
StreeTalk
公众服务：慢行导航
步行出行评分
起点
输入地址
终点
输入地址
开始计算
游览路线规划
出发
31.2369, 121.5028
8条线路
10.00
公里
开始搜索
测试版本，若查询时间过长，请刷新页面
线路4: 6.91分, 距离: 10.55公里
线路7: 6.55分, 距离: 11.6公里
线路5: 6.33分, 距离: 10.97公里
线路1: 6.25分, 距离: 11.77公里
线路3: 6.2分, 距离: 10.8公里
线路8: 5.75分, 距离: 9.93公里
线路2: 5.53分, 距离: 10.97公里
线路6: 5.46分, 距离: 10.4公里
评分
距离
复原
清空查询
健康出行规划(开发中)
城市评估：空间风貌(开发中)
StreeTalk 街道感知评分指数
该段街道的安全感评分约为: 8.26分
显示路段街景
餐饮
开关热力图
Powered by CityTech | 城室科技

StreeTalk慢行导航系统

STREETALK: A NAVIGATION SYSTEM FOR PEDESTRIANS AND CYCLISTS

刘浏，张帆，周博磊，王舟童，李颖欣* / Liu LIU, Fan ZHANG, Bolei ZHOU, Zhoutong WANG, Yingxin LI
*上海市普陀区宁夏路777号财通坤田大楼4楼，邮编：200063
customer@citorytech.com

http://doi.org/10.15302/J-LAF-20180210
收稿时间 RECEIVED DATE / 2018-03-19
中图分类号 / TU98, TP18
文献标识码 / B

摘要

不同于当前以出行效率为主要考量因素的传统导航系统，StreeTalk慢行导航系统更注重行人和骑行者的感受，特别是对安全感和舒适度的需求。通过运用物体检测识别及场景语义分割等技术，城市街景的深度图像特征被提取出来。之后，结合深度学习模型，机器可以预判人类对周边环境的感知，并为数以千万计的街景照片评分，从而建立起可以直观显示街道安全和舒适程度的导航系统。相关技术不仅可以提供定制化、全方位的慢行导航服务，为城市研究及决策、城市景观设计提供相关依据，也为未来城市生活创造了更多可能。

关键词

城市街道；安全感与舒适度；慢行导航系统；深度学习；物体检测识别；场景语义分割

ABSTRACT

Different from the conventional efficiency-driven navigation systems, StreeTalk navigation system is developed for pedestrians and cyclists to optimize their travelling safety and comfort. Through the application of technologies including object detection and scene parsing, the characteristics of urban streetscapes can be extracted. By combining deep learning models, machines can imitate human's perception and evaluate streetscapes automatically, and intuitively show users the street safety and comfort level of different commuting options. Relevant technologies can not only be applied in offering panorama navigation services for pedestrians and cyclists, but also better support urban research, inform decision-making and urban landscape design, and explore more possibilities for future urban life.

KEY WORDS

Urban Street; Safety and Comfort; Navigation System for Pedestrians and Cyclists; Deep Learning; Object Detection; Scene Parsing

整理 田晓劼 佘依爽 田乐
译 张健 田乐
EDITED BY Xiaojie TIAN Yishuang SHE Tina TIAN
TRANSLATED BY Angus ZHANG Tina TIAN

背景

目前，中国城镇化率已超过58%[1]，这意味着我国有近六成人口生活在城市中。随着城市环境日益拥挤、复杂，越来越多的市民倾向于在较短交通范围内选择步行、骑行的出行方式。对于不熟悉的目的地，人们也多会通过移动端导航来选择路线。遗憾的是，现有的导航系统在推荐慢行路线时均以“时间最短”为优先，忽视了沿途的景致，也忽视了人们对于安全感的需求。基于此，城室科技团队希望通过研发一款以城市街景及城市空间大数据为基础、以人工智能云计算为技术依托的智能导航系统，以提升公众的慢行生活品质，并为城市研究及决策和城市景观设计提供相关依据。

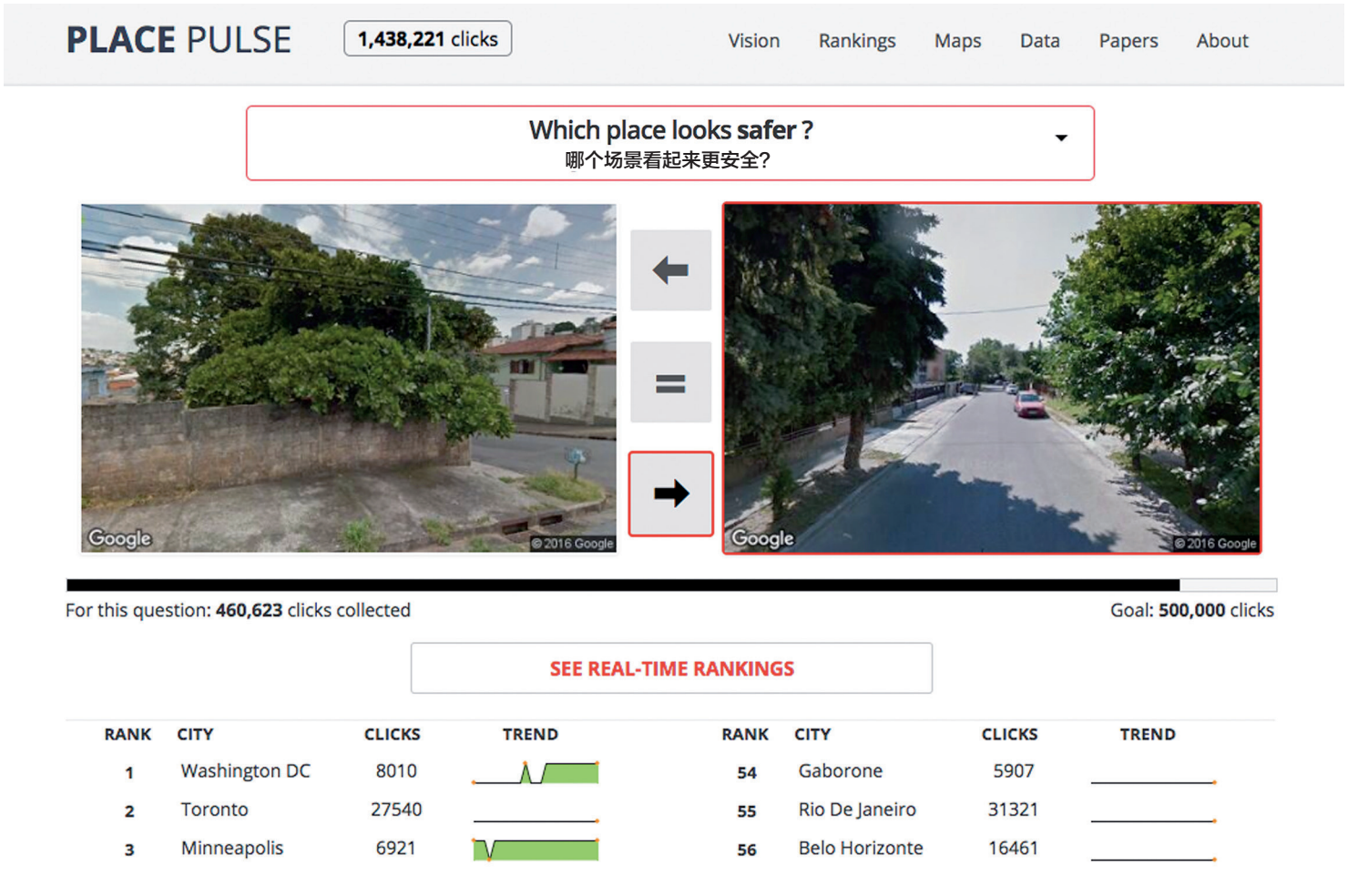
研究与技术

人们在日常出行时除了考虑时长因素，更需要考虑出行的安全感和舒适度。那么如何衡量和计算街道的安全感和舒适度？对街景照片进行比较和评分是一种行之有效的方法。美国麻省理工学院媒体实验室在几年前即采用这一思路来统计人们对于街道的安全感和舒适度的认知。该实验室设计了一个专门的网页，系统随机展示两张街景照片，由来自全世界的志愿者来选择哪张照片中的街道看起来更加安全、舒适。在积累了120余万条对比数据之后，每张被比较的照片都通过计算获得一个0~10的安全感和舒适度评分（分数越高表示街道的安全感和舒适度越高）。通过这一方法，实验室可以获取这些照片所示街道的评分数据[2][3]。

然而，如果要对上海这类特大城市数以千万计的街景照片进行安全感评分，采用上述方法并不现实，这就需要借助深度学习技术来解决问题。通过输入已有的街道评分数据，人工智能可以学习相关特征来对其他街道进行安全感评分，从而大幅提高工作效率。考虑到人类对于街道安全感和舒适度的认知基本一致，我们直接采用已有的国际数据，让计算机从已打分的照片中提取深度图像特征，即通过物体检测识别技术，从城市

1. 基于城市感知的StreeTalk慢行导航系统

1. StreeTalk is a navigation system for pedestrians and cyclists based on city sensing.



首席设计师：
刘浏
设计团队：
城室科技创业团队
设计时间：
2016年9月至今
所获奖项：
SODA上海开放数据创新应用大赛十强

CHIEF DESIGNER:
Liu Liu
PROJECT TEAM:
CitoryTech Entrepreneurial Team
DESIGN PERIOD:
September 2016 to present
AWARD:
Top Ten of Shanghai Open Data Apps (SODA) Competition

街道影像中分析并提取出300余种物体，包括行人、自行车、小汽车、公交车、路灯、花坛等，并通过运用场景语义分割技术对道路、天空、建筑等不同内容的图像进行像素级划分，从而建立起街道安全感知评分系统。基于该评分系统，并结合深度学习模型卷积神经网络①，我们将街道感知拆解为安全、美感、活力、整洁等多个维度，使机器从这些维度进行拟人化感知，从而预判人类对周边环境的感知。

2016年，团队选取上海市作为首个研究对象，在其市域范围内（含郊区）拍摄了230余万张一年四季的街景照片，并通过人工智能感知预判对这些照片进行分类。通过海量的数据标定与模型训练调整，团队打造了全球首个将神经网络模型应用在城市街道感知领域的实例——StreeTalk慢行导航系统。

StreeTalk将所有已打分的街景数据置入路网属性中，当人们使用StreeTalk导航时，系统可以调取该属性，自动计算出沿线每段道路的安全感和舒适度平均分，并优选出高分路线；使用者还可以查看具体路段的实景照片。此外，高分路线的推荐具有一定的智能性。例如，在夜晚，导航会优先推荐有明亮路灯或较多临街店铺的街道；而在炎炎夏日，则会首推绿树成荫的街道。

我们发现，评分较高的街景照片大多有着丰富的绿色空间、干净整洁的马路，而评分较低的街景照片则往往呈现出公共设施匮乏、人烟稀少、环境恶劣的状态。根据不同的评分，

① 卷积神经网络通过引入卷积自动分层提取街景特征，每一层由多个特征图组成，每一个特征图由前一层输出与一个卷积核进行卷积运算和非线性变换得出，随后进行池化操作，用于降低输出维度，同时获得一定的特征不变性。

2. 由美国麻省理工学院媒体实验室设计的用于统计人们对于街道安全感和舒适度认知的网站。

2. An online platform designed by MIT Media Lab to collect the statistics of people's evaluation on street safety and comfort level.

3-1. 物体检测识别技术
3-2. 场景语义分割技术
3-3. 物体检测识别和场景语义分割可以同时进行。
3-4. 每张被比较的照片都通过计算获得一个0~10的安全感和舒适度评分。

3-1. Object detection technology
3-2. Scene parsing technology
3-3. Object detection and scene parsing could be conducted simultaneously.
3-4. Each photo was scored from 0 to 10 on street safety and comfort.

城市管理者可以确定需要进行环境整治的街道；而设计师则可以通过吸取高分街道的街景特征优化设计方案。此外，通过StreeTalk慢行导航系统获得的包括慢跑和骑行爱好者出行路径在内的慢行交通数据，也可为城市慢行绿道的规划设计提供决策依据。

其他应用前景

用以支持慢行导航系统的物体检测识别技术、场景语义分割技术以及卷积神经网络还可为城市规划师及景观设计师提高工作效率、优化工作方法提供更多有效途径：

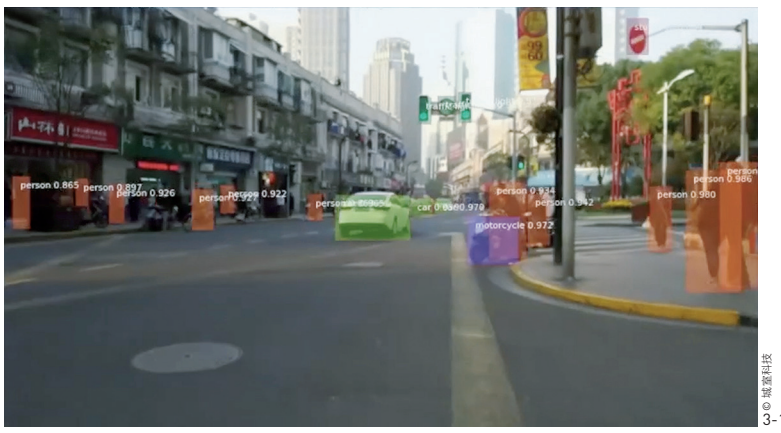
1）街道热力分析：通过物体检测识别技术对街景中的行人、车辆及各类设施进行识别并加以统计，可以迅速获取各个要素的数量，从而了解该街道的活力状态，并直观获悉活力点所处位置，进而为城市规划和街道设计提供指导；

2）城市色彩分析：通过对街景中的色彩进行提取和分析，可以得到城市街道的色彩图谱，从而为城市色彩设计、建筑立面设计及景观设计提供现状依据；

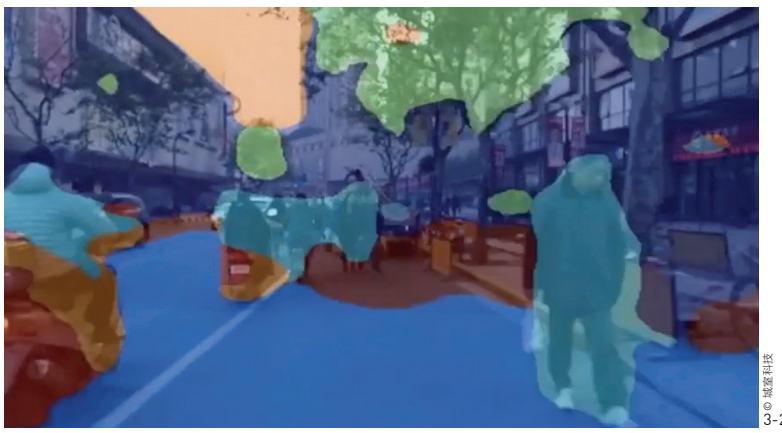
3）城市特色分析：街景特色辨析等测评系统可用于辅助城市风貌管理及设计前期调研。

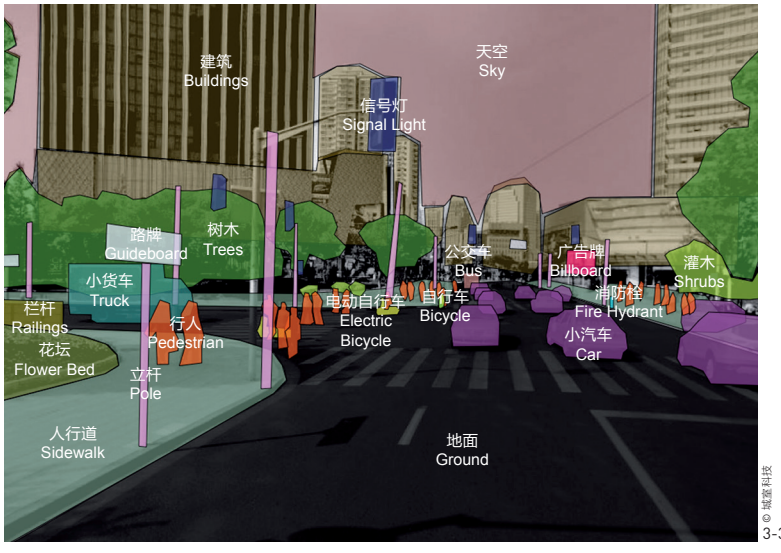
结语

不同于传统的步行导航系统，StreeTalk不仅考虑了出行时长因素，而且加入了对街道品质的考量，以为不同的出行需求提供多样化的选择。StreeTalk仅仅是城市街道感知神经网络模型的一个应用案例，通过运用物体检测识别技术、场景语义分割技术，以及卷积神经网络等，该模型可以对各种影响人们感知的要素加以分析。相关技术的应用不仅可以为城市研究及决策、城市景观设计、城市街区改善，以及整体环境品质的提升提供依据，也为未来城市生活创造了更多可能。**LAF**

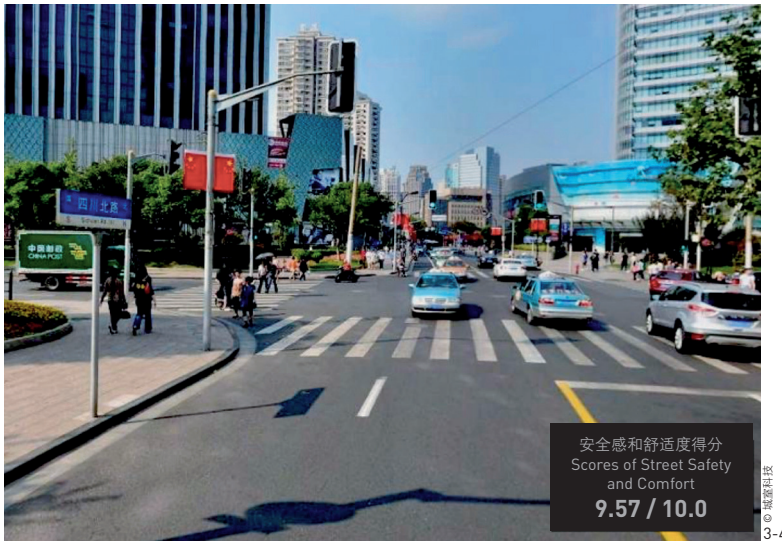


© 城窗科技 3-1

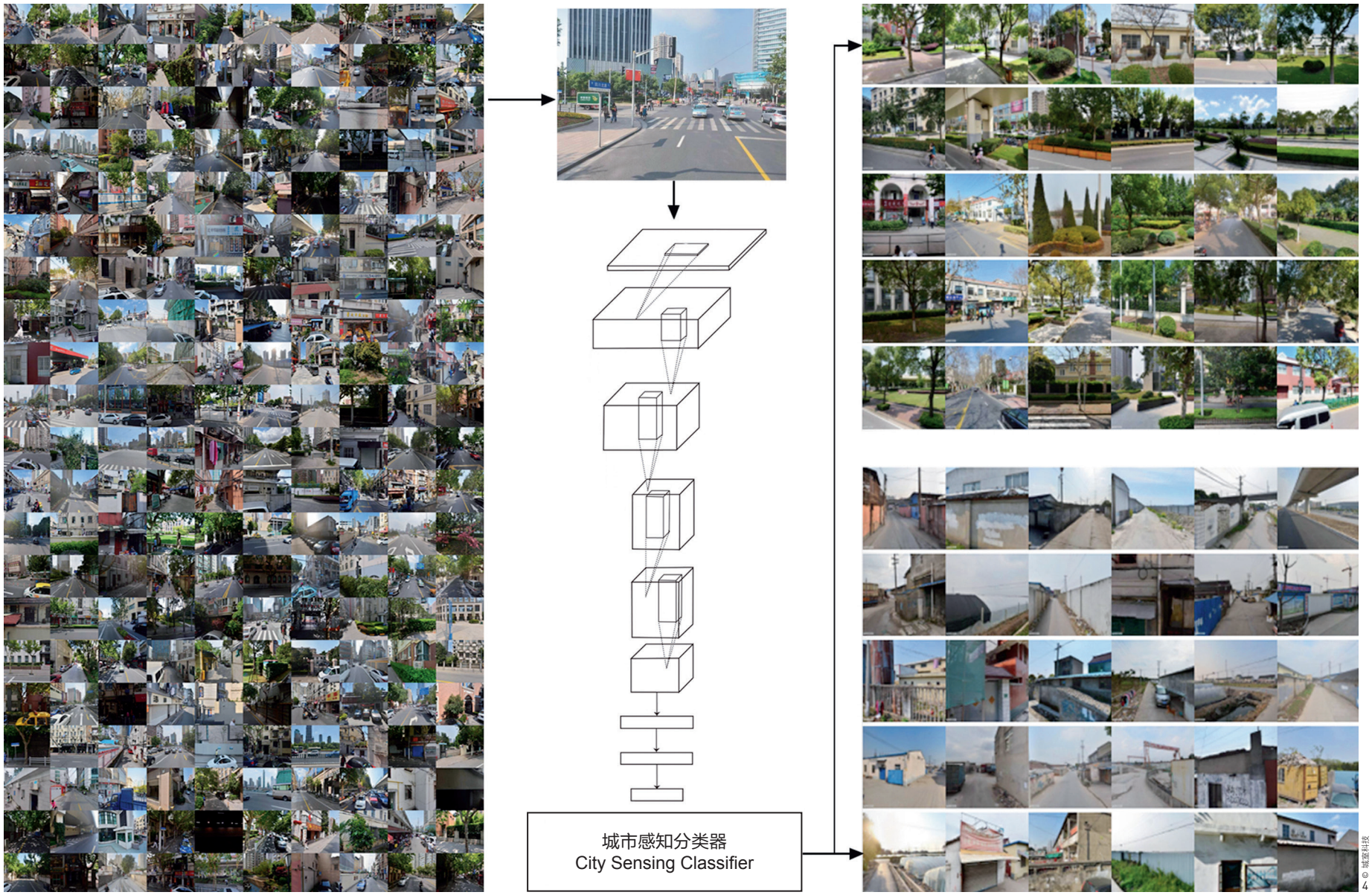

© 城窗科技 3-2



© 城窗科技 3-3



© 城窗科技 3-4



Background

At present, the urbanization rate in China has exceeded 58%[1]. As urban environments are increasingly dense and hybridized, there is a growing number of citizens tending to walk and cycle for short-distance commuting. It is commonly seen that when people go to an unfamiliar destination, they use mobile navigation systems for route guidance. However, most existing navigation systems are shortest-commuting-hour preferred, ignoring pedestrians, and cyclists' safety and comfort experiences along the routes. Thus, CitoryTech aims to develop a smart navigation system based on big data of urban streetscapes and supported by artificial intelligence cloud computing technology, to improve the quality of people's walking and cycling experience, while informing related urban research, decision-making, and public space design.

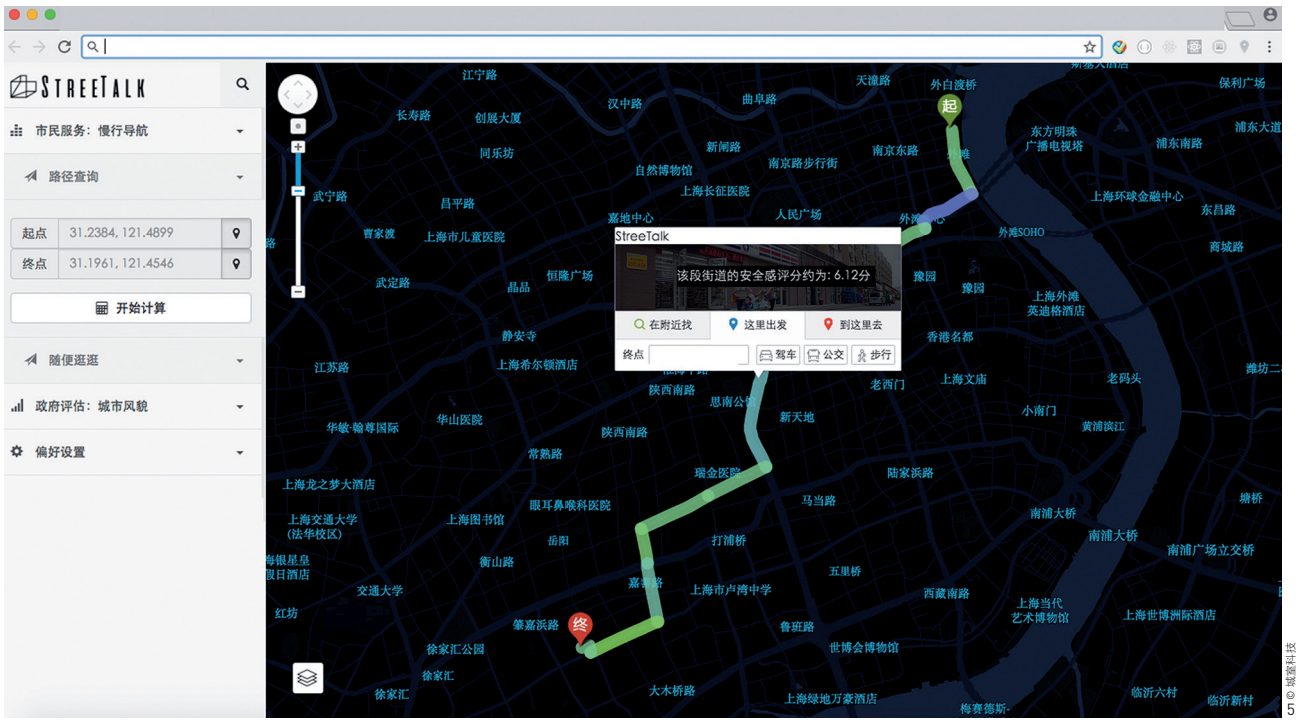
4. 通过对深度学习模型卷积神经网络（示意图）进行训练，机器可以预判人类对周边环境的感知。
5. StreeTalk慢行导航系统使用界面

4. By training convolutional neural network (schematic diagram), a deep learning model, the machine can imitate human's perception on streetscapes and score the entire street environment automatically.
5. The interface of StreeTalk navigation system

Research and Technologies

Besides the consideration of commuting hours, pedestrians and cyclists also prefer a safer and more pleasant travel experience. But, how to measure and evaluate the level of safety and comfort? Comparing street photos in pairs has been proven efficient in this very field. This method was adopted by MIT Media Lab several years ago to collect the statistics of people's evaluation on street safety and comfort level. The research team designed an online platform, through which two street photos were shown randomly each time allowing volunteer reviewers from all over the world to pick out the photo of place that looks safer and more pleasant. Based on 1.2 million-times comparisons, each photo was scored from 0 to 10 (the higher score, the higher level of street safety and comfort), and then an overall evaluation of city streetscapes was automatically generated based on the individual scores[2][3].

However, the same method is inefficient in manually scoring millions of street photos to evaluate the streetscape of metropolitans such as Shanghai. But deep learning technology can offer solutions, since Artificial Intelligence can extract relevant features with scored streetscape data input to evaluate other streets, which could greatly improve work efficiency. In view of people's perception and preference on a streetscape tends to be similar, the data drawn upon existing researches are adopted here to establish an automatically scoring system on street safety and comfort level, which allows the machine to extract image feature and to learn from each scored photograph. More than 300 categories of streetscape objects — including pedestrians, bicycles, cars, buses, street lamps, and flower beds — are extracted from existing photograph datasets through object detection technology. Meanwhile, by employing scene parsing technology, the photos are scanned at pixel-level to identify different urban scenes or domains, such as roads, sky, and buildings.



© 城家科技

By applying convolutional neural network[①], a deep learning model, the team programmed the machine to imitate human's perception on streetscapes from the dimensions of safety, aesthetics, vitality, and neatness, and score the entire street environment automatically.

In 2016, CitoryTech launched their pilot project in Shanghai, taking over 2.3 million photos of the city's streetscapes at all seasons, including suburbs. These photographs were categorized through Artificial Intelligence detection and identification. Then by combining massive data calibration and model training adjustments, CitoryTech created the world's first neural network model application in urban streetscape perception — StreeTalk navigation system for pedestrians and cyclists.

StreeTalk sets scored streetscape data as one of the properties of the city's road network, and automatically calculates the average scores of each street of each possible route, and then selects the higher-score route(s) for users, who are also allowed to view the panorama of specific streetscapes. In addition, StreeTalk is quite smart in route selecting. For example, during night hours, the route(s) with more lamps or more street shops would be recommended; while in summer, the route(s) with more tree-canopies would be recommended.

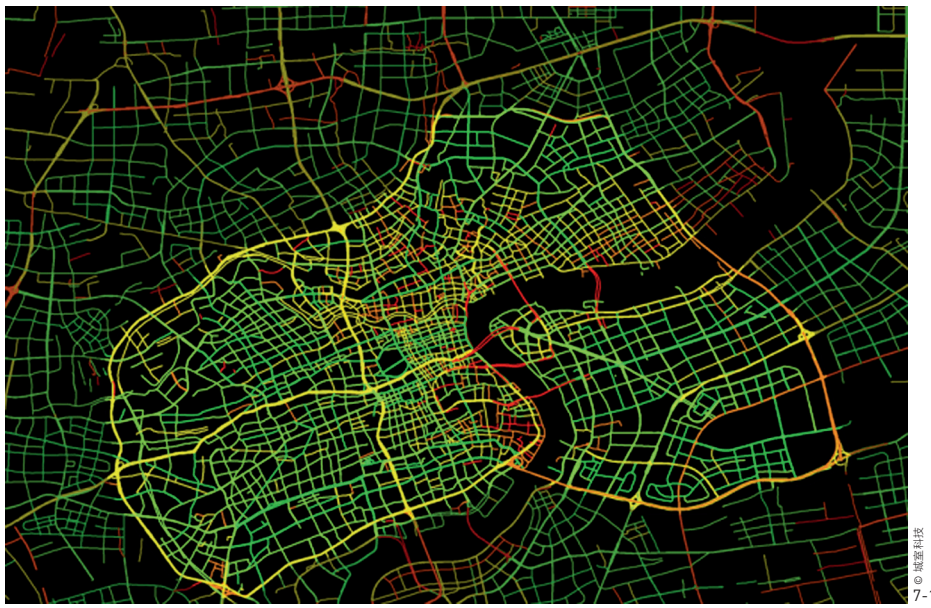
The research team found that most of the higher-score streets are rich in greenery, and clean and tidy, while the lower-score streets are often with poor landscape, lacking of public facilities and amenities. This might help city managers identify and locate the streets that need to be improved. For designers, they can leverage the system

① Convolutional neural network (CNN) learns and extracts image features layer by layer automatically. Each layer is made up of a number of feature maps, and each feature map receives several inputs, takes a weighted sum over them, passes it through an activation function, and outputs with a following layer. CNN has been proved efficient and invariant in image feature learning and representation.

6. 利用场景语义分割技术识别街景特征。
7-1. 相关技术应用：城市安全感地图
7-2. 相关技术应用：商圈热力分布图

6. Street features can be identified by scene parsing technology.
7-1. Relevant technology application: city security map
7-2. Relevant technology application: heat map of commercial districts

© 城室科技
6



7-1

7-2

to learn the streetscape features of higher-scored streets to further improve their design on urban spaces. Moreover, StreeTalk could record and analyze slow traffic information, including the data of jogging, walking, and cycling activities in the city, and support decision-makings for urban planning and design of urban greenways.

Other Application Prospects

Apart from being applied as navigation system for pedestrians and cyclists, technologies of object detection, scene parsing, as well as convolutional neural network, can provide methods for urban planners and landscape designers to enhance efficiency and to optimize design procedure:

1) Heat map analysis: Through object detection technology, pedestrians, vehicles, and various facilities and amenities in the streets are identified and statistically analyzed, which makes it convenient for urban designers to obtain the related information of dynamic flow and viability in local streets and neighborhoods, providing guidance for urban planning and streetscape design;

2) City color-spectrum analysis: By extracting and analyzing the colors in the streets, the city color-spectrum map could provide a basis for overall color design, building facades design, and landscape design;

3) Urban identity and streetscape quality analysis: Streetscape identity analysis and evaluation system could inform urban landscape management and pre-design research, and further inform decision-makings on urban identity and streetscape design.

Conclusion

Different from conventional efficiency-driven navigation systems, StreeTalk not only considers commuting hours, but also aims to provide diverse options for different travel requirements. By using object detection technology, scene parsing technology, and convolutional neural network, StreeTalk becomes an example of applying neural network model on urban streetscape perception. Relevant applications can provide bases for urban research and decision-making, urban landscape design, neighborhood renewal, and city environmental quality improvement, while offering more possibilities for a better future of urban life. **LAF**

REFERENCES

[1] National Bureau of Statistics of the People's Republic of China. (2018). National Economic and Social Development Statistical Bulletin of the People's Republic of China (2017). Retrieved from http://www.stats.gov.cn/tjsj/zxfb/201802/t20180228_1585631.html

[2] Dubey, A., Naik, N., Parikh, D., Raskar, R., & Hidalgo, C. A. (2016). Deep Learning the City: Quantifying Urban Perception at a Global Scale. Computer Vision — ECCV 2016 (pp. 196-212). Netherlands: Springer.

[3] Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The Collaborative Image of the City: Mapping the Inequality of Urban Perception. Plos One, 8(7), 1-12.