PREDICTING THE HEAT MAP OF STREET VENDORS FROM PEDESTRIAN FLOW THROUGH MACHINE LEARNING

XINYUE SHOU[1], PINYANG CHEN[2] and HAO ZHENG[3]
[1]*Bartlett School of Architecture, University College London, London, UK*
[1]*xinyue.shou.20@ucl.ac.uk*
[2]*Graduate School of Design, Harvard University, Boston, USA*
[2]*pinyangchen@gsd.harvard.edu*
[3]*School of Design, Shanghai Jiao Tong University, Shanghai, China*
[3]*zhhao@design.upenn.edu*

Abstract. Street vending is a recent policy advocated by city governments to support small and intermediate businesses in the post-pandemic period in China. Street vendors select their locations primarily based on their intuitions about the surrounding environment; they temporarily occupy popular locations that benefit their business. Taking the city of Chengdu as an example, this study aims to formulate the rules governing vendors‘ location selection using machine learning and big data analysis techniques, thus identifying streets likely to become vital street markets. We propose a semantic segmentation method to construct heat maps that visualize and quantify the distribution of street vendors and pedestrians on public urban streets. The image-based generative adversarial network (GAN) is then trained to predict the vendors' heat maps from the pedestrians‘ heat map, finding the relationship between the locations of the vendors and the pedestrians. Our successful prediction of the vendors' locations highlights machine learning techniques‘ ability to quantify experience-based decision strategies. Moreover, suggesting potential marketing locations to vendors could help increase cities' vitality.

Keywords. Machine Learning; Big Data Analysis; Semantic Segmentation; Generative Adversarial Networks.

1. Introduction

1.1. RESEARCH BACKGROUND

Street vending has become a common phenomenon in China as the country's urbanization process has progressed. For the last few decades, vendors' occupation of the streets has been considered illegal or informal. However, despite administrative opposition, street vending is still a prevalent phenomenon. Some scholars have pointed out the inevitability of their existence in the cities. S. Sassen suggested that vendors are a part of the informal urban economy, which is beneficial for small and intermediate businesses to adapt to the ever-increasing

PROJECTIONS, Proceedings of the 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2021, Volume 2, 569-578. © 2021 and published by the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong.

living cost in global cities (Sassen 1991). Street vending is a way of temporarily using urban public space, and vendors can become urban catalysts (Oswalt, Overmeyer et al. 2011) that serve as secondary diversity (Jacobs 2016) and even create urban ”soft edges” (Gehl 2013).

The pro-street economic policies in the post-pandemic era, supported by Chengdu’s government, is a strategy to recover the economic loss of the small-and-intermediate, self-employed economy. Streets are the primary and most essential components of the public realm, and the inclusiveness of vendors and street activities show the robustness of the urban public life and is a crucial measure of civilization (Kuntsler 1996). Popularity is a direct manifestation of the street vitality. The street vendors’ location choices strongly correlate with the distribution of pedestrian flows.

1.2. LITERATURE REVIEW

Previous research has analyzed street vendors as a component of the informal economy in global cities (Sassen 1991). Street vendor location selection strategies have similarities with small retail businesses. Although the big data approach has been widely used in the business location selection process in recent years, few scholars have used information technologies to analyze street vendors’ location selection decisions considering their unique selling strategies.

Heat mapping is a visualization method for geographical analysis that is widely used in quantitative and qualitative analyses in spatial planning and urban studies. The technique has contributed to the understanding of property prices (Elser 2011), crime rates (Sandig, Somoba et al. 2013), traffic planning (Hilton, Horan, et al. 2011), the spatial-temporal characteristics of bus travel (Yu and He 2017), and many other phenomena.

Machine learning builds on traditional big data analysis approaches by using input data and algorithms to estimate unknown future results. The technique learns from data to provide data-driven insights, decisions, and predictions. As inter-correlation does not impact machine learning approaches, machine learning can make better use of the full extent of available data (Reades, De Souza et al. 2019). Among machine learning models, the generative adversarial network (GAN) is a deep learning network that consists of a generator and a discriminator (Goodfellow, Pouget-Abadie et al. 2014). GAN aims to model natural image distribution by forcing the generated samples to be indistinguishable from natural images, which enables the network to learn and generate higher-order features (Schmidhuber 2015).

Since extensive data about human society is increasingly accessible, GAN is now applied to analyze many image-based data in the architecture and urban design realm. Shen et al. (2020) trained a dataset to predict the architectural filling on urban plan drawings based on color-labeled roads, green land, rivers, and other empty space elements. Huang and Zheng (2018) applied GAN to recognize and generate apartment plan drawings by assigning a unique color label to each room type.

1.3. PROBLEM STATEMENT

Previous site selection approaches by on-site observation and big data extraction and analysis rely heavily on researchers' empirical understandings and subjective judgments and can only yield analyses of sites. These approaches can predict neither the patterns governing vendors' location selection nor the distribution on sites with similar features as the analyzed sites. Moreover, on-site observation and recording are limited in their ability to summarize all the spatial and demographic features that affect street vendors' decisions.

Although big data approaches enable researchers to deduce street vendor location selection and distribution based on various factors, their predictive power is restricted by the human capability for data analysis. The processes of big data and simple linear regression analysis also involve a substantial amount of empirical presumption when seeking correlations between population distribution and street vendors.

Each of the aforementioned methods can only make a "best guess" about the existing data for which the analysis is valid. In the urban design and planning realm, a "best guess" is not enough. The full probability density distribution for a new prediction is of major interest. Machine learning provides a way to derive semantic understanding from the collected data.

1.4. PROJECT GOAL

This paper proposes a machine learning method to predict vendor distribution patterns with the input of pedestrian distribution heat maps along urban public streets in Chengdu, China. We paired pedestrian distribution heat maps with street vendor heat maps to train datasets for image-to-image translation by pix2pixHD (Isola, Zhu et al. 2017), a GAN-based neural network. Our ultimate goal is to create an app that recommends street vendor sites (Figure 1).

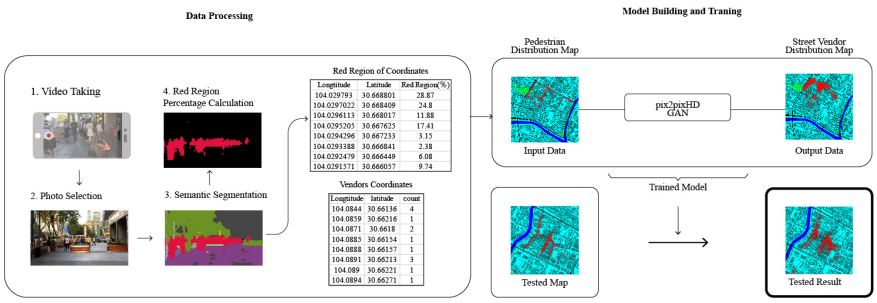


Figure 1. The overview of the workflow.

2. Methodology

2.1. STUDY AREA

In this study, we chose Chengdu, China as the study area. In the post-pandemic era in China, Chengdu was among the first cities to encourage street vending to

recover the economic losses of small-and-intermediate, self-employed businesses. In order to collect the data for the distribution of the vendors and pedestrians, we conducted a street survey from 15 July, 2020 to 30 August, 2020, from 6:00 p.m. to 8:00 p.m. Beijing Time. Rainy days are excluded from our street survey. In the summer in Chengdu, most street vending occurred in the evening. During the period from 6 p.m. to 8 p.m., the number of vendors usually reaches its daily peak. This period's real-time pedestrian flow data has the best correlation with street vendor distribution patterns.

2.2. VENDOR SITE DATA COLLECTION

During the one and half period of time of on-site observation, we explored eight areas of the city. The potential trained and tested sites were selected according to the following criteria: 1) The sites that we chose for creating training datasets were relatively dense in population, and the sites attracted a certain number of vendors every evening. 2) The tested sites and the trained sites had to have similar features. 3) Most sites were located in residential areas with a mixture of (or next to) at least two types of functions 4) The regions were selected primarily based on the organization of main urban arterial roads. Each research region was bounded by at least three arterial roads. Finally, four regions were selected as representative sites for training data, and two regions were selected for testing data.

We used a six-foot geo-coordinate positioning and recording mobile phone application to record the location of every street vendor that we observed during the on-site research period. The latitude and longitude coordinates of every street vendor we observed were recorded on the sites based on the GCJ-02 coordinate system. Vendors clustered within 3 meters were considered to be at the same geo-coordinate.

2.3. PEDESTRIAN FLOW DATA COLLECTION

Population heat maps can show the agglomeration patterns and density distribution of populations. Some mapping software, such as Baidu Maps, AMaps, and Google Maps, provides real-time heat maps by obtaining geographic location information from mobile phone users‘ locations. However, the data collected by smartphone users' locations do not distinguish between people in private and public spaces, in cars, and out on the streets. Also, population data collected within some semi-private building compounds such as residential blocks, schools, and governmental buildings are considered invalid, as street vendors are not allowed to enter those spaces. Apart from that, the data collected by smartphone users‘ locations may cause data bias. The number of smartphones cannot accurately reflect the population, as not all people bring smartphones with them on the street all the time, and many people may turn off their phones' location function. This is especially true in some old residential blocks, which overlap significantly with the areas covered by this study.

Therefore, we constructed heat maps that reflect the distribution of pedestrian flow on the urban public streets as real-time data. We propose a semantic segmentation method to construct heat maps to visualize and quantify the

distribution of street vendors and pedestrians on urban public streets. We applied the pyramid scene parsing network (PSPNet), a deep neural network model trained with the Cityscape dataset, to segment street images (Zhao, Shi et al. 2017). PSPNet incorporates suitable global and local features, assigning each pixel in an image a category label. A single PSPNet yields the mIoU accuracy 80.2% on Cityscapes. In our study, we aim to predict the percentage of pedestrian pixels in input images.

We used video recordings to collect pedestrian flow data along each observed street. The street was regarded as lines or polylines. We recorded the latitude and longitude coordinates of the start and end points of each segment. The videos were exported as a series of static images. As we attempted to maintain a constant walking speed, the images may be regarded as evenly distributed on each section of the streets. Five hundred photographs were generated approximately every 8 meters. A single photo that could best represent the population density at its geo-coordinate was chosen for the semantic segmentation process. The population density was calculated based on the proportion of color blocks of pixels recognized as people within the entire picture (Figure 2).

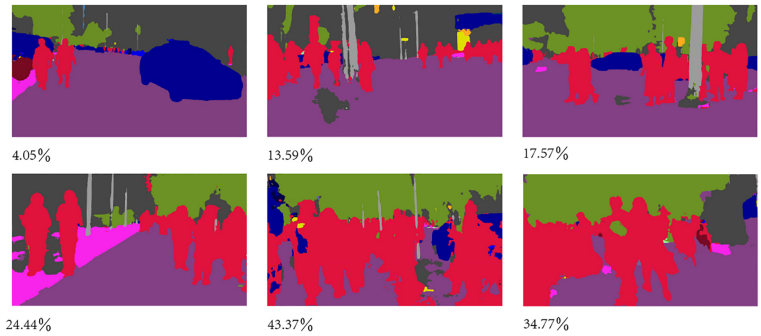


Figure 2. Percentage of people by results of semantic segmentation.

2.4. HEAT MAPPING

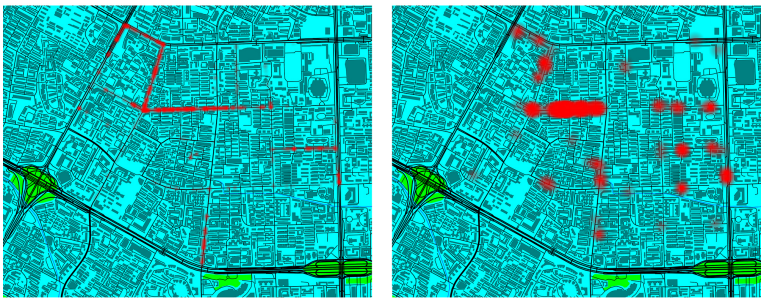
After the data was collected, the heat mapping process followed three parts: 1) The selected pictures were marked with a coordinate based on the GCJ-02 coordinate system. 2) We used the Amap API to create heat maps. AMap is a provider of digital map content, navigation, and location service solutions in China. AMap and the recording app are both based on the GCJ-02 coordinate system. 3) We used the AMap API to create user-defined map styles and add heat map layers from user data.

A labeling rule (Figure 3) was created to label five types of regional surfaces in the map: land, street, building, waterway, and green space. We established a clear differentiation in either G or B values between each pair of regional surface types so that the machine could recognize and distinguish them more accurately.

Regional Surface	RGB values R	G	B	Hex	Opacity	
Land	0	255	255	#00B4B4	100%	
Green Space	0	255	0	#00FF00	100%	
Highway/Road	0	0	0	#000000	100%	
Building	0	128	128	#005050	100%	
Water System	0	0	255	#0000FF	100%	

Figure 3. Heatmap color labeling principles.

The value in the R channel in all regional surfaces was 0, representing the heat maps. The legend bars show the corresponding relationship between the R value and the density of pedestrian flow and vendor distribution. The pedestrian and street vendor distribution heat maps of each block covered the same geographical region, and the paired images were unified to the same size. All the testing images were cut into smaller sections of 400×400 pixels. Then, 4,606 pictures were generated and marked in sequence. The input images showed the heatmap of pedestrians, and the output images showed the heatmap of the vendors (Figure 4).



Input Image -
Pedestrian Flow Heat Map

Output Image -
Vendor Distribution Heat Map

Figure 4. Heat map for training at Yulin District.

2.5. MODEL TRAINING

With the training dataset, we applied pix2pixHD (Isola, Zhu et al. 2017) to train the GAN model. There are two neural networks in GAN, the generator and the discriminator. The generator attempts to output fake images to cheat the discriminator, while the discriminator is trained to distinguish the fake images.

The pix2pixHD method is a conditional GAN framework for image-to-image translation. It consists of a generator G, which translates semantic label maps to realistic-looking images; and a discriminator D, which distinguishes real images

from the translated ones. It contains a coarse-to-fine generator, a multi-scale discriminator, and a robust adversarial learning objective function. The training process was completed on a computer with a GeForce RTX 2060 Graphics Card.

The loss values of the generator and discriminator were recorded during the training process. Figure 5 compares the loss values of the two models. The training is a process in which the generator and discriminator "compete" with each other. When the discriminator loss value is higher, the generator loss value is lower. This changing trend proves the accurateness of the training results.

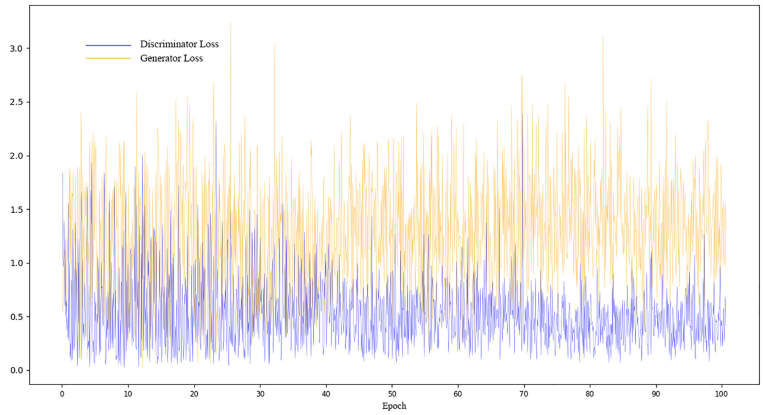


Figure 5. Loss values of Pix2pixHD training.

3. Results

3.1. TRAINING ACCURACY

The network went through all 4,916 pictures during every epoch. After one epoch of training, a representative input image was sent to the neural network. We could determine the completeness of the training by the output image.

Figure 6 shows the image pairs in each training epoch for the training model. We compared the synthesized images of epochs 18, 44, 90, and 100 with their ground truth images. At the beginning of training, the synthesized images produced inaccurate prediction results at a higher rate. In epoch 100, inaccuracies still occurred occasionally, but the output images showed the changing pattern from input images to the ground truth images. Thus, we decided to stop training at epoch 100 for the prediction of street vendor distribution heat maps.

3.2. VENDOR SITE PREDICTION

Figure 7 shows the predicted heatmap of vendor sites in other areas in Chengdu. We collected the pedestrian's heatmap as the input to the neural network and provided suggestions of the site choice based on the output vendors' heatmap.

The generated maps show that, in most cases, points with high pedestrian density and multi-section intersections can attract more street vendors. However,

the trained results show a similar pattern as the input data in that the density of vendors and the density of pedestrian flow did not have an accurate positive correlation. Streets with a continuous flow of people tended to have a high density of street vendors. In the trained dataset, about four sites had a high pedestrian flow density but a low density of street vendors. Approximately three sites with low pedestrian flow density attracted street vendors. However, in the test result, no input images without pedestrian distribution (with a number in the R channel) produced a result with vendor distribution.

On the streets where street markets can be formed, the density of vendors at both ends of a street was not significantly high in our model's predictions. Vendors were more concentrated in the middle of the street in places with a higher density of pedestrians. In contrast, on the streets where pedestrian flow density was relatively low, street vendors were more likely to sell at the junction of the road. This is probably because these places have a higher visibility rate.

The results of this study did not reflect the direct impact of road width and building texture on the distribution of mobile vendors. However, we believe that there is a certain correlation between the density of pedestrian flow and road level. Secondary roads are considered more walkable than main urban arteries but also have a higher pedestrian flow density than tertiary and quaternary streets.

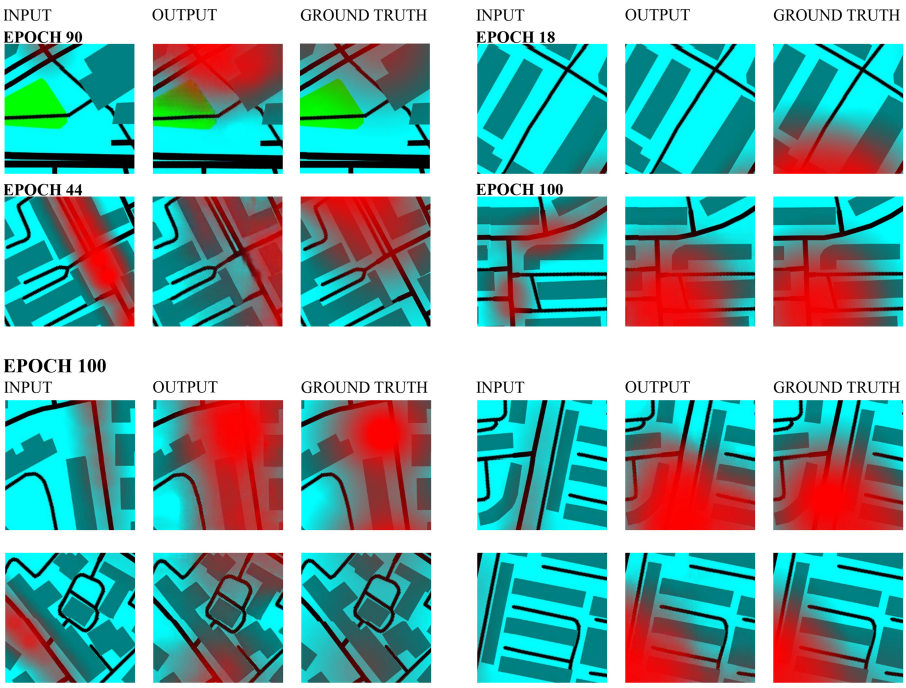


Figure 6. Training results at different epochs.

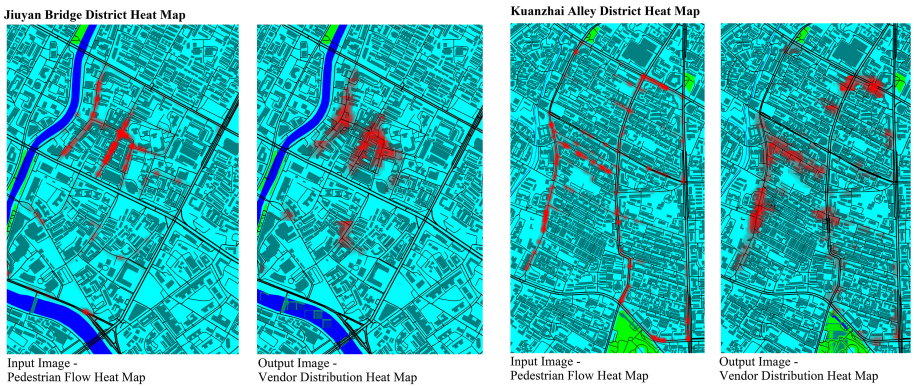


Figure 7. Testing results for predicting the vendors' site heatmap.

4. Conclusion and Discussion

Our machine learning model successfully predicts the heatmap of vendor sites based on the pedestrian heatmap. The successful prediction of vendor location highlights the ability of machine learning techniques to quantify experience-based decision strategies. In the future, we aim to create an App that can display larger scale vendor-site prediction heatmaps based on our proposed methodology. This APP is not only used as a positioning software, but it also entails social meanings: While it guides street vendors of site selection, it can also attract residents to that place. Suggesting additional marketing locations to vendors could help increase urban vitality. Futural urban design needs to take the former 'informal' economy into equal account as other economic sectors. The mapping could also provide designers and decision-makers data support to make a more inclusive city, and our research can largely prevent data biases caused by some data sources.

The question of whether vendors belong to urban spaces reduces to another question: how citizens make use of public space and how much they are willing to stay outdoors for street activities. This study can only illustrate where vendors are likely to stay mainly based on population flow, but the underlying reasons why people are willing to stay at specific locations cannot be explained without considering other socioeconomic conditions. The analysis of streetside behavior is also related to micro urban features, neighborhood dynamics, and the richness and types of nearby POI, etc. Therefore, more indicators should be added to the research to improve the accuracy of the forecasted results for future research.

This study used a labor-intensive method to collect the data due to the current inadequate street view image data from the pedestrian perspective. The cityscape dataset, which is based on street views taken in several European cities, may cause a mismatch with China's second-tier cities, including Chengdu. The better functioning of our App relies on real-time site-specific street view image data and the combination with other data sources from relevant location-based systems.

References

Elser, J.: 2011, *System and methods for comparing real properties for purchase and for generating heat maps to aid in identifying price anomalies of such real properties*, Google Patents.

Gehl, J.: 2013, *Cities for people*, Island press.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y.: 2014, Generative adversarial nets, *Advances in neural information processing systems*, **2014**, 2672-2680.

Hilton, B. N., Horan, T. A., Burkhard, R. and Schooley, B.: 2011, SafeRoadMaps: Communication of location and density of traffic fatalities through spatial visualization and heat map analysis, *Information Visualization*, **10(1)**, 82-96.

Huang, W. and Zheng, H.: 2018, Architectural Drawings Recognition and Generation through Machine Learning, *ACADIA 2018*, Mexico City, Mexico, 156-165.

Isola, P., Zhu, J. Y., Zhou, T. and Efros, A. A.: 2017, Image-to-image translation with conditional adversarial networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1125-1134.

Jacobs, J.: 2016, *The death and life of great American cities*, Vintage.

Kuntsler, J. H.: 1996, *Home from nowhere: remaking our everyday world for the twenty-first century*, Nova York: Touchstone.

Oswalt, P., Overmeyer, K. and Misselwitz, P.: 2011, *Urban catalyst: Strategies for temporary use*, Birkhauser Basel.

Reades, J., De Souza, J. and Hubbard, P.: 2019, Understanding urban gentrification through machine learning, *Urban Studies*, **56(5)**, 922-942.

Sandig, J. D. E., Somoba, R. M., Concepcion, M. B. and Gerardo, B. D.: 2013, Mining online gis for crime rate and models based on frequent pattern analysis, *Proceedings of the World Congress on Engineering and Computer Science*.

Sassen, S.: 1991, *The global city*, New York.

Schmidhuber, J.: 2015, Deep learning in neural networks: An overview, *Neural networks*, **61**, 85-117.

Shen, J., Liu, C., Ren, Y. and Zheng, H.: 2020, Machine Learning Assisted Urban Filling, *CAADRIA 2020*, Bangkok, Thailand, 2:681-690.

Yu, C. and He, Z. C.: 2017, Analysing the spatial-temporal characteristics of bus travel demand using the heat map, *Journal of Transport Geography*, **58**, 247-255.

Zhao, H., Shi, J., Qi, X., Wang, X. and Jia, J.: 2017, Pyramid scene parsing network, *Proceedings of the IEEE conference on computer vision and pattern recognition*.