Mathematical and Computational Applications

MDPI

Article

Detection of Abnormal Pedestrian Flows with Automatic Contextualization Using Pre-Trained YOLO11n

Adri n Núñez-Vieyra, Juan C. Olivares-Rojas *, Rogelio Ferreira-Escutia, Arturo Méndez-Patiño, José A. Gutiérrez-Gnecchi and Enrique Reyes-Archundia

National Technological Institute of Mexico (TecNM), Technological Institute of Morelia (ITM), Morelia 58120, Mexico; adrian.nv@morelia.tecnm.mx (A.N.-V.); rogelio.fe@morelia.tecnm.mx (R.F.-E.); arturo.mp@morelia.tecnm.mx (A.M.-P.); jose.gg3@morelia.tecnm.mx (J.A.G.-G.); enrique.ra@morelia.tecnm.mx (E.R.-A.)

* Correspondence: juan.or@morelia.tecnm.mx; Tel.: +52-443-181-3542

Abstract: Recently, video surveillance systems have evolved from expensive, human-operated monitoring systems that were only useful after the crime was committed to systems that monitor 24/7, in real time, and with less and less human involvement. This is partly due to the use of smart cameras, the improvement of the Internet, and AI-based algorithms that allow the classifying and tracking of objects in images and in some cases identifying them as threats. Threats are often associated with abnormal or unexpected situations such as the presence of unauthorized persons in a given place or time, the manifestation of a different behavior by one or more persons compared to the behavior of the majority, or simply an unexpected number of people in the place, which depends largely on the available information of their context, i.e., place, date, and time of capture. In this work, we propose a model to automatically contextualize video capture scenarios, generating data such as location, date, time, and flow of people in the scene. A strategy to measure the accuracy of the data generated for such contextualization is also proposed. The pre-trained YOLO11n algorithm and the Bot-SORT algorithm gave the best results in person detection and tracking, respectively.

Keywords: automatic contextualization; Bot-SORT; confusion matrix; pedestrian flow; tracking; video surveillance; YOLO; YOLO11n

check for updates

Academic Editors: Mario Graff and H ctor-Gabriel Acosta-Mesa

Received: 3 February 2025
Revised: 1 April 2025
Accepted: 15 April 2025
Published: 17 April 2025

Citation: Núñez-Vieyra, A.; Olivares-Rojas, J.C.; Ferreira-Escutia, R.; M ndez-Patiño, A.; Guti rrez-Gnecchi, J.A.; Reyes-Archundia, E. Detection of Abnormal Pedestrian Flows with Automatic Contextualization Using Pre-Trained YOLO11n. *Math. Comput. Appl.* **2025**, *30*, 44. https://doi.org/10.3390/mca30020044

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In recent years, video surveillance systems have become an important support in the search for the clarification of all types of crimes. According to [1], in Mexico, some crimes, such as street robbery, vehicle theft, and home burglary, have decreased in recent years, a situation that coincides with the growth in the installation of video surveillance systems.

A typical video surveillance system is defined as a technological tool based on the use of video cameras to support actions such as police deployment, emergency response, crime prevention, and the administration of justice [2]. That is to say, historically, the use of video surveillance systems has been solely to respond to events that indicate a crime has been committed.

However, over the past two decades, video surveillance systems have evolved from being generators of extensive databases, often useful only after a crime has been committed, to becoming intelligent systems capable of detecting and identifying threats, tracking them through multiple cameras, and generating summaries for use by law enforcement authorities. To achieve this automatic understanding, video surveillance systems must go