOIs, and the decomposable or non-decomposable attribute is applicable for the AOI. To ensure spatial query, we assume the space decomposition as:

a) a child accessible area will be fully embedded in its parent accessible area;

b) all child accessible areas of a parent accessible area will be without overlap;

DNG: directed navigation graph, the connections between POI to POI or AOI are modeled as the topology information, and it is applied for auxiliary navigation. Such as supporting multi-floor transition, specifying the door nearby the hallway for a destination room. However, to provide assistive way-point guidance to the user without controlling the user's mobility behavior, ISANA is using metric path planning rather than DNG for navigation guidance.

2.2.2 Cogntive Object Tracking

Real-time object-oriented perception during navigation is an admirable functionality for BVI, since it involves with the particular perception concerns for the user, such as the awareness of people approaching/leaving, hazard sign fence, safety banner sign, stairs or cleaning cart in front. Deep learning approach provides a powerful tool for the real-time cognitive image recognition. However, to perform this real-time task in embedded system (such as portable mobile phone) is still in earlier research and is a very challenging task.

In this paper, we select the awareness of people approaching/leaving as the problem to verify the real-time cognitive perception and tracking capability on a mobile phone during the assitive navigation. There are two main reasons to choose it for the experimental evaluation. First the detection and tracking of moving objects (person) during the navigation is more challenging than static objects, and it is a typical scenery of real environment involving the dynamic of both agent and objects. Second, the on-line labeled human data is more accessible compared with other subjects which facilitates our evaluation procedure. Technically the verified approach in this section will be applicable for real-time detection and tracking of any other object types and multiple types in both of static and dynamic environments with proper training.

The cognitive real-time people detection and tracking for blind navigation in this section involves four main steps: Data and data preprocessing, Model training, Model reference for detection and Real-time tracking.

1) Data and data preprocessing

There are various datasets for the people detection in the forms of RGB or RGB-Depth. Conventionally most of the datasets are designed for the whole/upper body detection for surveillance applications, however during indoor assistive navigation, the camera is working as the vision substitution for BVI and is with the closer view for the environment. Thus to gain robust detection, we choose the head region labeled datasets for the CNN training and verification, since these regions in the camera view will be independent without overlapping and suffers less with the clothes color variations than whole/upper body regions.

The HollywoodHeads person head region labeled dataset is selected for our training and evaluation. It includes $369,846$ human heads annotated in $224,740$ movie frames (Vu, Osokin, and Laptev). Most of the camera views are close to the people surrounding and make it ideal for the cognitive detection during indoor navigation.

The pre-processing of the datasets normally involves data cleaning, augmentation and normalization. Since the selected dataset is with large amount of data and is well annotated from high quality image frames, we performed three pre-processing steps for our off-line training. First, the training data input is in the form of $.jpg$, thus the HollywoodHeads dataset is converted from the $.jpeg$ to $.jpg$ format. Then for the annotations, we parser the input *.xml* labeling boxes for annotated boxes, perform the normalization for boxes starting positions and size in the frame and export it as genetic *.txt* for training. Finally a subset of $50,000$ annotated images from the HollywoodHeads dataset is selected as the validation dataset.

2) TinyYOLO model fine-tuning and off-line training

You-only-look-once (YOLO) is one of the state-of-the-art object detection CNN model for real-time application and has been training based the VOC and COCO datasets (Redmon and Farhadi). Compared with conventional classifier-based model, it divides the whole input image into cell regions and predicts the bounding boxes and probabilities for each cell. Thus it only read the input data once and makes it extremely fast for real-time detection Scenarios.

The TinyYOLO CNN model is fine-tuned and is applied for this research on the mobile phone and runs along with BVI assisive navigation functionalities. TinyYOLO is a type of Darknet [1] reference network and runs much faster than standard YOLO convolutional neural network, since it is with only 9 convolutional layers, 6 max-pooling layers and 2 fully connected layers at the end as the region layers. The architecture of the TinyYOLO CNN model is as shown in Fig. 5.

To train the customized detection model, we fine-tune the TinyYOLO network and train the model. The original TinyYOLO network is trained for twenty defined categories. For our detection, we define it as the specified one category of “head” region.

As the post-processing step after the off-line training, the latest checkpoint of the TinyYOLO network graph

[1] https://pjreddie.com/darknet

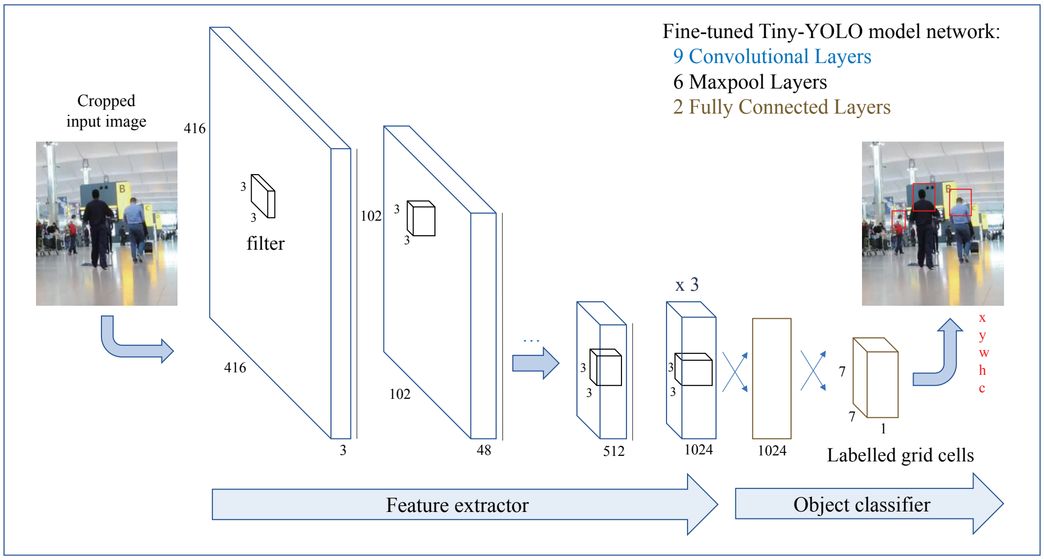

Figure 5: Fine-tuned TinyYOLO CNN network structure, output: x, y, w, h, c as the staring coordinates, width and height and catagory of detected object bounding boxes

model is converted into protobuf *.pb* file along with the meta-data *.meta* JSON file using Darkflow [2]. These two files contains the network graph anchors and labels information. The protobuf file can be used to migrate the model graph to more general platforms, including the mobile devices applied in this research.

3) Model reference for detection

With the TinyYOLO protobuf model file, we implement the recognition reference on the Tango phone using Tensorflow. The whole process is running along with ISANA only on the mobile phone during the navigation.

The software module for model reference is implemented in a hierarchical and object-oriented design. In the top, a *NavigationActivity* class is designed to handle the Google Tango API sensor setup and accessing, and it is the main Activity extended from *CameraActivity* class. Then we design *DetectorActivity* class for the detection processing, which handles the TinyYOLO configurations and initialization, and image pre-processing for necessary data format conversion (from *ARGB* to *YUV*420 in our design). Finally the *TinyYoloDetector* class is implemented using the Tensorflow on the bottom level for the object classification inference. The whole visualization of the App GUI with camera view is rendered using *Canvas* and *GLSurfaceView* class. Due to the GPU CUDA driver limitation on the Tango device, we perform the detection on the CPU and takes around 963.4*ms* for each detection based on our statistical average estimation.

[2]https://github.com/thtrieu/darkflow

4) Real-time tracking

Since the TinyYOLO detection runs on a background thread, and it is not sufficient for smooth tracking and visualization due to low detection rate. A *MultiBoxTracker* is then applied on the UI thread to track each detected object. The tracker extracts the optical flow key-points of *FAST* features at every frame and update the tracking position using the normalized cross-correlation when it is higher than a marginal correlation threshold, otherwise dropping the track. Finally each detected object is assigned with an unique tracking id. The tracker runs efficiently with around 10.6*ms* per frame on the ISANA App and shows the smooth tracked bounding boxes for the moving people.

2.2.3 Cognitive Scene Tell

It should be noted that one of the most challenge and promising task for the assistive navigation is to provide a complete sense of field for visually impaired people. Recent techniques, CNN and RNN, enable us to do basic image caption for almost all life scenarios, which are originally proposed based on translation model. Base on the CNN, we propose to utilize the CNN+LSTM visual caption techniques (Vinyals et al.Fang et al.Xu et al.Wu et al.) for scene understanding, that is, the capability to tell BVI "what's going on in front?". The CNN+LSTM framework deploys CNN model for feature detection as input for a recurrent neural network (RNN) based translator, then the decoder generates the caption based on the image features. Furthermore, we propose a 3D annotated caption on top of the RGB scene captioning for spatial relationship.

1) A **deep CNN** model performs classification for the input original image. The neural network does not doing soft-max by itself, it tries to output the features of each object detected in the image independently.

2) An **attention model** is introduced for feature importance evaluation, where the attention model shifts weight of each learned feature. Then it encodes based on the adapted weights for decoding.

3) The final **LSTM decoder** is a normal decode which can translate the encoded features for text caption.

The reason of introducing attention model is that typical vectorized encoding approaches for CNN lack representation of importance for each feature, thus it is hard to converge as well as perform accurate representation. The attention model here can select the features which have higher importance in post representation, thus generates caption which is closer to real representation of the image.

2.3 Experiments

There are four modules implemented in this research to supported cognitive blind navigation. The assessment of the modules of the whole cognitive framework is conducted in the Tango phone and Cloud Server.

The Cloud Server in this research is equipped with two GeForce GTX 1080 cards and each of them has 2560 GPU cores and 8GB memory on board. The server system is running with Ubuntu 14.04 LTS with dependencies installed. On the Cloud Server, the indoor maps editor is implemented based on Qt Creator and SQLite using Netbeans IDE, the Tiny-YOLO CNN model is trained using Darknet. and the scene tell service is implemented based on Tenforflow and VGGNet model running on Ruby on Rails framework. The ISANA assitive navigation App is implemented on a Lenovo Phab 2 Pro Tango phone based on Google Tango VPS SDK, and the Tiny-YOLO CNN reference is integrated with ISANA for real-time moving people detection and tracking.

The Draknet is applied for the Tiny-YOLO CNN model training on the Server and it is programmed in the C. For the off-line training, we set the batch size as 128 which means the number of images for each training step, the training subdivisions as 8, the learning rate as 0.001, the momentum as 0.9 and the decay as 0.0005. The filters of final fully connected layers is defined as $(classes+5) \times 5 = 30$ for our one category case. During the training procedure, the average loss (AL) is observed to determinate the training procedure when it decreases lightly which is selected around the 6200 iterations in our training. Then to avoid over-fitting in the trained datasets, we look back into the intersection over union (IOU) of historical weight files and select the one with maximum IOU as the final trained model, which is 78.32% for our case. IOU is a metric to determine how accurately is our trained model and is defined as area of overlap (AOO) over area of union (AOU). Finally we convert the trained graph model to a protobuf file with the size of 63MB, which is in the proper size to be migrated into Android ISANA App.

The real-time object CNN detection and tracking on Tango phone is integrated into ISANA. During the navigating procedure, the App will read the image data from Tango API and perform multi-person detection at the rate of around $1HZ$. A Multi-Box tracker is implemented for smooth tracking at the camera frame rate. Fig. 6 shows the ISANA App screenshot with multi-person tracking during the navigation traveling, and the evaluation demos of ISANA system can be seem at demo video [3], and with real-time cognitive people tracking demo video [4].

During the navigation, the user can interact with ISANA to perform a scene-tell understanding for the view in

[3] http://tinyurl.com/isanaeccv
[4] https://youtu.be/eB_Yxr93Tmc

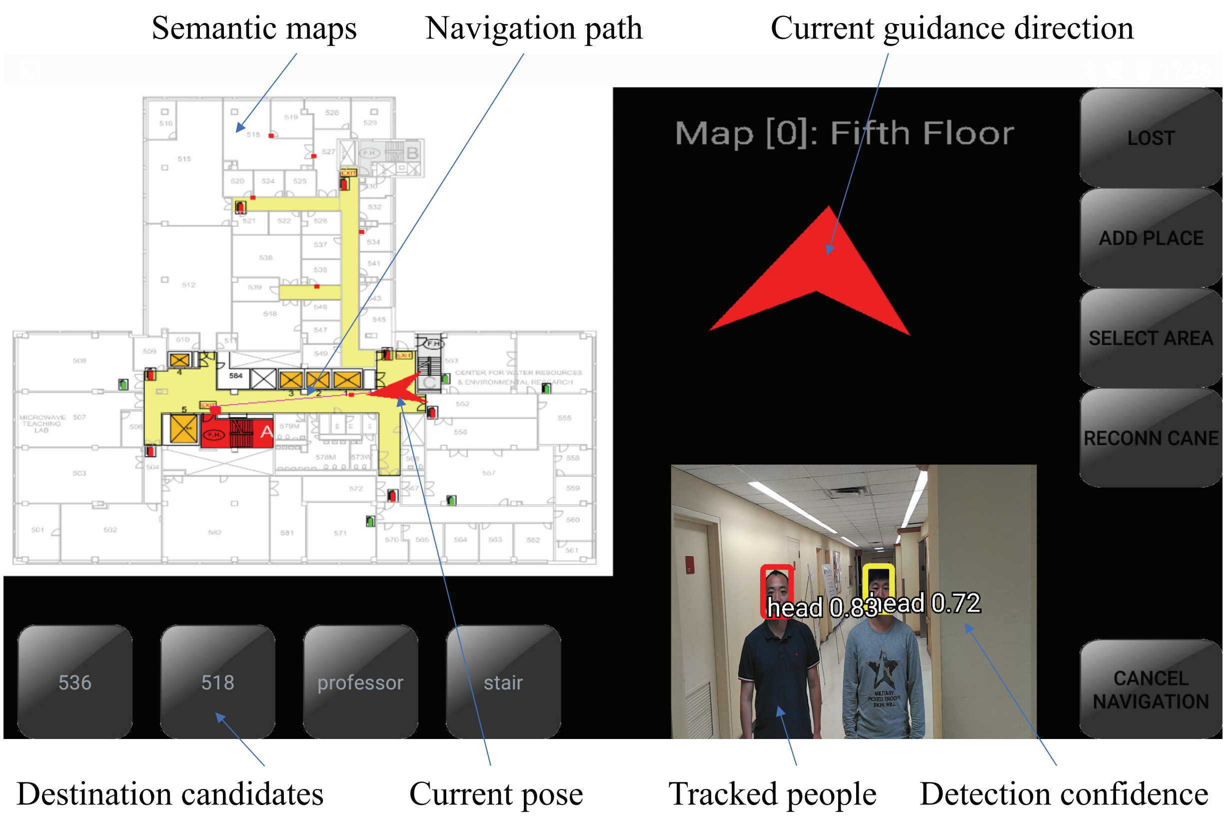

Figure 6: ISANA App screenshoot: real-time cognitive detection for moving people, and tracking which is indicated by different color

Figure 7: Automatically generated scene description from CNN+LSTM network for our testbed building scenes

front of him/her. The CNN+LSTM network is implemented using Tensorflow based on VGGNet recognition model. The detection results of the testbed building for the scene understanding and tell are as shown in Fig. 7.

3. Conclusions

In this research, we augment our mobile intelligent situation awareness and navigation aid (ISANA) App to help BVI with cognitive indoor assitive navigation based on deep learning CNN. The system consists of indoor maps editor to parse the semantic maps from building architecture model to support spatial-context awareness, real-time cognitive objects (people) recognition and tracking on the mobile phone and query-based supplement scene tell computing on the Cloud Server. Our future research will focus on working with BVI subjects to integrate the user preference into deep learning model.

Acknowledgment

This work was supported in part by U.S. Federal Highway Administration (FHWA) grants DTFH 61-17-C-00007 and DTFH 61-12-H-00002. Dr. J. Xiao thanks Google Project Tango for providing a grant to CCNY Robotics Lab as well as free Tango devices and technical support. The authors acknowledge Barbara Campbell for her valuable feedback and suggestions on ISANA, and would like to thank Prof. Hao Tang, Luciano C. Albuquerque, Norbu Tsering for their contributions to this research.

Works Cited

Ahmetovic, Dragan, et al. “NavCog: turn-by-turn smartphone navigation assistant for people with visual impairments or blindness.” *Proceedings of the 13th Web for All Conference*. ACM, 2016. 9. Print.

Albuquerque, Luciano Caraciolo and Jizhong Xiao. “2D Vector Map and Database design for Indoor Assisted Navigation.” *Master Thesis* (2017). Print.

Angin, Pelin, Bharat K Bhargava, et al. “Real-time mobile-cloud computing for context-aware blind navigation.” *International Journal of Next-Generation Computing* 2.2 (2011): 405–414. Print.

Au, Anthea Wain Sy, et al. “Indoor tracking and navigation using received signal strength and compressive sensing on a mobile device.” *IEEE Transactions on Mobile Computing* 12.10 (2013): 2050–2062. Print.

Fang, Hao, et al. “From captions to visual concepts and back.” *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015. 1473–1482. Print.

Hu, Feng, et al. “Indoor Localization for the Visually Impaired Using a 3D Sensor.” *Journal on Technology and Persons with Disabilities* (2016): 192–203. Print.

Li, Bing, et al. “ISANA: wearable context-aware indoor assistive navigation with obstacle avoidance for the blind.” *European Conference on Computer Vision*. Springer, 2016. 448–462. Print.

Lyardet, Fernando, Diego Wong Szeto, and Erwin Aitenbichler. “Context-aware indoor navigation.” *Ambient Intelligence*. Springer, 2008. 290–307. Print.

Möller, Andreas, et al. “A mobile indoor navigation system interface adapted to vision-based localization.” *Proceedings of the 11th international conference on mobile and ubiquitous multimedia*. ACM, 2012. 4. Print.

Redmon, Joseph and Ali Farhadi. “YOLO9000: better, faster, stronger.” *Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on*. IEEE, 2017. Print.

Rong, Xuejian, et al. “Guided Text Spotting for Assistive Blind Navigation in Unfamiliar Indoor Environments.” *International Symposium on Visual Computing*. Springer, 2016. 11–22. Print.

Tang, Hao, et al. “Indoor Map Learning for the Visually Impaired.” *Journal on Technology and Persons with Disabilities* (2017): 191–206. Print.

Van der Bie, Joey, et al. "Guiding the visually impaired through the environment with beacons." *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, 2016. 385–388. Print.

Vinyals, Oriol, et al. "Show and tell: A neural image caption generator." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015. 3156–3164. Print.

Vu, Tuan-Hung, Anton Osokin, and Ivan Laptev. "Context-aware CNNs for person head detection." *International Conference on Computer Vision (ICCV)*. 2015. Print.

Wu, Qi, et al. "What value do explicit high level concepts have in vision to language problems?" *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016. 203–212. Print.

Xu, Kelvin, et al. "Show, attend and tell: Neural image caption generation with visual attention." *International Conference on Machine Learning*. 2015. 2048–2057. Print.

Yi, Chucai and Yingli Tian. "Assistive text reading from natural scene for blind persons." *Mobile Cloud Visual Media Computing*. Springer, 2015. 219–241. Print.

Zhang, Chi, et al. "GROPING: Geomagnetism and crowdsensing powered indoor navigation." *IEEE Transactions on Mobile Computing* 14.2 (2015): 387–400. Print.

Zhuang, Yuan, et al. "Evaluation of two WiFi positioning systems based on autonomous crowdsourcing of handheld devices for indoor navigation." *IEEE Transactions on Mobile Computing* 15.8 (2016): 1982–1995. Print.