



HOKKAIDO UNIVERSITY

Title	A Study on Public Safety Prediction using Satellite Imagery and Open Data [an abstract of dissertation and a summary of dissertation review]
Author(s)	Najjar, Al-ameen
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第12644号
Issue Date	2017-03-23
Doc URL	https://hdl.handle.net/2115/65764
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	doctoral thesis
File Information	Alameen_Najjar_abstract.pdf, 論文内容の要旨



学 位 論 文 内 容 の 要 旨

博士の専攻分野の名称 博士（情報科学） 氏名 Al-ameen Najjar

学 位 論 文 題 名

A Study on Public Safety Prediction using Satellite Imagery and Open Data
(衛星画像およびオープンデータを利用する公共安全性予測に関する研究)

Data-driven mapping is critical for the sustainable development of cities. Maps visualize patterns and trends about cities that are difficult to spot in data otherwise. For example, a road-safety map made from several years' worth of traffic-accident reports pinpoints roads and highways vulnerable to accidents. Similarly, an accurate crime map clearly highlights neighborhoods where criminal activities abound. Such insights are highly needed to inform sustainable city-planning decision-making and policy. Therefore, public-safety mapping is crucial for urban planning and development worldwide. Accurate mapping requires accurate longitudinal data collection, which is both highly expensive and labor intensive. Data collection is manual and requires skilled enumerators to conduct. Therefore, generating city-scale public-safety maps is beyond affordable to countries that lack data or the means required for collecting it, such as low- and middle-income countries. Therefore, taking manual data collection out of the equation will quicken the mapping process in general, and make it possible where data does not exist or collecting it is not affordable.

Recent advances in imaging and space technology have made high-resolution satellite imagery increasingly abundant and affordable. Satellite imagery is a rich medium of visual cues relevant to environmental, social, and economic aspects of urban development. Moreover, given the recent breakthroughs made in the field of computer vision and pattern recognition, it is logical to attempt predicting public safety at a city scale directly from satellite imagery. In other words, investigating an automatic approach to public-safety mapping based on the assumption that visual information captured in satellite imagery can be used as a proxy indicator of public safety.

In this study, we discuss our approach to city-scale public safety prediction directly from satellite imagery using tools from modern machine learning and computer vision. In particular, we focus on two types of public safety maps, namely road safety maps and crime maps. We re-defined the problem of public safety mapping as a supervised image classification problem, in which a city-scale satellite map is treated as a set of raw high-resolution satellite images each of which is assigned a safety label predicted using a model learned from a separate set of training data. To obtain training data we have leveraged official police reports collected by police departments, and released as open data. We validated and tested the robustness of the learned models for both road safety and crime rate prediction tasks over four different US cities, namely New York, Chicago, San Francisco, and Denver. We also investigated the reusability of the learned models across different cities of the same group.

This thesis consists of 5 chapters. Chapter 1 discusses the motivation of the study, introduces public safety mapping, describes the contributions we made in this study, and explains how the thesis is organized. Chapter 2 overviews modern image classification and discusses its main challenges. This

chapter explains the basics of supervised image classification focusing on both flat and deep classification architectures. Chapters 3 and 4 contain a detailed explanation of the contributions we make in this thesis. In Chapter 3, we detail our proposed approach for obtaining high-resolution satellite images labeled with public safety (road-safety level, and crime rate). To this end, we leverage open data in the form of official police reports manually collected by several police departments across the United States. We used crime incident reports, and traffic accident reports released by four police departments, namely New York Police Department (NYPD), Chicago Police Department (CPD), Denver Police Department (DPD), and San Francisco Police Department (SFPD). We used these reports to produce five datasets (two road safety datasets, and three crime datasets) of labeled satellite images on which our models are verified, tested and trained. Our approach to obtain labeled satellite images from official police reports is straightforward. We start by discretizing the location information contained in the reports. This is done via dividing the city map into square regions using a square grid. Each region is then assigned a safety score calculated as a sum of all incidents reported within its boundaries. These scores are then clustered by frequency around three bins representing three different safety labels (Low, neutral, and high safety). In order to avoid learning a biased model, the obtained classes are resampled so that each has the same number of instances. At this point, police reports are transformed into a set of square regions each labeled with a corresponding safety label. Finally, each region is represented with a square satellite image centered at its center location. Chapter 4 introduces our framework to public safety mapping from satellite imagery and open data. Given a source and a target cities, our goal is to predict for the target city a city-scale public-safety map directly from its satellite map (raw satellite imagery) using a model learned from the source city's data represented in its satellite map and its official police reports (open data). For prediction, we have used both flat and deep models trained on labeled satellite imagery. Empirical evaluation on four different US cities showed that visual information captured in satellite imagery can be effectively used as a proxy indicator of road safety and crime rate. Moreover, the learned models are reasonably reusable across different cities. Finally, in Chapter 5 we conclude our study, and discuss different scenarios to extend our work to be more suited for cities in the developing world.