

ZebrAI Crossing: Detecció de passos de zebra i semàfors de vianants

Albert Capdevila Estadella (1587933), Levon Kesoyan Galstyan (1668018), Luis Martínez Zamora (1668180)

1 INTRODUCCIÓ

ZebrAI Crossing és un model de visió per computador que, donada una imatge en un entorn urbà, detecta si hi apareix un pas de zebra i en determina la seva orientació i posició respecte la càmera. A més a més, també detecta la presència de semàfors per a vianants i indica el seu estat (en vermell, en verd...).

Aquest és un model pensat per al desenvolupament d'OrionWay, un robot autònom que guia persones amb discapacitat visual per entorns urbans. OrionWay és un projecte de l'assignatura Robòtica, Llenguatge i Planificació del Grau en Enginyeria Informàtica a la Universitat Autònoma de Barcelona.

1.1 Objectius

Donada una imatge, volem:

- Identificar si hi apareix un pas de zebra.
En cas afirmatiu:
 - Determinar l'orientació del pas de zebra.
 - Determinar la posició de l'inici del pas de zebra.
- Identificar si hi apareix un semàfor de vianants.

En cas afirmatiu:

- Indicar si el semàfor es troba en verd.

Així doncs, aquests objectius seran útils per a posicionar i orientar el robot OrionWay quan apareguin passos de zebra, i només creuar-los quan el semàfor estigui en verd.

2 ESTAT DE L'ART

En el punt de partida d'aquest projecte, hem pogut trobar diversos repositoris útils i articles existents que tracten de resoldre un problema similar. Aquests són els recursos que hem investigat que ens han resultat més útils:

2.1 Projectes

ImVisible [1]

El projecte tracta del desenvolupament d'una aplicació mòbil que mitjançant la càmera del dispositiu analitza l'entorn per tal de detectar passos de zebra, la seva orientació i, en cas que estiguin regulats per un semàfor, detectar-ho i saber si es troba en verd o en vermell. Tot el procés de reconeixement d'imatges es fa mitjançant xarxes neuronals convolucionals.

Crosswalks-detection-using-YOLO [2]

El projecte tracta del desenvolupament d'un detector de passos de zebra utilitzant YOLO, ens serà útil la resolució de problemes que s'explica al projecte, tant la unificació de les diferents deteccions que pot fer el programa a un mateix pas de zebra com la part de tenir diversitat d'exemples d'altres marques de circulació que es troben a la carretera.

Real-time detection of people and bicycles in pedestrian crossings - CVC [3]

El projecte tracta d'una detecció instantània mitjançant una única càmera de vianants i ciclistes a passos de zebra, amb l'objectiu de poder avisar als conductors que s'apropen a l'àrea d'encreuament i altres com la necessitat de control policial, una mala configuració de semàfors...

PedestrianTrafficLightsAndCrosswalkDetection [4]

El projecte tracta de la creació d'una xarxa neuronal convolucional per la identificació en temps real de passos de zebra i semàfors, tenint en compte l'estat d'aquest últim.

2.2 Articles

Zebra Crossing Detection and Time Scheduling Accuracy, Enhancement Optimization Using Artificial Intelligence [5]

Els autors proposen un sistema de detecció de passos de zebra orientat a ajudes electròniques per a persones amb discapacitat visual. La solució divideix la imatge en finestres de tamany fix que es processen seqüencialment amb un classificador de regressió logística entrenat sobre features extrems per una CNN (anomenada LYTNNet). Un segon model de regressió prediu l'orientació del pas. Amb aquesta estratègia milloren la corba precision-recall respecte a enfocaments basats només en transformades de Hough i redueixen l'RMSE de l'angle estimat. No obstant això, l'enfocament patch-based incrementa el temps d'inferència i no contempla la detecció simultània de semàfors, de manera que caldria un mòdul addicional per als objectius de ZebrAI Crossing.

Zebra-crossing Detection for the Partially Sighted [6]

Aquest article relata el procés de fer un algorisme de detecció de passos de zebra, amb l'objectiu de ser útil per a Technological Aid aimed at helping Partially Sighted (TAPS). Utilitza transformacions de Hough i detecció de variacions d'intensitat per a identificar les línies paral·leles que defineixen el pas de zebra de manera precisa. A més, per aplicar restriccions que permetin diferenciar els passos de zebra de les imatges d'escales, que visualment són similars, es proven tres mètodes: la cerca d'una homografia, el càlcul de la normal mitjançant la línia de fuga obtinguda a partir de línies equidistants, i l'ús de dos punts de fuga. En tots els casos, el pendent (slope) de les escales permet diferenciar-les dels passos de zebra.

ZebraRecognizer: efficient and precise localization of pedestrian crossings. [7]

ZebraRecognizer està pensat per a dispositius mòbils i combina la càmera amb l'acceleròmetre per rectificar la perspectiva abans de detectar. L'algoritme segueix quatre etapes: (i) pre-processat amb suavitzat adaptatiu, (ii) càlcul de la matriu de rectificació mitjançant sensors inercials, (iii) detecció de segments de línia amb EDLines i (iv) agrupació + validació de candidats segons solapament i consistència de color. Els autors reporten 100% de precisió amb recall 0,78 i 12 ms per imatge en escriptori (uns 32 ms en mòbil).

Aquest article encaixa molt bé amb el nostre enfocament sense deep learning perquè, primerament, la rectificació geomètrica

amb els sensors del mòbil és una manera senzilla i eficaç de corregir la perspectiva abans de buscar patrons i, a més, EDLines dona línies clares i sòlides que ens ajuden a validar bé la forma dels rectangles alterns del pas de zebra.

Zebra-crossing detection based on cascaded Hough transform principle and vanishing point characteristics [8]

L'article presenta un estudi per detectar passos de zebra utilitzant un filtratge de Gauss, detecció de vores de Canny i una transformada en cascada de Hough basada en el punt de fuga, sense extreure regions d'interès. També es compara amb la transformada de Hough estàndard, demostrant que la tècnica en cascada obté millors resultats.

Advanced Zebra Crosswalk Detection Using Deep Learning Techniques for Smart Transportation Systems [9]

Aquest estudi amplia el problema a un escenari de trànsit intel·ligent: utilitza SSD amb transfer-learning sobre un conjunt de dades de nou classes (pas de zebra, vianants, vehicles, etc.). Demostra altes precisions i recall en temps real i planteja la seva integració en dispositius portàtils.

Encara que el cor és deep learning, la descripció inclou avaluació en temps real amb FPS i confusion matrix. La part de xarxa no és d'interès per a nosaltres, però la metodologia experimental és un bon model de validació.

ZebraRecognizer: Pedestrian crossing recognition for people with visual impairment or blindness [10]

L'article tracta de la creació d'una aplicació mòbil que té l'objectiu de detectar passos de zebra en temps real el qual informa a l'invident mitjançant instruccions d'àudio on s'obté la informació del voltant de la persona a partir de les característiques del mòbil com l'altura, la inclinació, zoom i resolució de la càmera, etc... A la imatge obtinguda primerament es transforma en escala de grisos i s'aplica un suavitzat per reduir soroll, on posteriorment mitjançant una homografia on "desprojecta la imatge" per tal de que es trobi en un pla i així es treballi amb qualitats reals. Posteriorment apliquen una detecció de vores, utilitzant regressió ortogonal per definir una única línia comuna.

2.3 Datasets

Roboflow (july_6, capstone-for-detection, capstone-for-detection1) [11, 12, 13]

Els tres datasets, proporcionats pel mateix usuari, contenen un gran nombre d'imatges de passos de vianants i primers plans de llums de semàfors per a vianants. Cada imatge està classificada en tres categories possibles: *green_pedestrian_light*, *red_pedestrian_light* i *zebra*, podent pertànyer a més d'una classe simultàniament. A més, per a cada classe detectada es proporciona una regió rectangular que delimita l'element corresponent (Veure Fig. 1a).

El gran desavantatge d'aquest conjunt de datasets és que la regió rectangular no indica l'orientació dels passos de zebra i sovint els delimita de manera recta i imprecisa (veure Fig. 1b).



(a) Múltiples classes

(b) Regió rectangular

Fig. 1: Exemples d'imatges dels datasets [11, 12, 13]

Roboflow - crosswalk [14]

Aquest dataset és semblant als anteriors, però només inclou passos de zebra, amb regions una mica més precises. El seu principal inconvenient és que té molt poques imatges (261).

ImVisible [1]

El projecte ImVisible utilitza un dataset que conté imatges de passos de zebra amb semàfors per a vianants en diverses condicions, com ara dia, nit i pluja. Les imatges tenen les anotacions: [file_name, class, x1, y1, x2, y2, blocked tag], on class indica el tipus de llum del semàfor de vianants, x1, y1, x2, y2 indiquen l'inici i el final de la línia perpendicular al pas de zebra en el seu punt mig, i blocked tag indica si el semàfor per a vianants està obstruït visualment en la imatge. L'autor del projecte proporciona les imatges del dataset en tres resolucions diferents (768x576, 876x657, 4032x3024).



Fig. 2: Exemples d'imatges de ImVisible

3 PROPOSTA

3.1 Descripció de les dades

En aquest projecte, utilitzarem un conjunt de dades fabricat a partir de dos datasets diferents trobats en repositoris de GitHub: **ImVisible [1]** i **GlobalStreetscapes [15]**.

Aquest conjunt de dades té dos propòsits:

1. Avaluar el rendiment de la part del model encarregada d'identificar els passos de zebra i els seus atributs.
2. Entrenar, validar i provar la part del model encarregada d'identificar l'estat dels semàfors per a vianants.

Hem triat el dataset d'ImVisible perquè proporciona una gran quantitat d'imatges anotades amb informació que ens permet extreure la orientació del pas de zebra, la posició de l'inici d'aquest i l'estat del semàfor de vianants.

Ara bé, per a fer una bona avaluació, també necessitarem imatges que no continguin passos de zebra, motiu pel qual hem decidit utilitzar també imatges de GlobalStreetscapes. Dins d'aquest immens conjunt de dades, hem seleccionat imatges amb els atributs `platform = "walking surface"`, `panoramic status = false`, i `view direction = "front"` per assegurar-nos que siguin similars a les de l'altre dataset.

Així doncs, hem incorporat un nou atribut als existents en el dataset d'ImVisible (Explicació a l'apartat 2.3). Aquesta mesura l'orientació del pas de zebra en relació amb una trajectòria recta de creuament, i es calcula a partir de la recta definida per les coordenades (x_1, y_1) i (x_2, y_2) :

$$\theta = \tan^{-1} \left(\frac{x_2 - x_1}{y_2 - y_1} \right)$$

En la aquesta figura es pot veure un exemple:

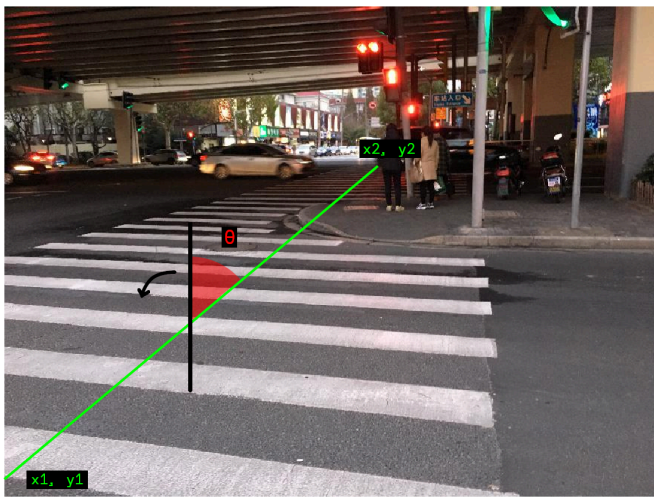


Fig. 3: Angle d'orientació del pas de zebra

En aquest cas, farà falta girar la càmera $-\theta$ per poder creuar el pas de zebra avançant cap endavant.¹

A més, considerarem les coordenades $((x_1, y_1) \circ (x_2, y_2))$ més properes a la càmera com el punt d'inici del pas de zebra:

$$(x, y) = (x_i, y_i : y_i > y_{3-i})$$

En l'exemple anterior, $x=x_1$ i $y=y_1$.

En conclusió, els atributs finals que té cada imatge del dataset que hem construït són: `[file.name, x, y, theta.rad, theta.deg, class, blocked tag]`. **Recordatori:** `class` indica el tipus de llum del semàfor de vianants i `blocked tag` indica si el semàfor per a vianants està obstruït visualment en la imatge.

3.2 Tècniques a utilitzar

Hem pensat que aquest projecte combinarà dues metodologies: **tècniques tradicionals** per a detectar els passos de zebra i els seus atributs, i **tècniques basades** en Deep Learning per a la detecció dels semàfors i el seu estat.

Així doncs, per a la tasca de detecció dels passos de zebra, algunes de les tècniques considerades són:

- Detecció de vores (Laplacià, Canny,...)

- Filtratge Gaussià
- Morfologia matemàtica
- Transformades de Hough / CHT
- Homografies per detectar *slope*
- Detecció de variacions de color / intensitat

En canvi, per a la detecció dels semàfors i el seu estat crearem una **xarxa neuronal (convolucional)** o bé investigarem alguna eina existent com ara **YOLO**. La idea és que utilitzarem el dataset fabricat per validar i decidir el millor model.

REFERÈNCIES

Projectes

- [1] S. Yu et al., *ImVisible: Pedestrian Traffic Light Dataset, Lyt-Net Neural Network, and Mobile Application for the Visually Impaired*. GitHub repository.
- [2] N. Kuznetsov, *Crosswalks Detection using YOLO: A Supervised Method for Detecting Pedestrian Crosswalks*. GitHub repository.
- [3] CVC UAB, *Real-time detection of people and bicycles in pedestrian crossings*.
- [4] R. Santos, *PedestrianTrafficLightsAndCrosswalkDetection*. GitHub repository.

Articles

- [5] M. A. Rahman et al., *Zebra Crossing Detection and Time Scheduling Accuracy Enhancement Optimization Using Artificial Intelligence*. ResearchGate, 2021.
- [6] D. Bradley and A. Dunlop, *Zebra Crossing Detection for the Partially Sighted*.
- [7] D. Ahmetovic et al., *ZebraRecognizer: efficient and precise localization of pedestrian crossings*. 2014.
- [8] J. Zhang and L. Wang, *Zebra-crossing detection based on cascaded Hough transform principle and vanishing point characteristics*. De Gruyter, 2022.
- [9] P. Sharma and R. K. Singh, *Zebra Crossing Detection using Image Processing Techniques*. IJESSE, 2022.
- [10] Y. Chen et al., *ZebraRecognizer: Pedestrian crossing recognition for people with visual impairment or blindness*. Pattern Recognition, 2016.

Datasets

- [11] Dkdkd. *july_6 Dataset*. Roboflow Universe, juny 2024.
- [12] Dkdkd. *capstone for detection Dataset*. Roboflow Universe, juny 2024.
- [13] Dkdkd. *capstone for detection1 Dataset*. Roboflow Universe, juny 2024.
- [14] Esera. *crosswalk Dataset*. Roboflow Universe, agost 2024.
- [15] UALSG. *Global Streetscapes Dataset for Crosswalk Detection*. GitHub repository.

¹És una simplificació que no té en compte la distorsió provocada pels punts de fuga.