

São Paulo City Homicides and Their Surroundings: from Non-Negative Tensor Factorization to Pattern Identification

Jaqueline Alvarenga Silveira¹, Sérgio Adorno², Marcelo Nery³, Germain García⁴, Afonso Paiva⁵, Luis Gustavo Nonato⁶

Universidade de São Paulo, USP, Brasil

alva.jaque@usp.br¹, sadorno@usp.br², mbnery@usp.br³, germain.garcia@ucsp.edu.pe⁴, apneto@icmc.usp.br⁵, gnonato@icmc.usp.br⁶

Abstract

This study aims to detect patterns around intentional homicide in the city of São Paulo, aiming to understand the relation between criminal records and the different variables involved in the analysis. The study makes use of different data sources and a combination of Non-Negative Tucker Tensor Decomposition and hierarchical clustering to extract patterns associated with homicides. Moreover, a visualization tool, able to handle spatiotemporal data was developed to support the visual analysis of the founded patterns. The visualization tool revealed a relationship between homicides and the involved variables.

Introduction

The problem of homicide affects several places and has been identified as a complex event characterized by variations in behavioral style, level of violence, and means used during the action, motive, and personal interaction. Thus, understanding the urban around homicides can be educational for the community in general due to its implications for the prevention and reduction of violent crimes. The aim of this study is to identify patterns potentially related to homicides in São Paulo, aiming to understand the relationship between homicides and external variables involved in the analysis. To do so, we make use of different data sources and the Non-negative Tucker Tensor Decomposition combined with hierarchical clustering to extract homicide related patterns.

Methodology

Tensor decompositions and factorizations is natural modeling mechanism to identify patterns from various data sources, allowing us to capture multi-aspects and multi-linear structures in lower dimensions. In our application, a specific group of records is summarized by getting the most prevalent occurrences inside the given area. Thus, the tensor resulting from these aggregations is non-negative.

The methodology used in this study for the detection of patterns in the surroundings of a site is based on the decomposition of non-negative tensors and comprises three main stages (see Fig. 1): tensor modeling (according to the indexes generated as discussed in the previous sections), extraction of patterns and visual representation.

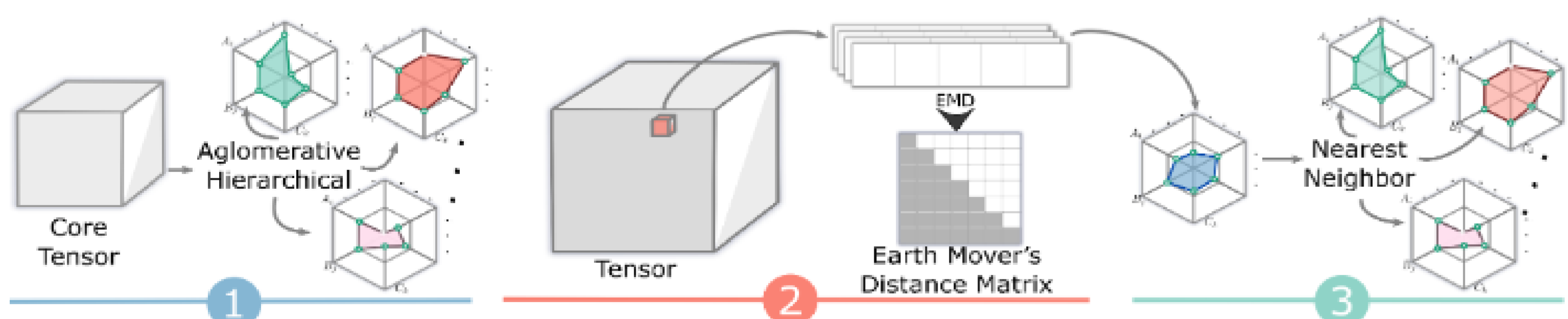


Fig. 1: The pattern approximation process: the patterns returned by the tensor core are clustered (1), a signature is created for each pattern within the tensor (2), and the remaining patterns are grouped within the most relevant ones (3).

Results

Results showed homicides seem to be related to the presence of Cortiços (a large building divided into apartments, usually in a poor area of a city) and trees, Fig. 2 and Fig. 3. The improvement of the urban infrastructure seems to reduce the number of homicides Fig. 4 (a) and (b).

Acknowledgments

This work was supported by São Paulo Research Foundation (FAPESP) - Brazil. The views expressed are those of the authors and do not reflect the official policy or position of the São Paulo Research Foundation.

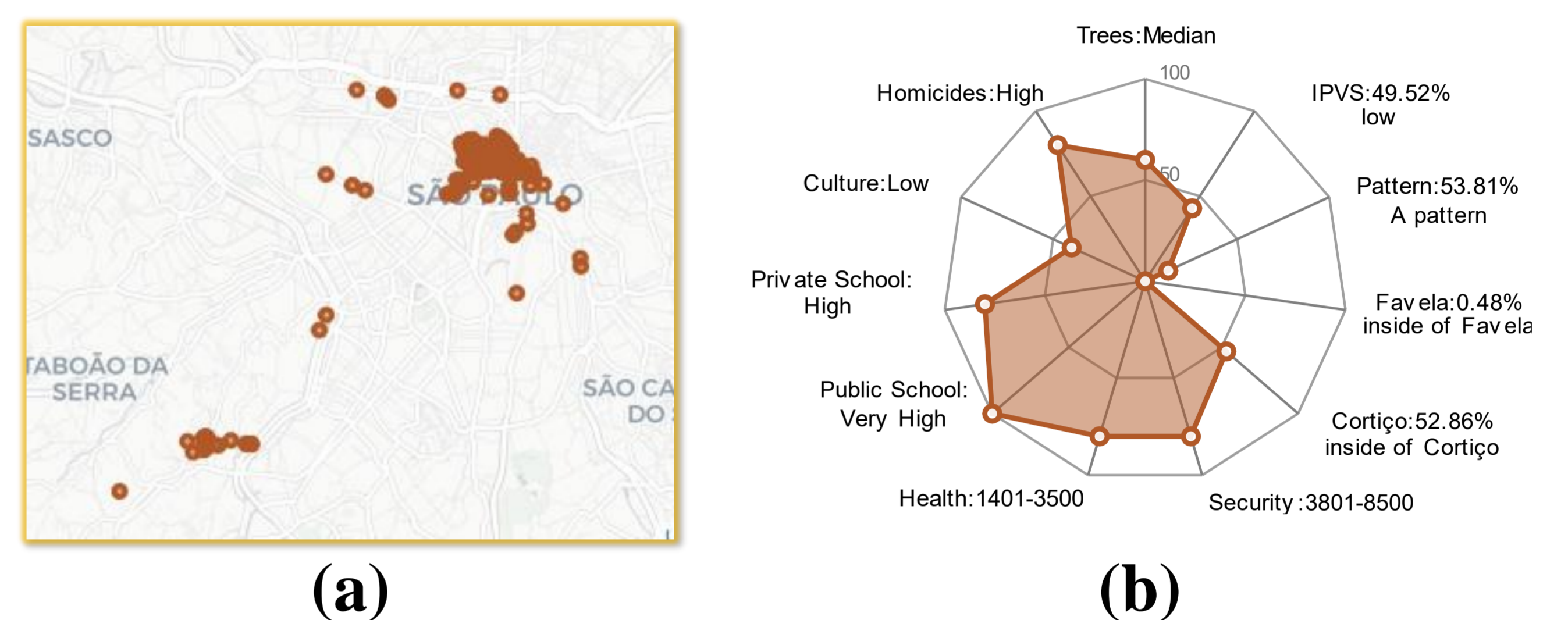


Fig. 2: According to the pattern in (b), there is a high concentration of homicides in the Cortiços.

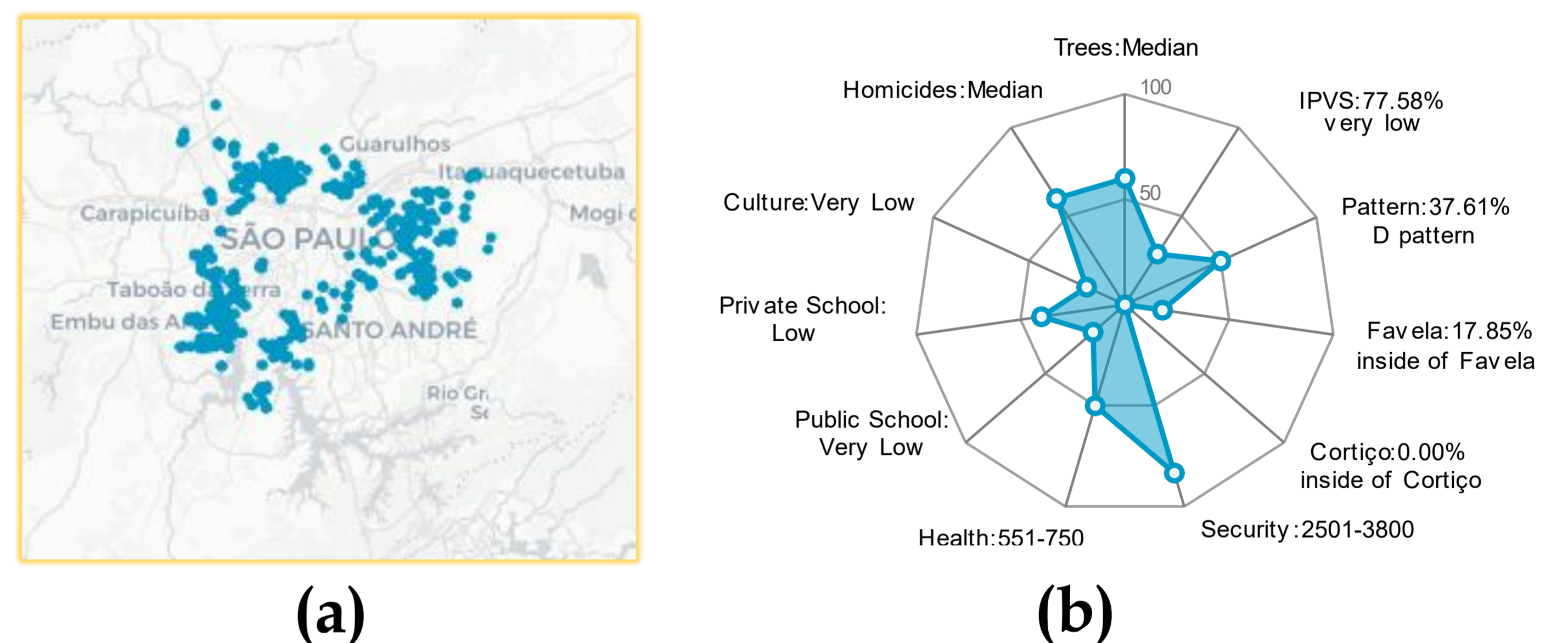


Fig. 3: According to the pattern in (b), there is a relationship between trees and homicides.

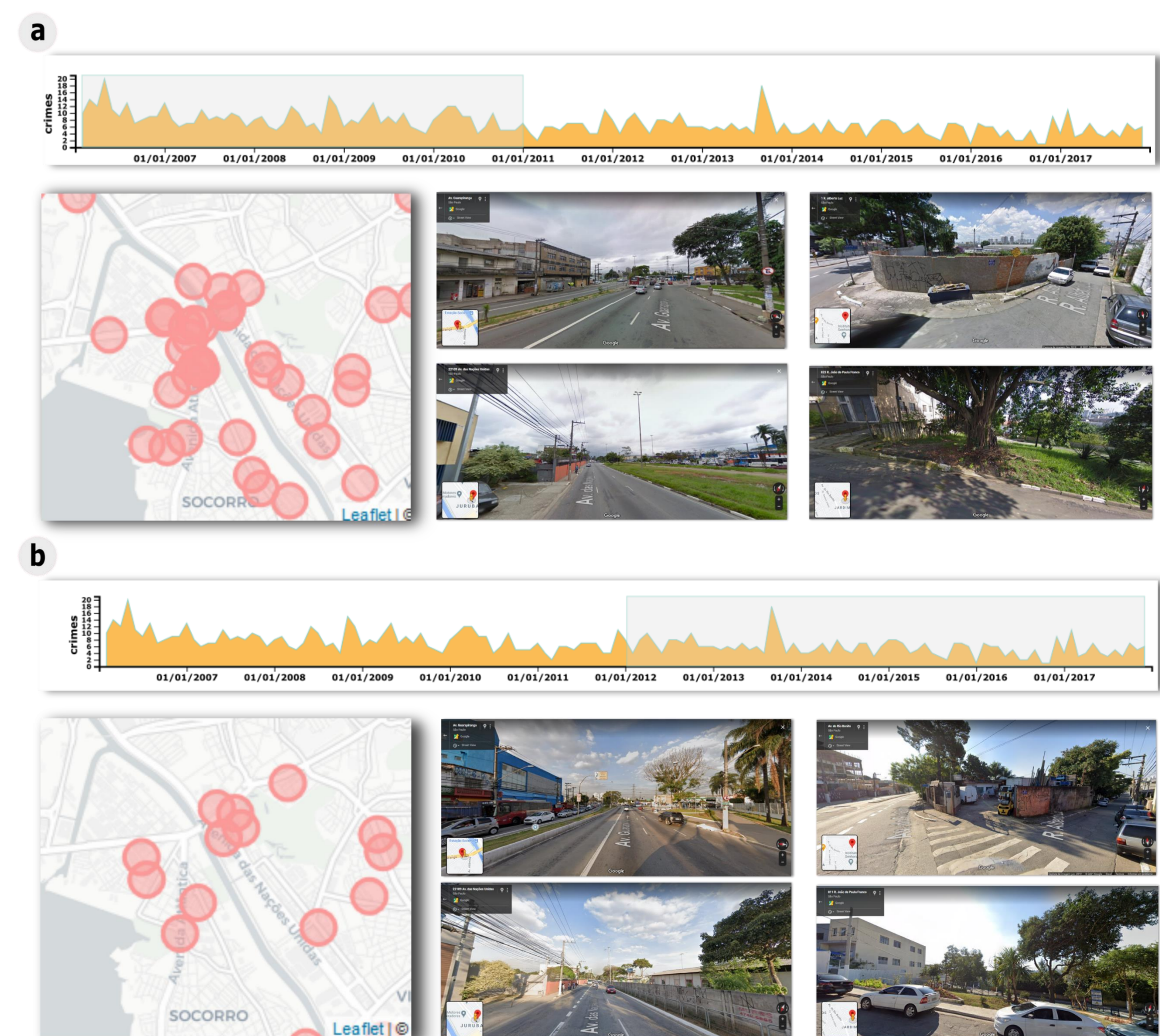


Fig. 4: The improvement of infrastructure seems to influence homicide decrease.

Conclusions

The results showed that there is a relationship between intentional homicides and some variables of the patterns (presence of trees and Cortiços). Also, the improvement of infrastructure seems to decrease the homicide rate. Therefore, the tensor decomposition can be applied to different data sources and different contexts.