

Redes neurais convolucionais para detección de obxectos en entornos subterráneos

Pablo Santana González

Enrique Paz Domonte*

George Nikolakopoulou†

Christoforos Kanellakis†



Escola de
Enxeñaría
Industrial

Universidade de Vigo

Traballo de fin de Mestrado

Mestrado en Enxeñaría Industrial

Universidade de Vigo *, Luleå Tekniska Universitet †



Contidos

- Introducción e obxectivos
- Deteción de obxectos e redes neurales
- Localización en robótica
- Demostración
- Conclusións

Introducción

- Exploración de zonas de risco con vehículo non tripulado
- Entornos subterráneos: cerrados, sen luz natural e de acceso ou espazo restrinxido
- Minas, cavernas, estacións de metro, sótanos, etc.



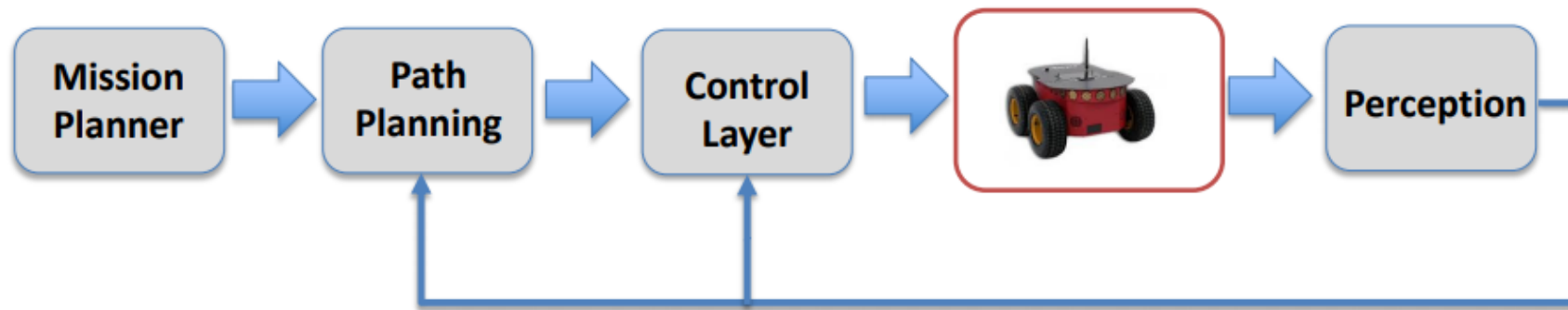
Introducción

- Exploración SubT autónoma é fronte de investigación no mundo da robótica
- DARPA SubT Challenge:
 - Busca, detección e localización dun set de artefactos representativos de actividade humana
 - *Tunnel Systems, Urban Underground, and Cave Networks*



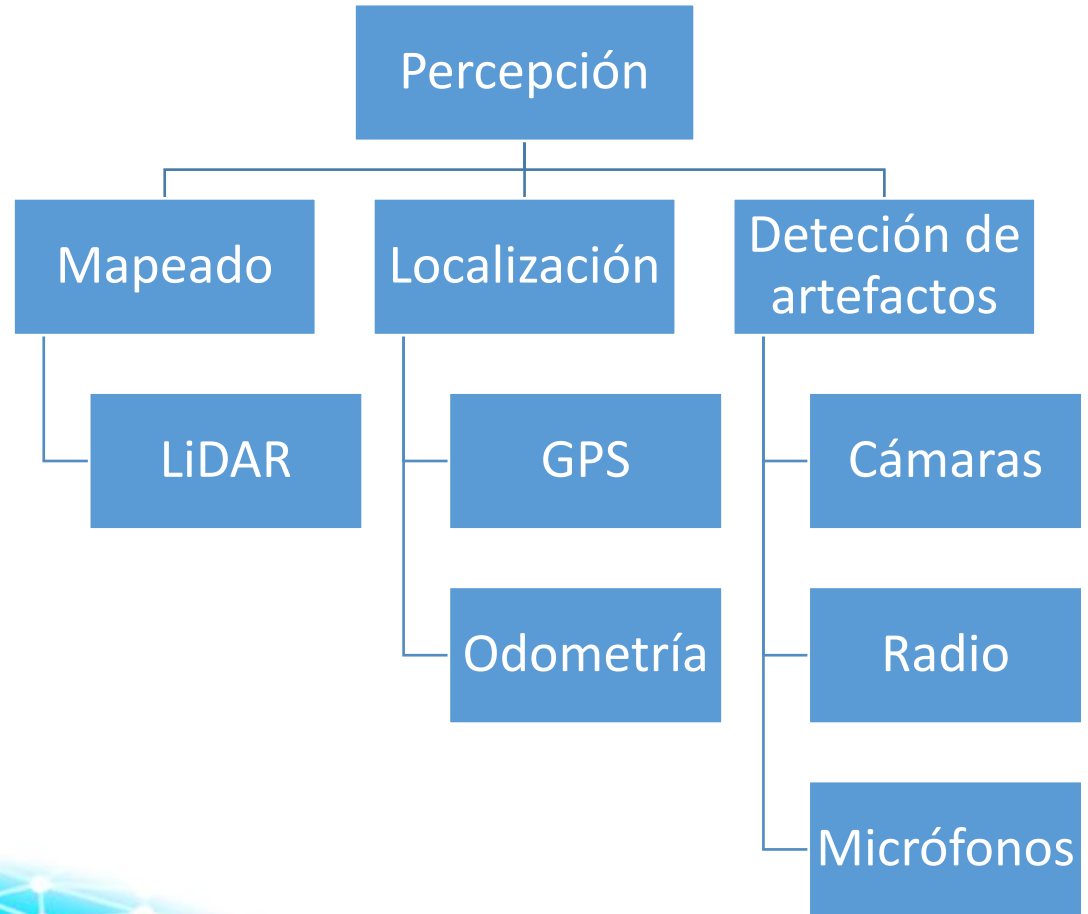
Introducción

- Robot: Máquina capaz de percibir, planear e actuar de forma autónoma

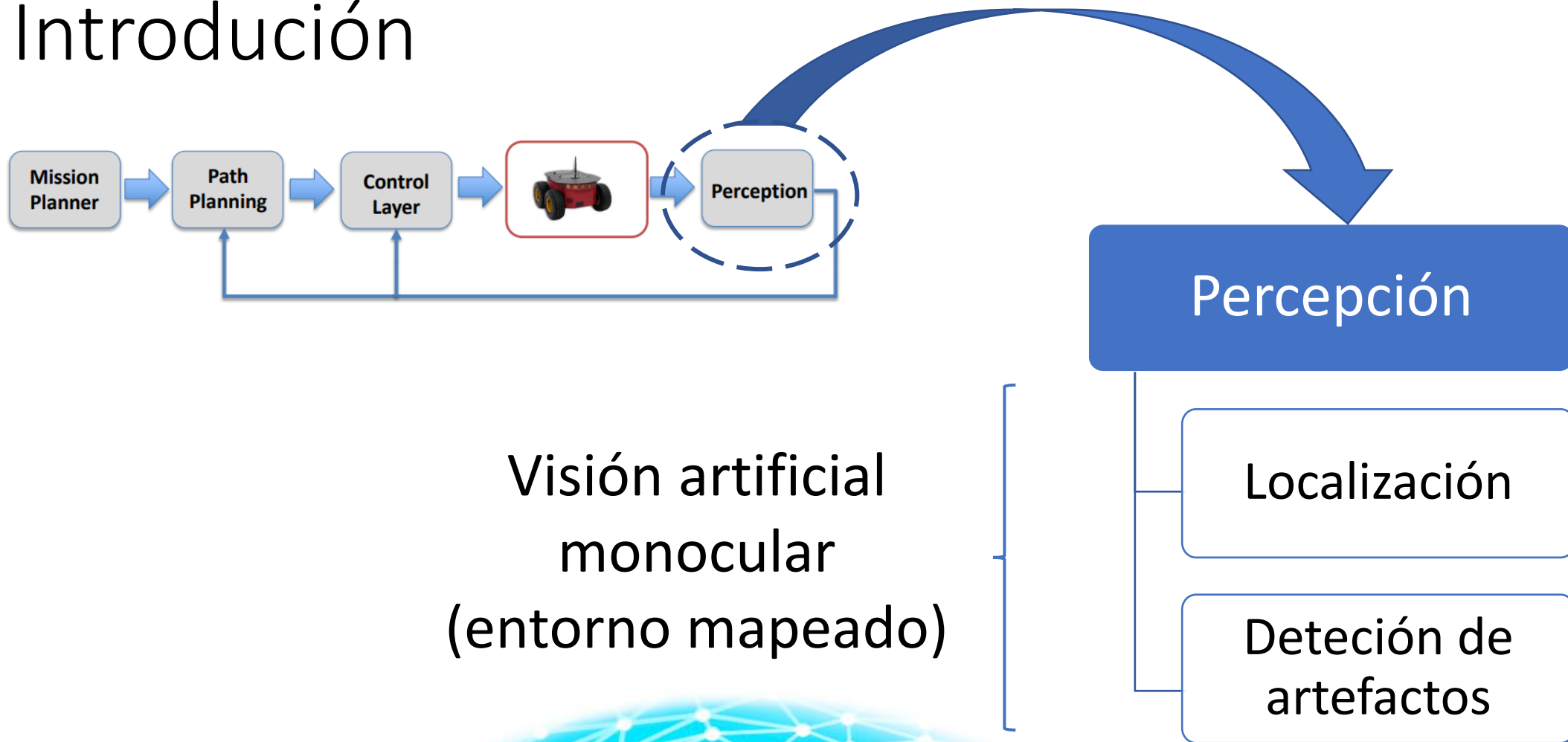


Introducción

- Típica tarea de exploración
 - Mapeado
 - Localización do axente (robot)
 - Deteción de artefactos (obxectos)
- Multiple sensórica
- Realizable só con cámara e visión artificial



Introducción

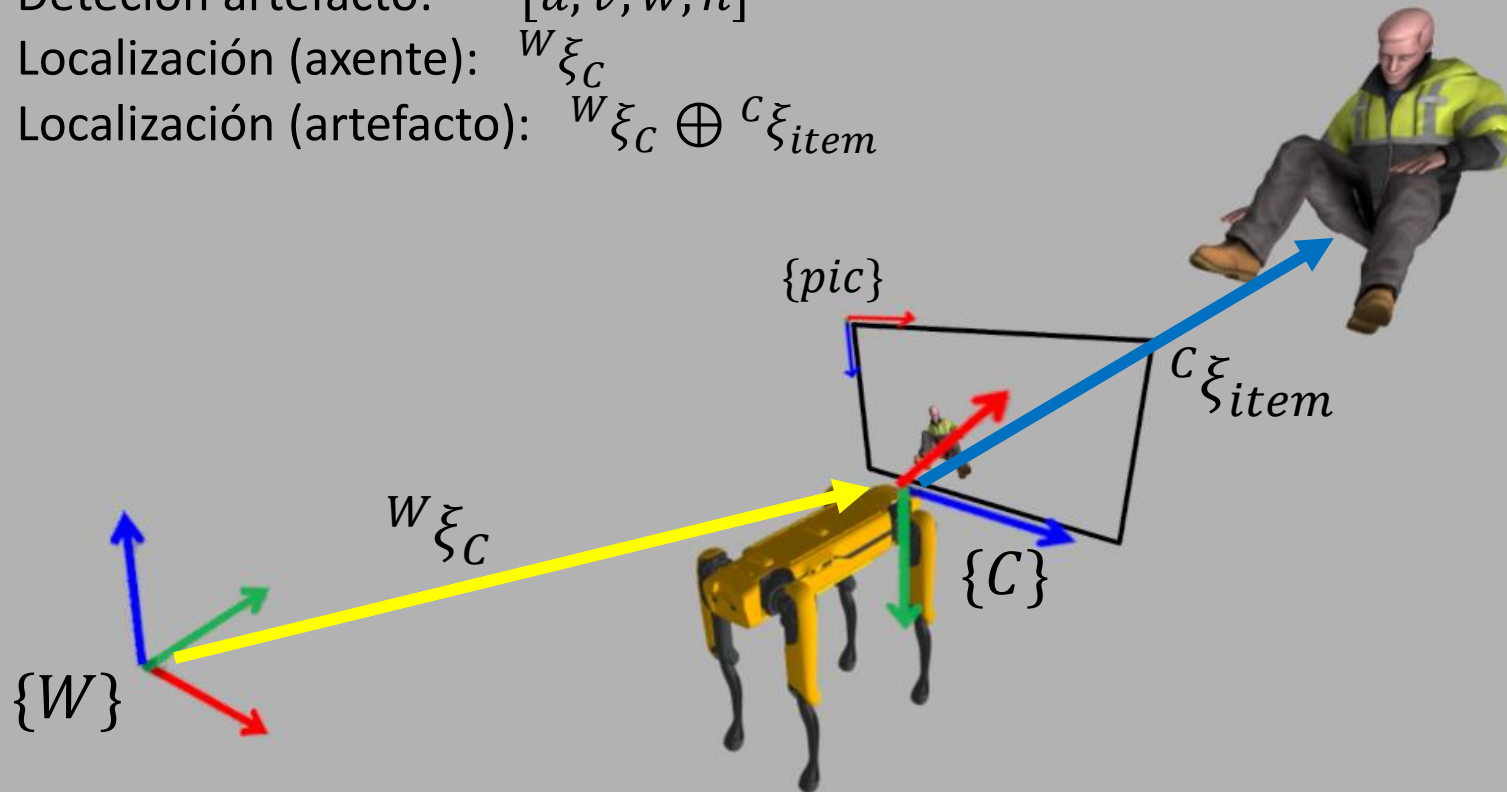


Introducción

Detección artefacto: $pic[u, v, w, h]$


Localización (axente): ${}^W\xi_C$

Localización (artefacto): ${}^W\xi_C \oplus {}^C\xi_{item}$



Obxectivos

- Plantear a detección dos obxectos do DARPA SubT Challenge con visión artificial
 - Construír un set de datos para entrenar redes neurais
 - Entrenar e comparar formalmente un set de modelos de detección de obxecto estado do arte nos artefactos deste reto
- Construír unha capa de percepción monocular para busca de obxectos
 - Acadar a localización do axente baseado nun sistema fiducial
 - Acadar a localización de obxectos e implementar unha capa de percepción en tempo real



DARPA SubT con detección de obxectos

- Detección de obxectos
- Recollida de datos
- Selección de modelos
- Entrenando redes neurais
- Resultados

Detecção de obxectos

- Técnica que busca atopar obxectos (clases) en imaxes debuxando unha caixa arredor deles (bounding box)
- Baséase en redes neurais convolucionais:
 - Constituída por capas convolucionais
 - Modelo estadístico de regresión que infire regras a partir de datos exemplo

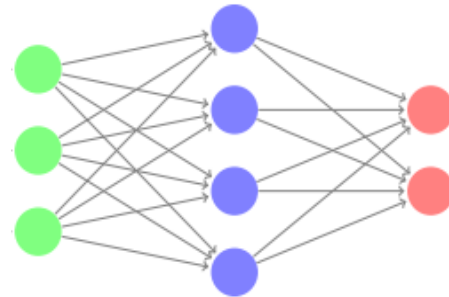


Detección de obxectos

Imaxe etiquetada



Rede neural



Predicción da rede

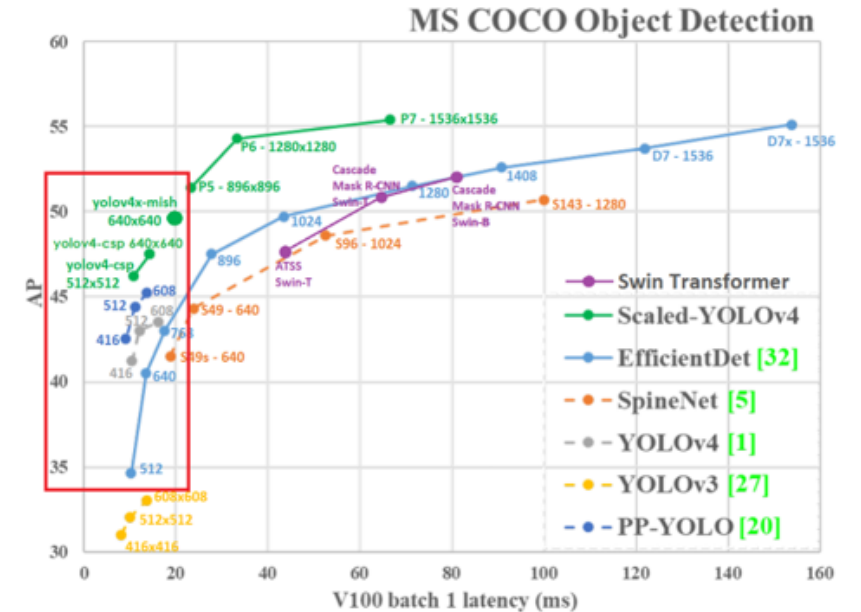


Actualización
("aprendizaxe")

Deteción de obxectos

- Benchmarking: comparación de distintos modelos sobre os mesmos datos
- Métricas de interese:
 - Precisión: AP / mAP
 - Velocidade de inferencia: FPS, ms
- Interesa realizalo para o noso caso

Regular detectors



Recollida de datos

- Datos representativos ao caso de estudo:
 - Datos sintéticos xenerados
 - Datos reais importados
 - Datos reais xenerados
- Overfitting: división en tres splits
 - Entrenamento: train + valid
 - Evaluación e comparación: test



Dataset



Entrenamento

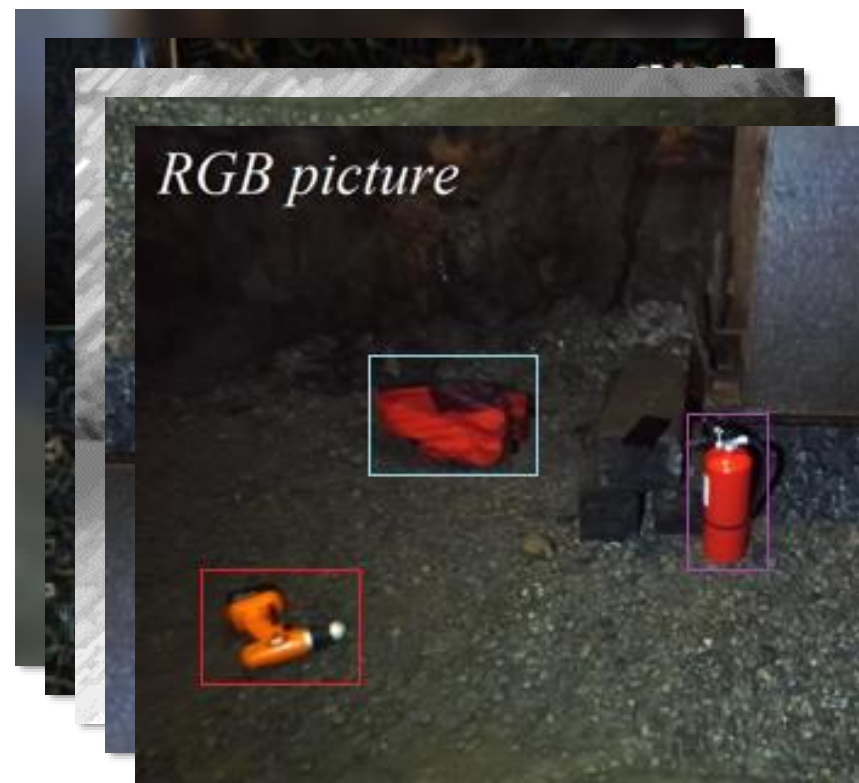
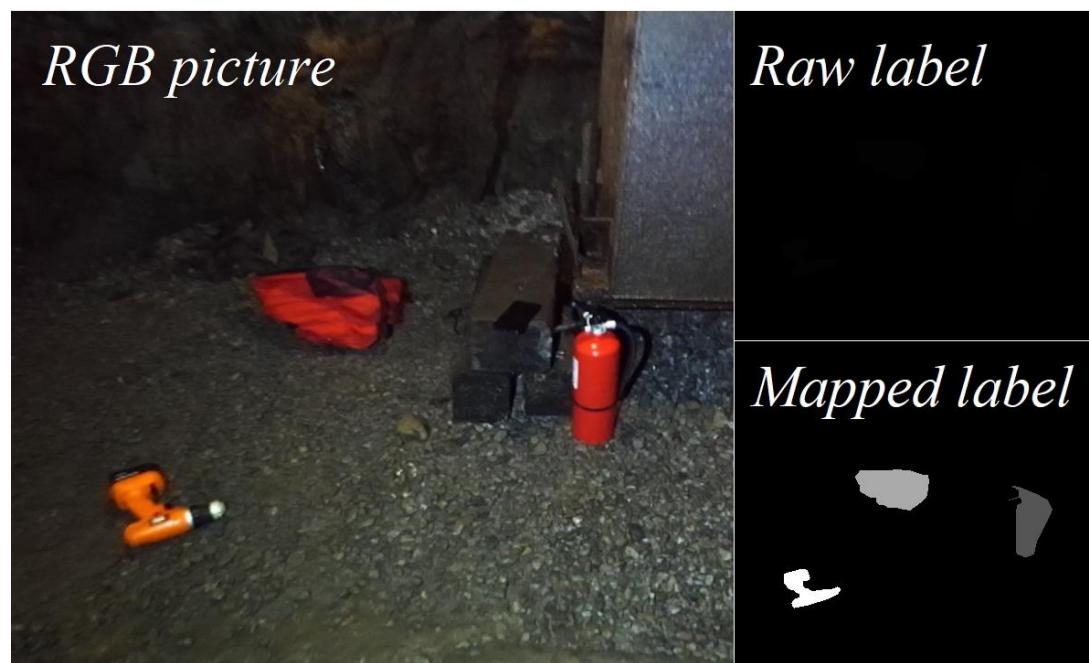
Validación

Test

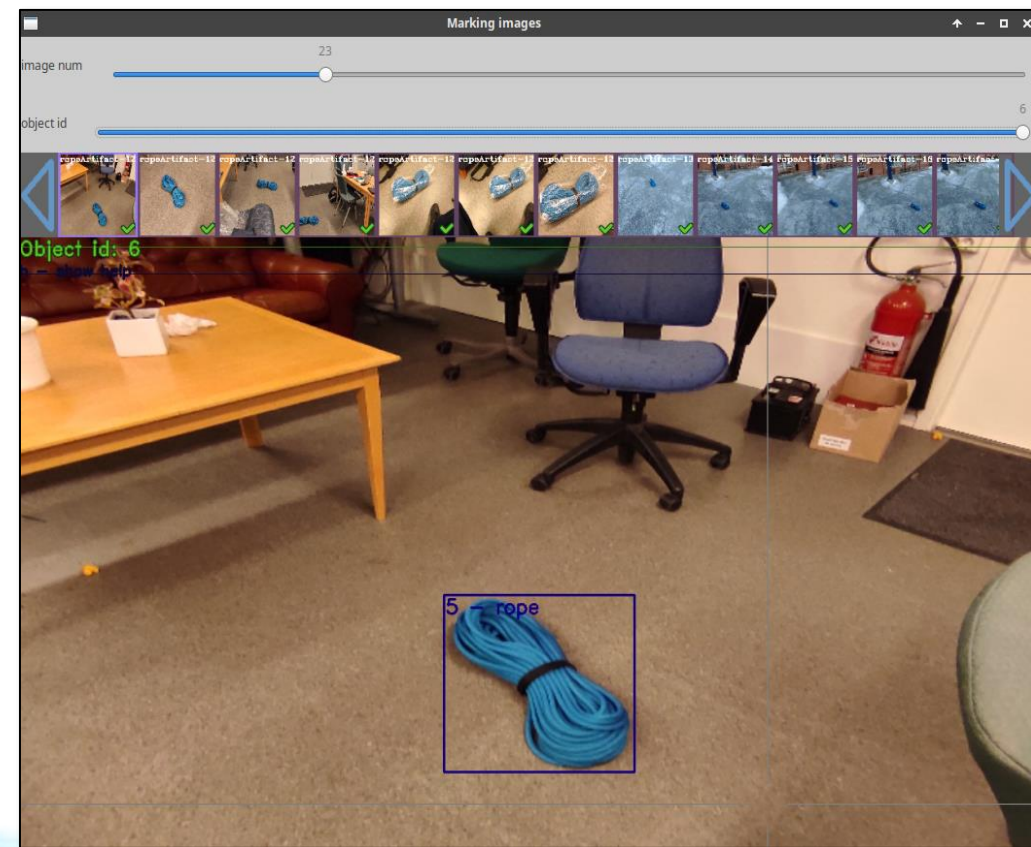
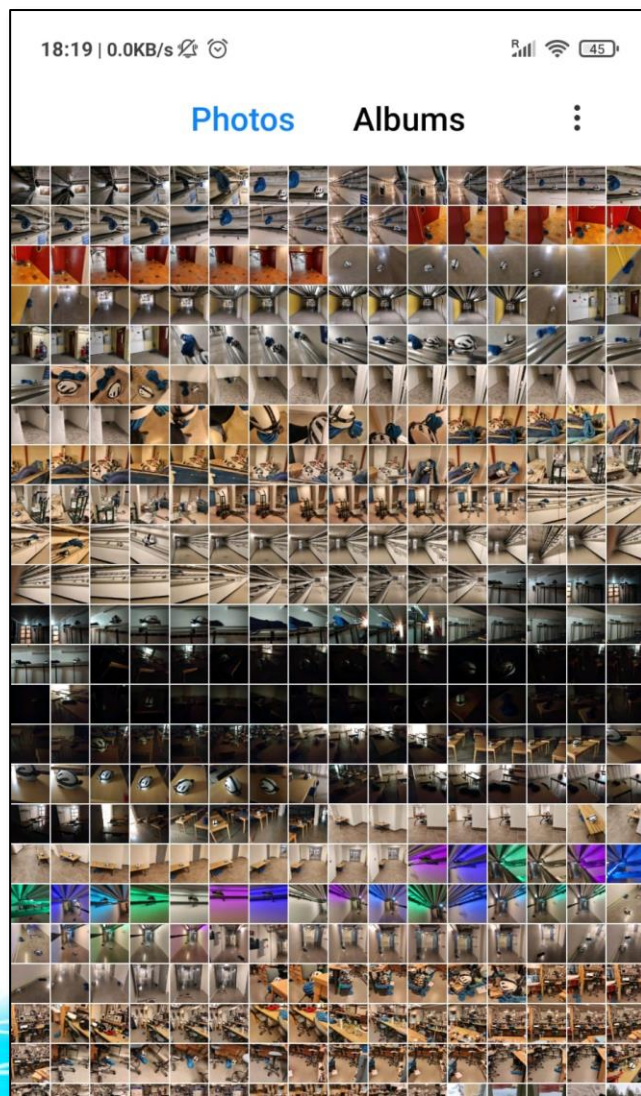
Recollida de dats



Recollida de dats

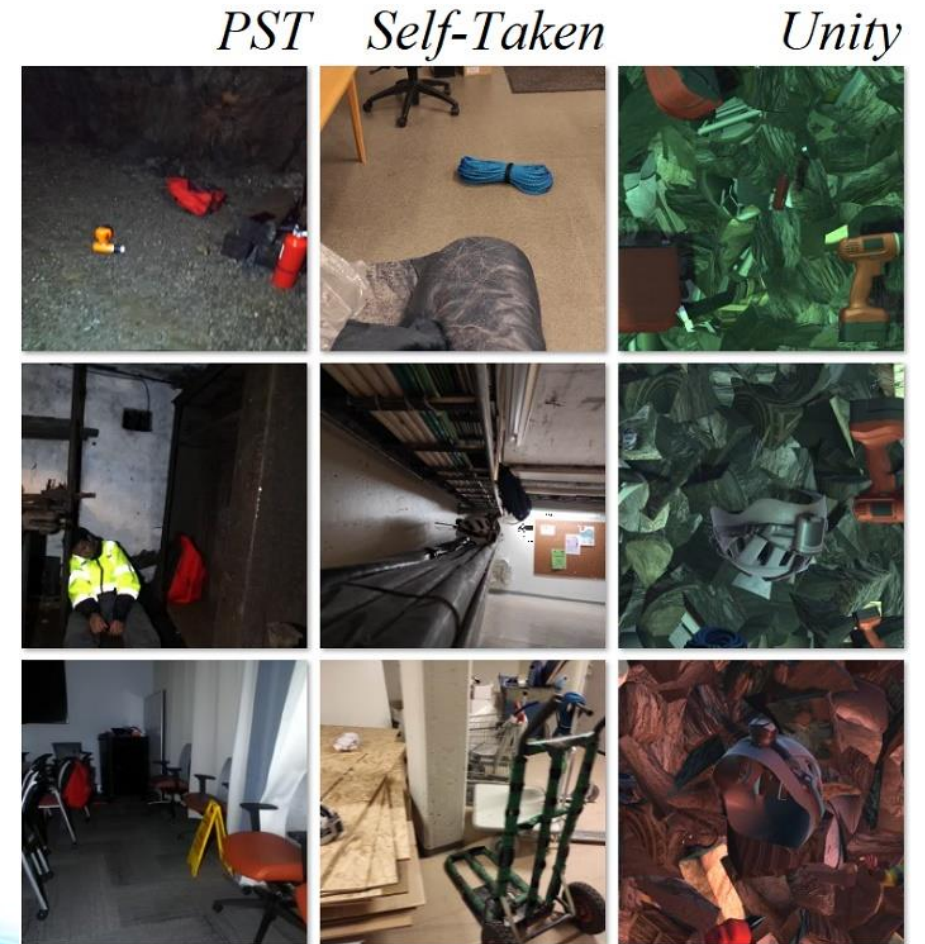


Recollida de dats



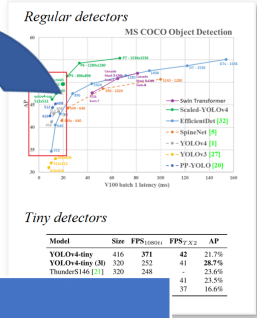
Recollida de datos

- PPU-6:
 - Set de entrenamiento mixto real-sintético
 - Set de validación real
 - Set de test real
- Unity-1000
 - 1000 imaxes sintéticas



Entrenando redes neurais

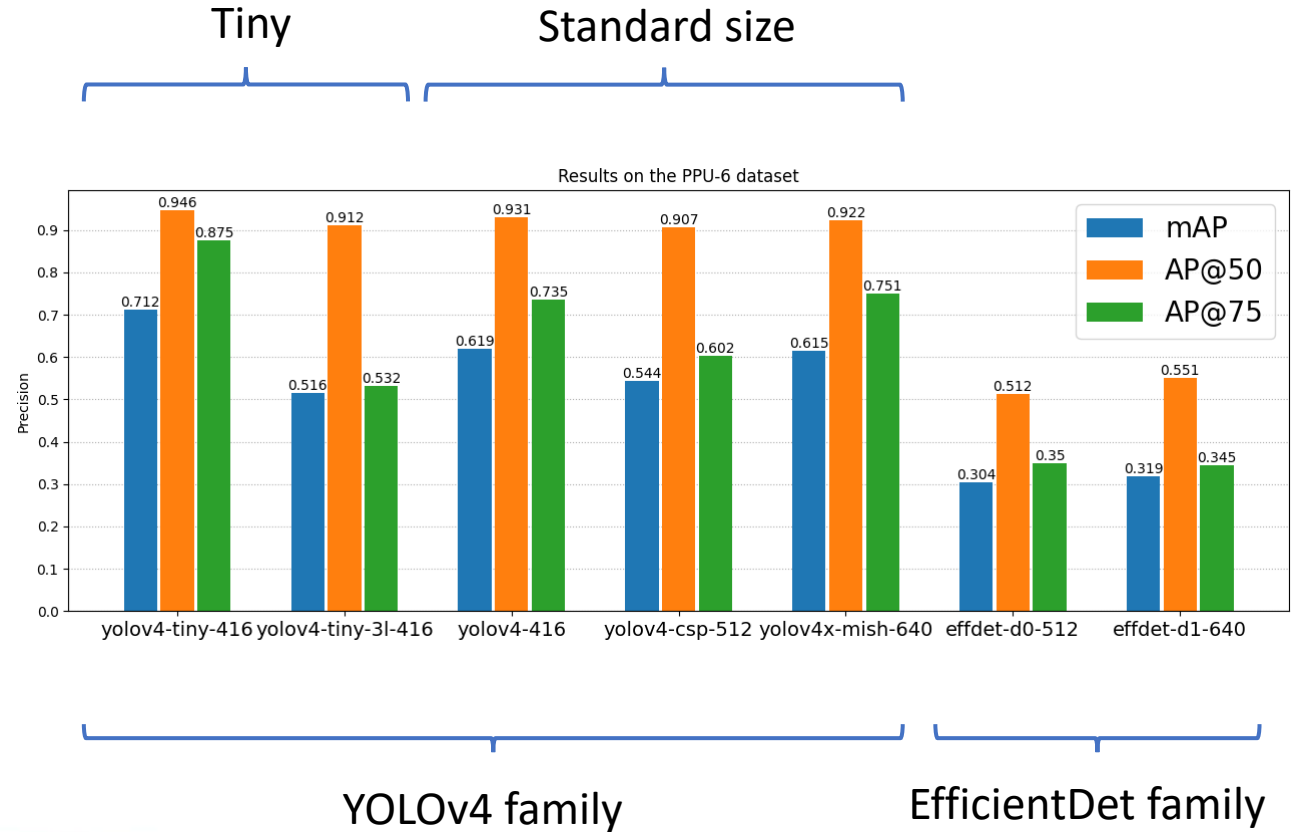
- Selección de modelos:
 - A partir de bechmarking MS COCO
 - Compromiso AP-FPS
 - Tiny & estándar
- Estratexia de entrenamiento:
 - 3 etapas
 - 13 configuraci3ns diferentes



Network	Platform	Resolution
yolov4-tiny	darknet	416x416
yolov4-tiny-3l	darknet	416x416
yolov4	darknet	416x416
yolov4-csp	darknet	512x512
yolov4x-mish	darknet	640x640
efficientdet-d0	google-automl	512x512
efficientdet-d1	google-automl	640x640

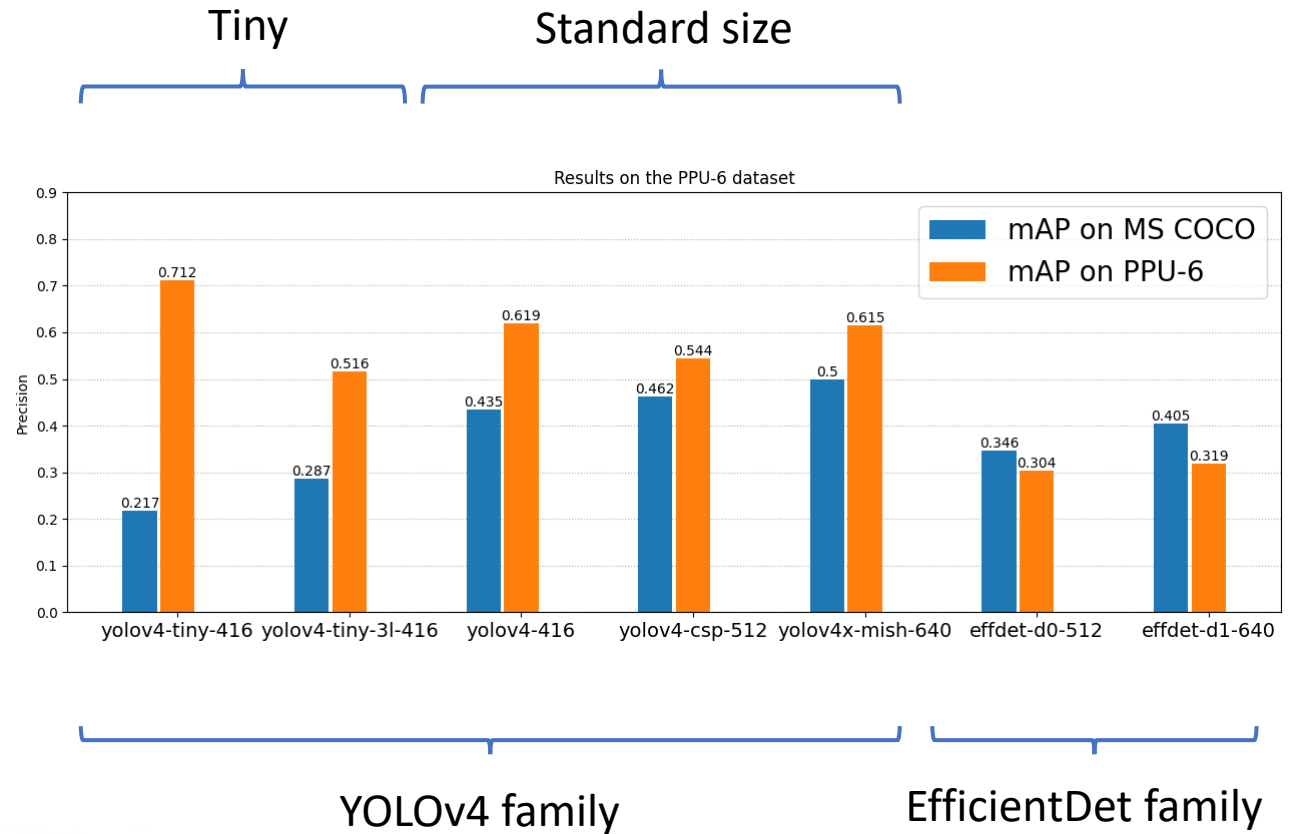
Resultados

- Resultados variados, pero AP@50 estable
- Modelos portables (tiny) competitivos cos estándar
- Familia YOLO notablemente mellor que EfficientDet



Resultados

- Resultados variados, pero AP@50 estable
- Modelos portables (tiny) competitivos cos estándar
- Familia YOLO notablemente mellor que EfficientDet



Resultados adicionais

Deteccions sobre imaxe escura

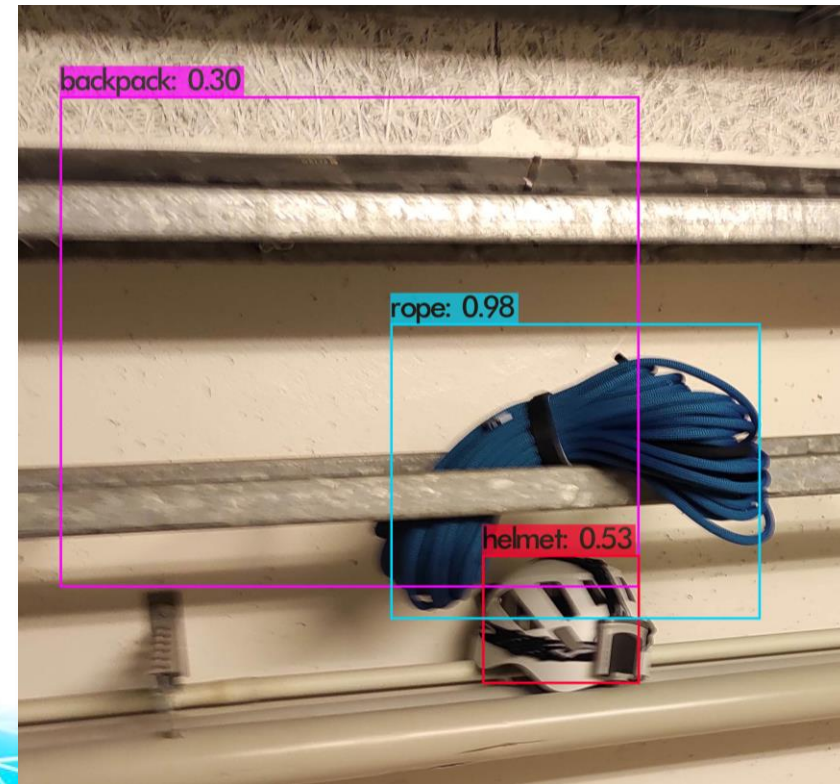
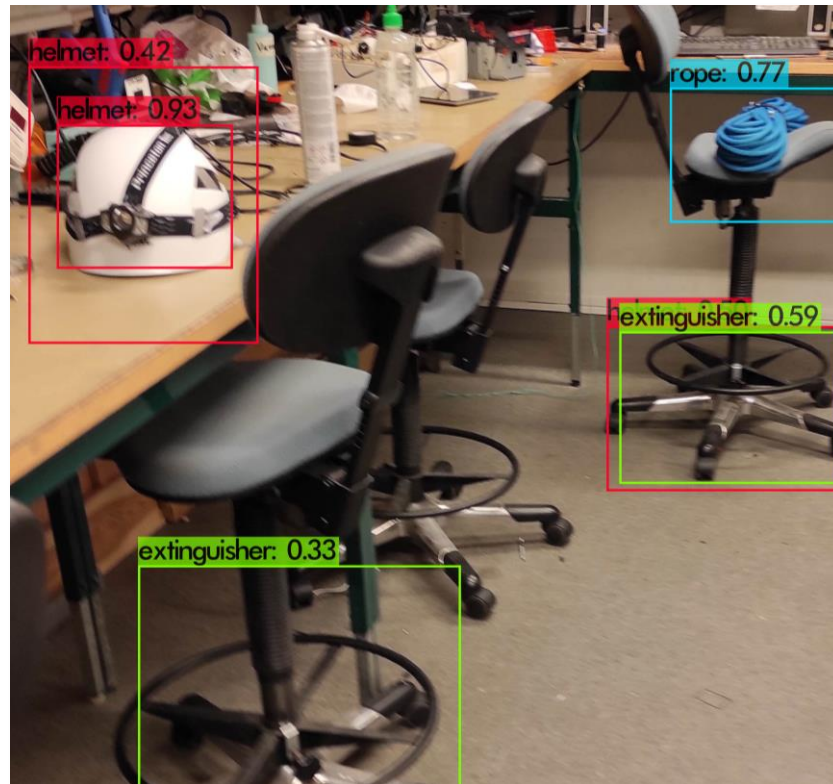



Mesma imaxe, filtrada



Resultados adicionais

Deteccions dun YOLOv4 entrenada exclusivamente en imaxes sintéticas



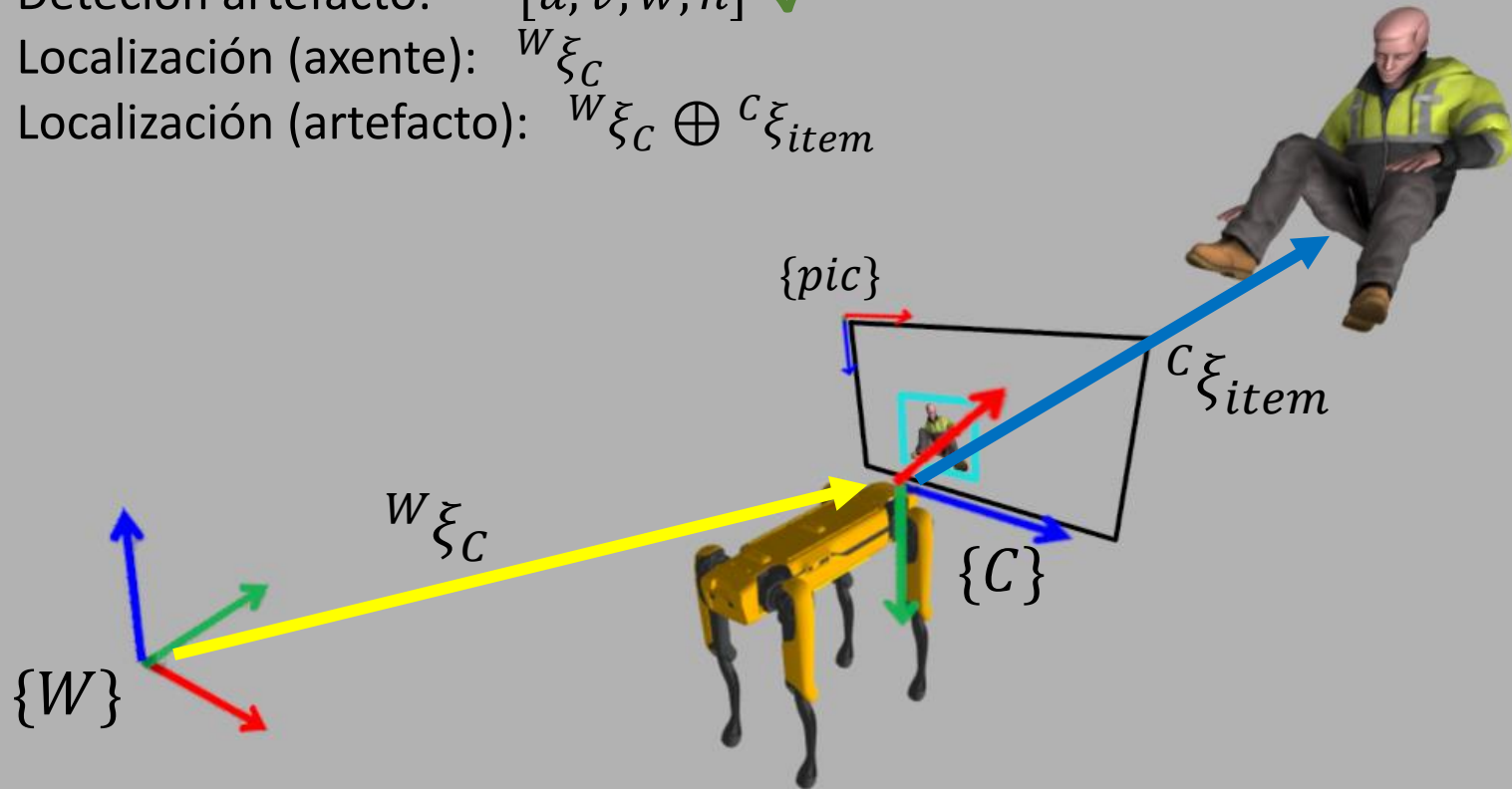


Localización e capa de percepción

- Calibración de cámara
- Localización (autolocalización)
- Localización de obxectos
- Mapas de verosimilitude
- Capa de percepción

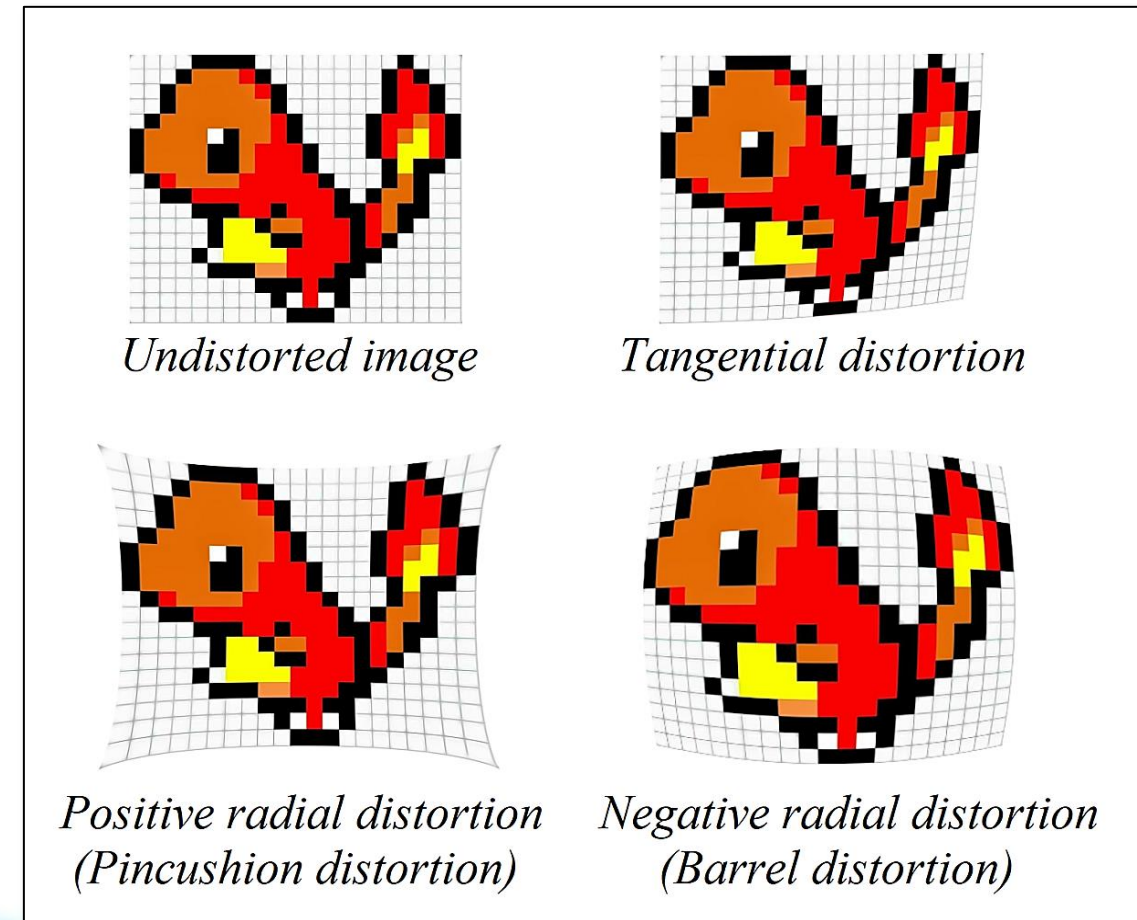
Localización e capa de percepción

Detección artefacto: $pic[u, v, w, h]$ ✓
Localización (axente): ${}^W\xi_C$
Localización (artefacto): ${}^W\xi_C \oplus {}^C\xi_{item}$



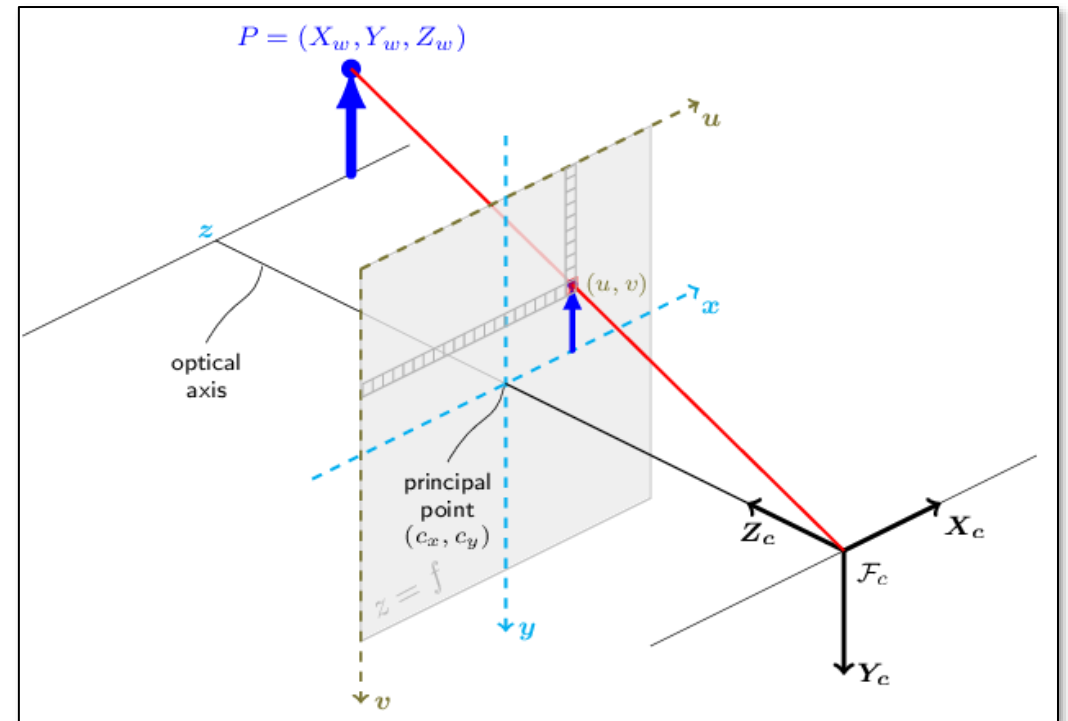
Calibración de cámara

- Cámaras reales producen imaxes con distorsións:
 - Radial
 - Tanxencial
- Calibración permite:
 - Compensación de distorsións
 - Obter a matriz intrínseca da cámara



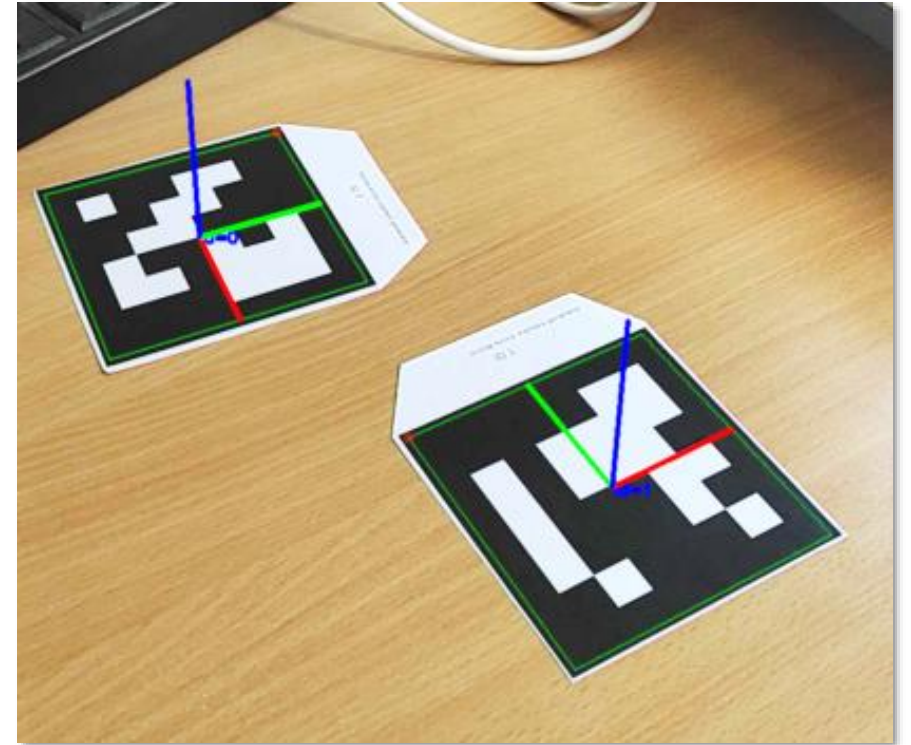
Calibración de cámara

- Cámaras reales producen imaxes con distorsións:
 - Radial
 - Tanxencial
- Calibración permite:
 - Compensación de distorsións
 - Obter a matriz intrínseca da cámara

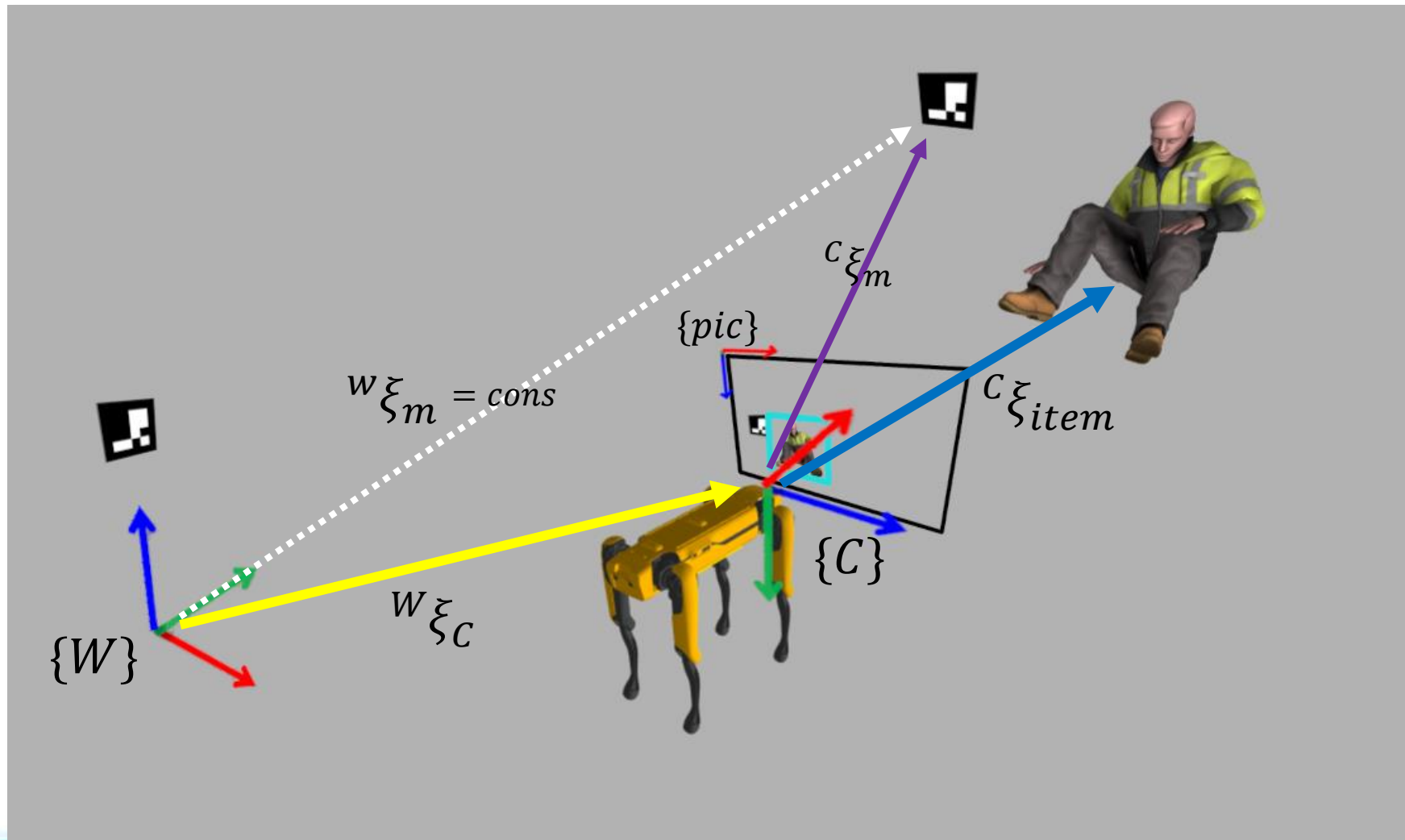


Localización

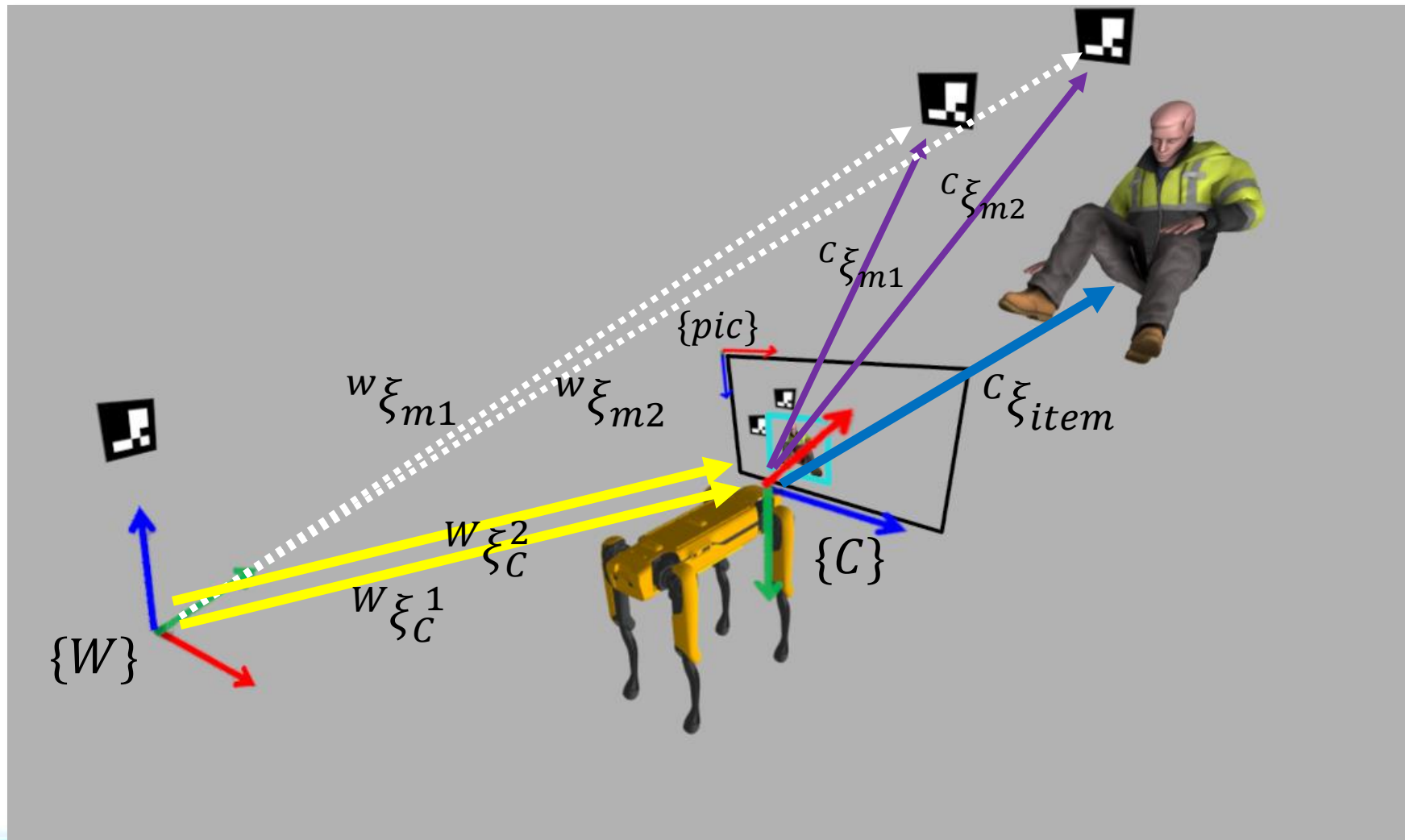
- Sistema fiducial:
 - Información de indentificación
 - Referencia posicional
- Pose 3D facilmente detectable desde imaxes libres de distorsión
- Mundo mapeado cun layout de marcadores ArUco coñecido



Localización

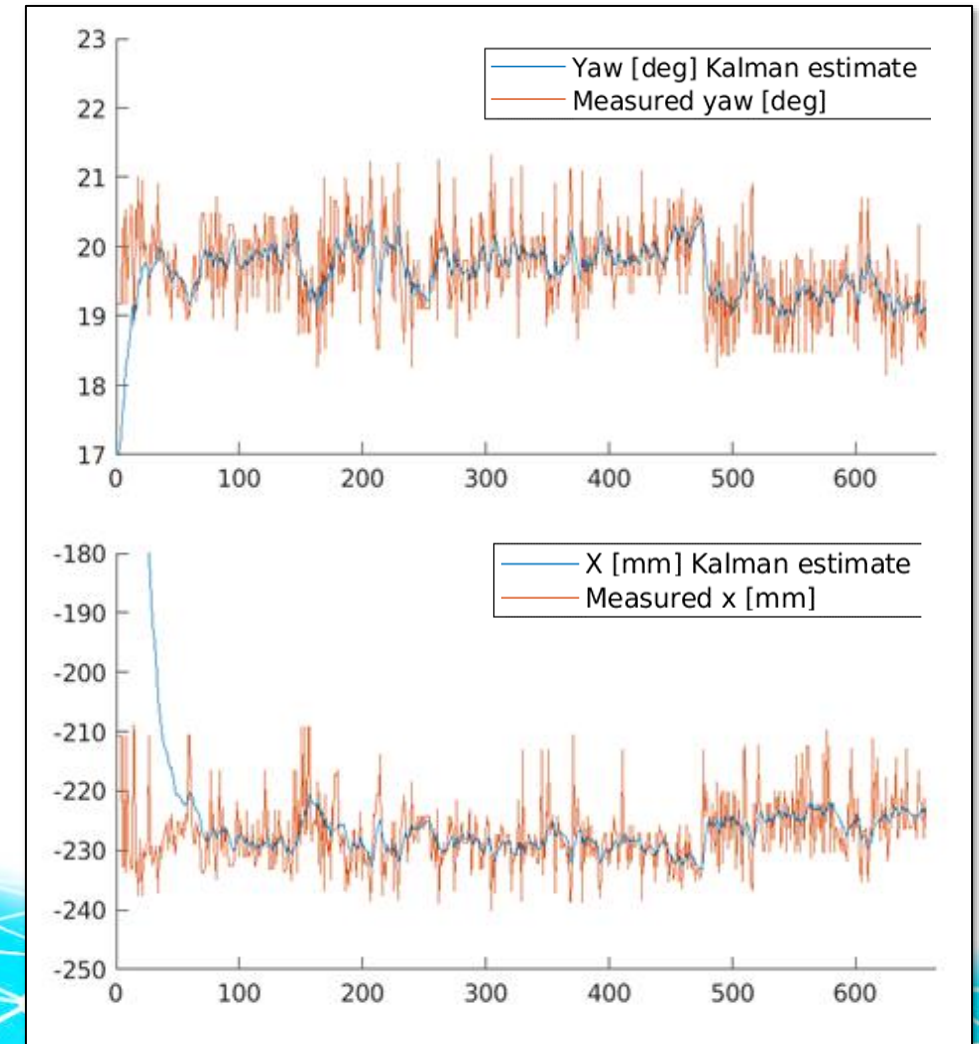


Localización

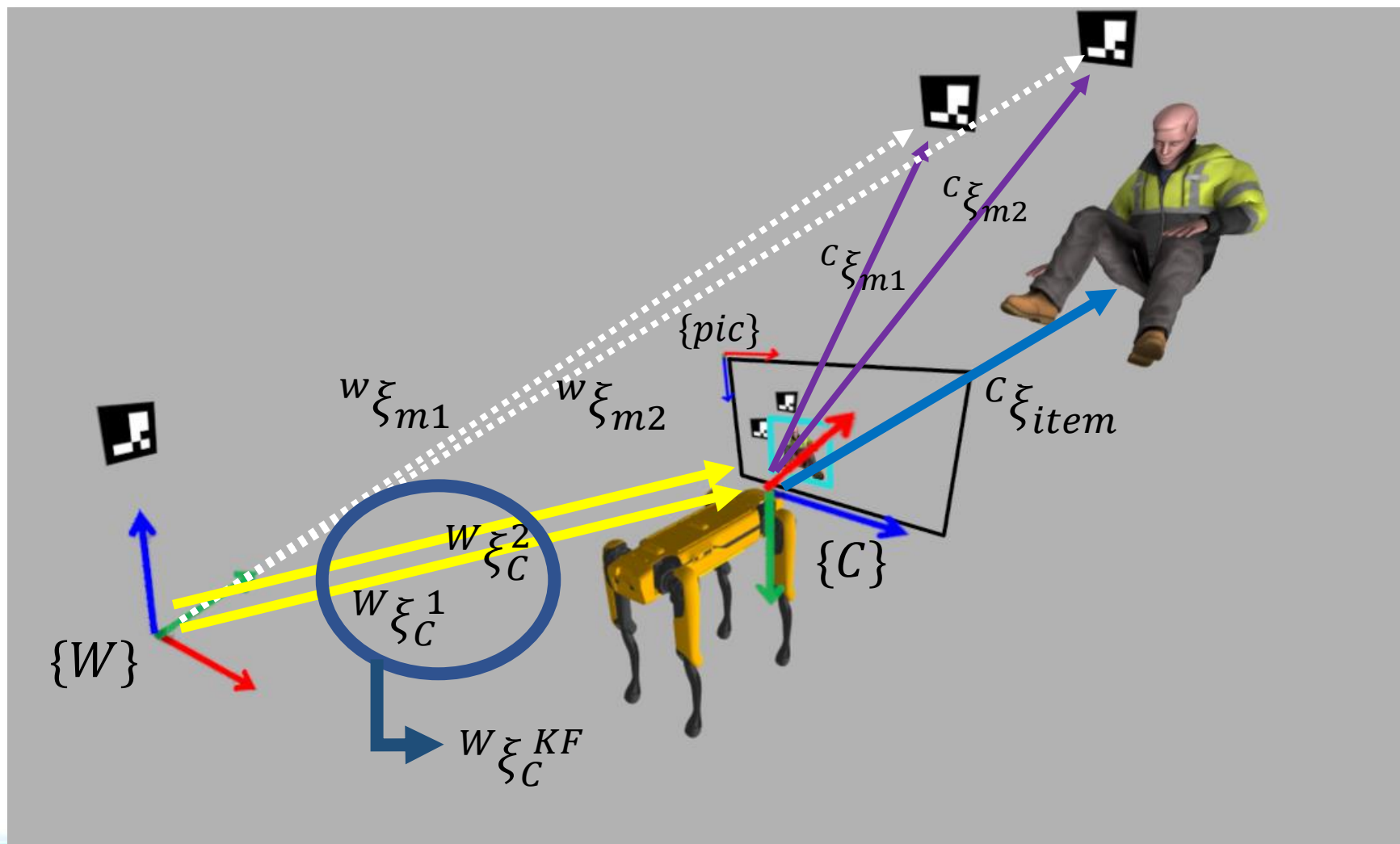


Localización. Filtro de Kalman

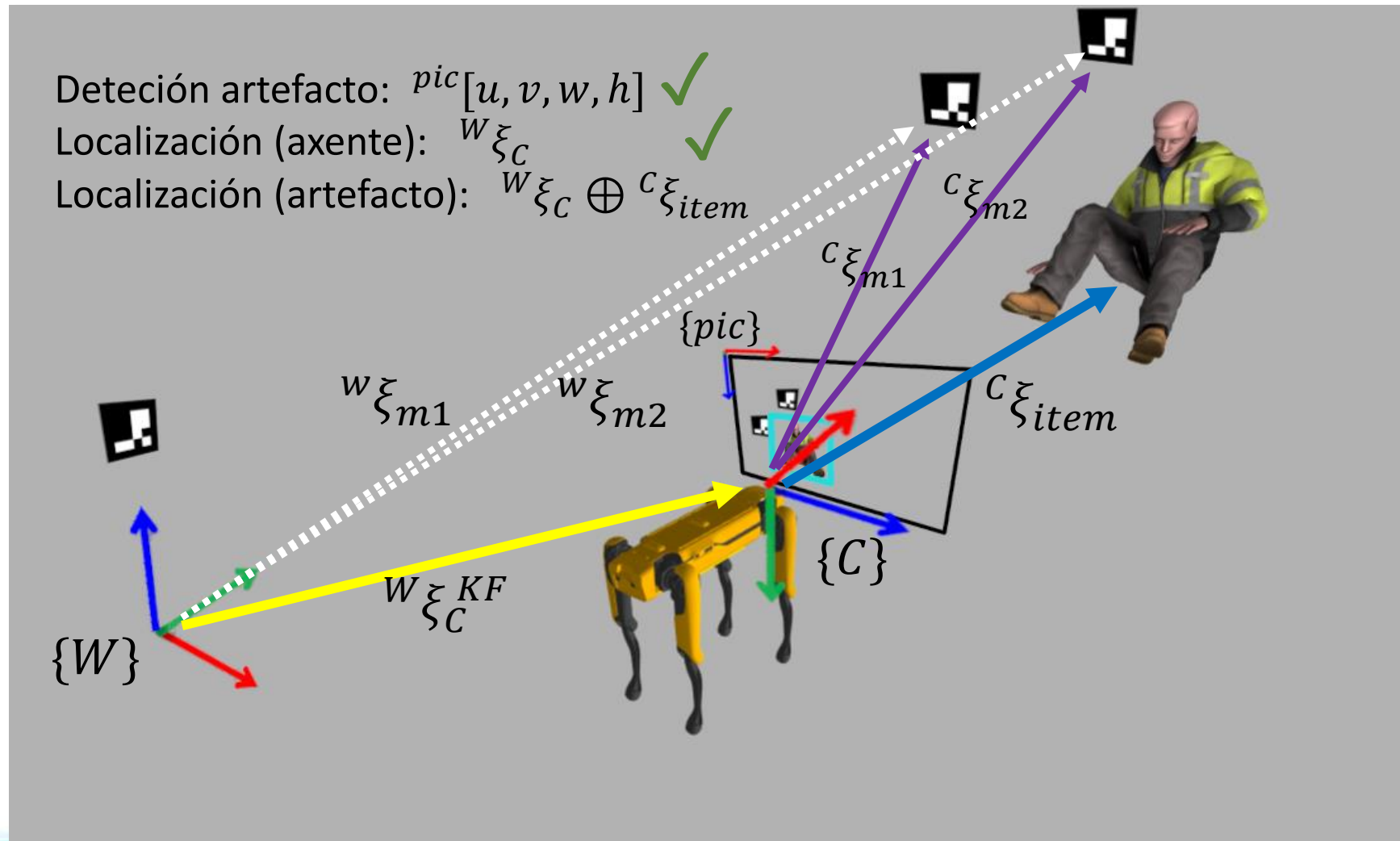
- Filtro de ruido blanco da pose:
 - Posición dentro de tolerancia ($30mm$)
 - Yaw propaga demasiado erro (2°)
- Número variable de ArUcos:
 - Fusión sensorial
 - Modelo de estado e filtro de Kalman de tamaño variable



Localización. Filtro de Kalman

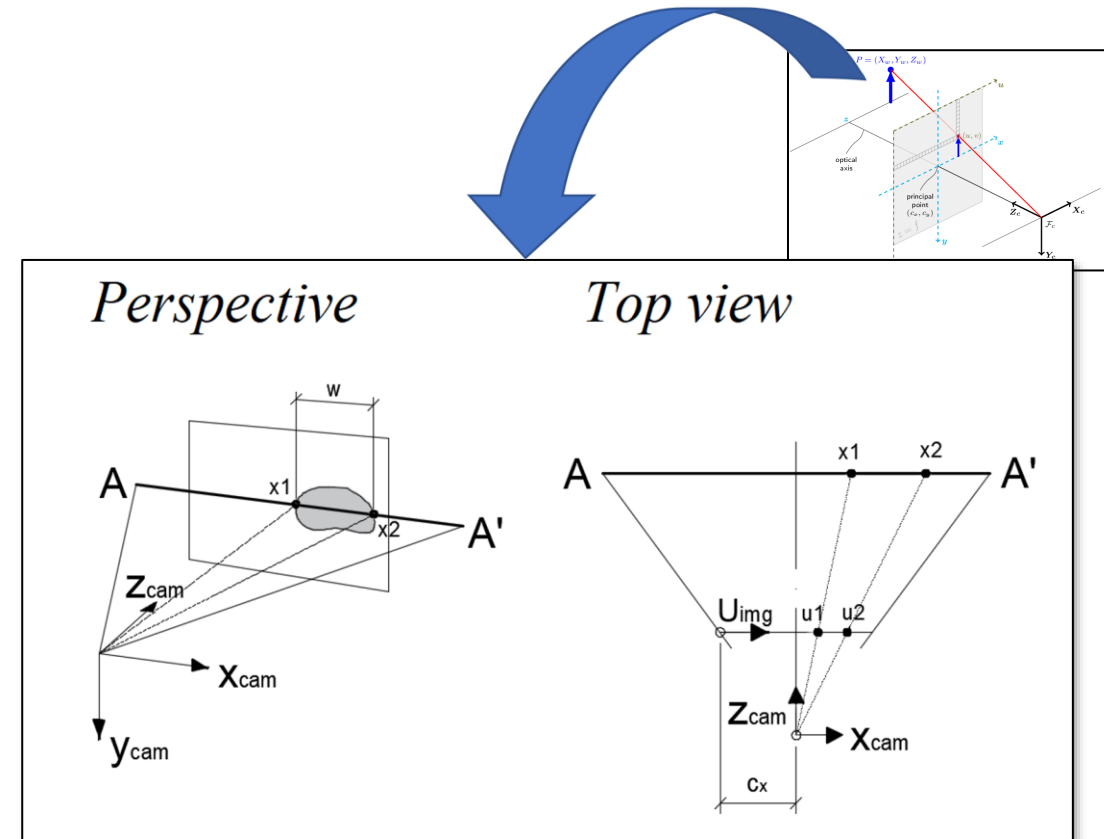


Localización de artefactos

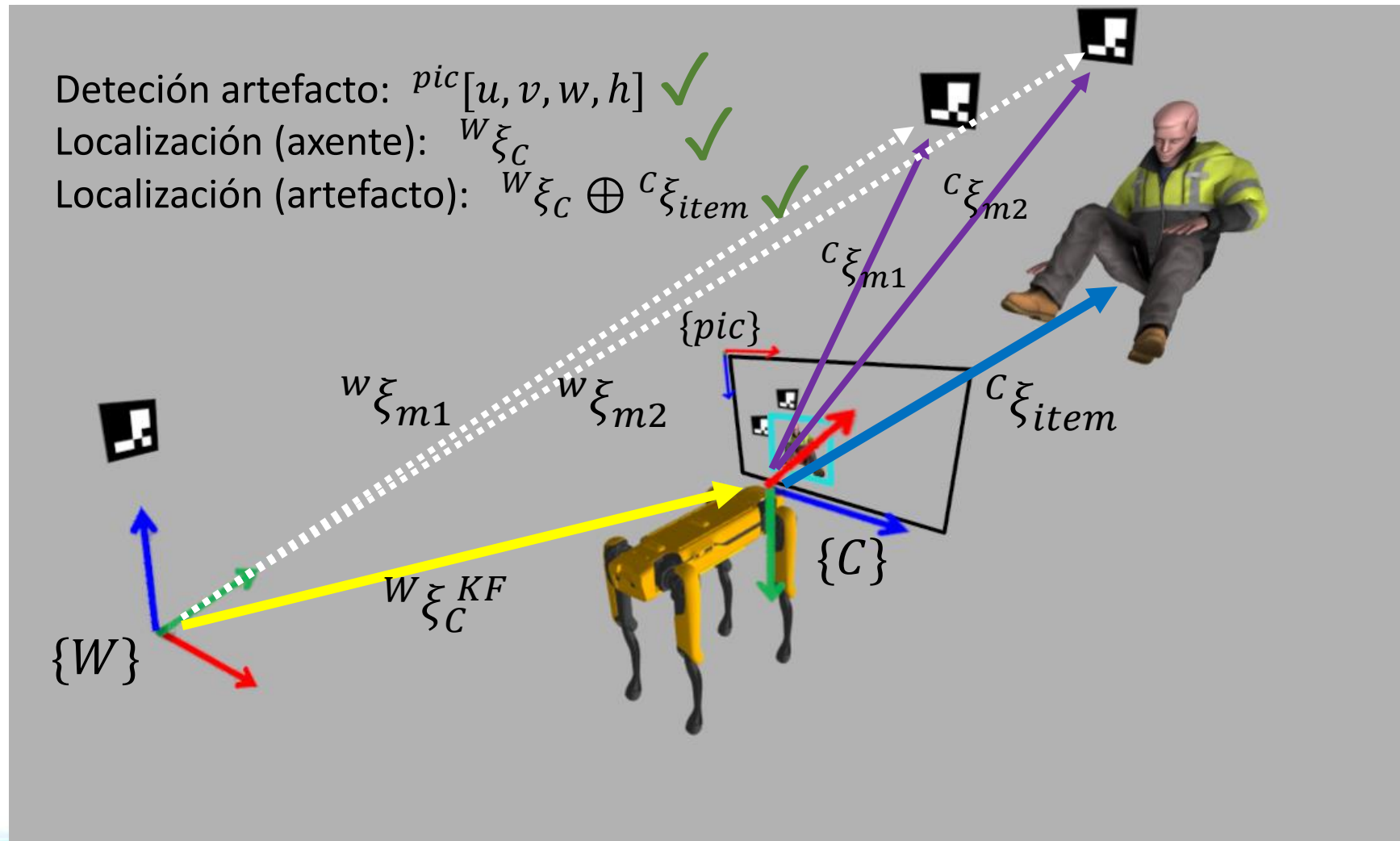


Localización de artefactos

- Modelo + matriz intrínseca de cámara permite calcular posición 2D relativa objeto-cámara
- Parte de:
 - Ancho estimado de objeto coñecido
 - Imaxe libre de distorsións
 - Imaxe pasada por detector de obxectos

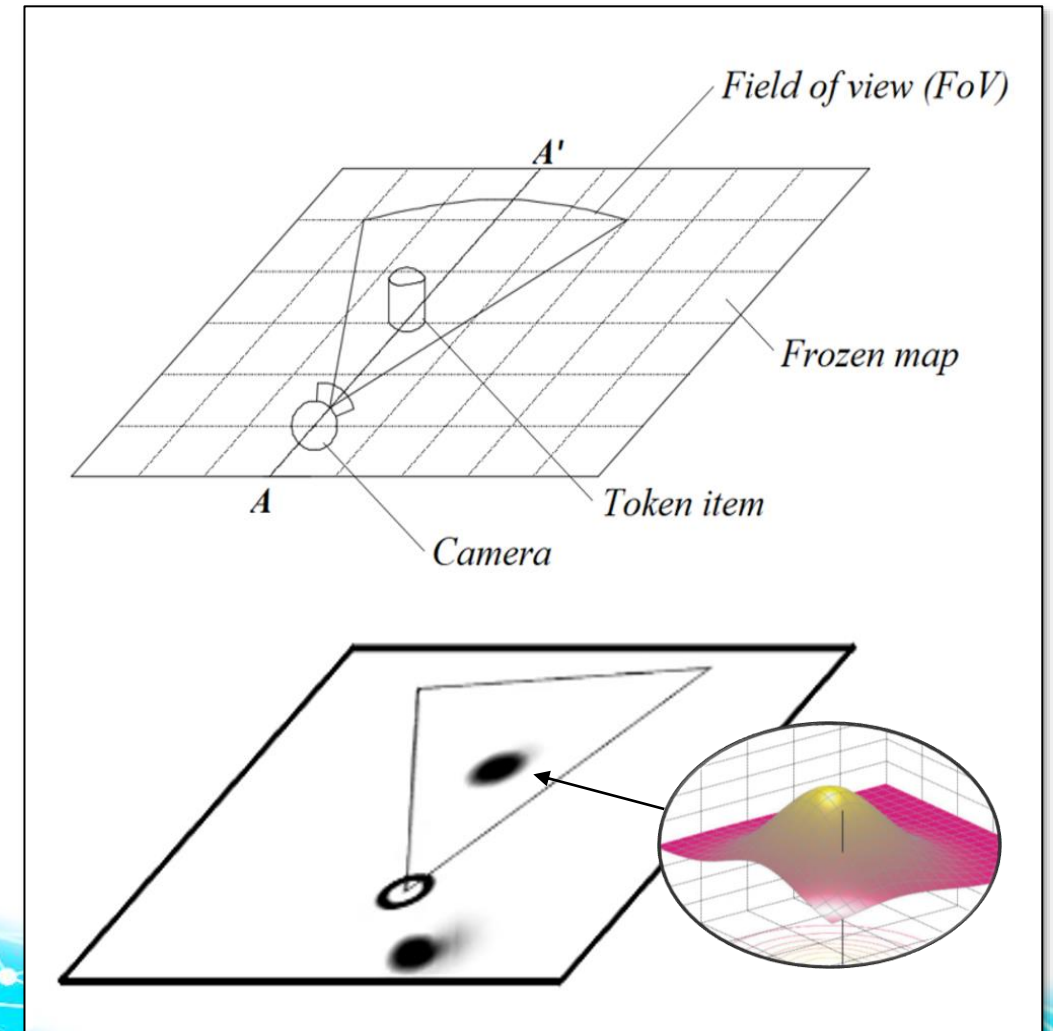


Localización de artefactos

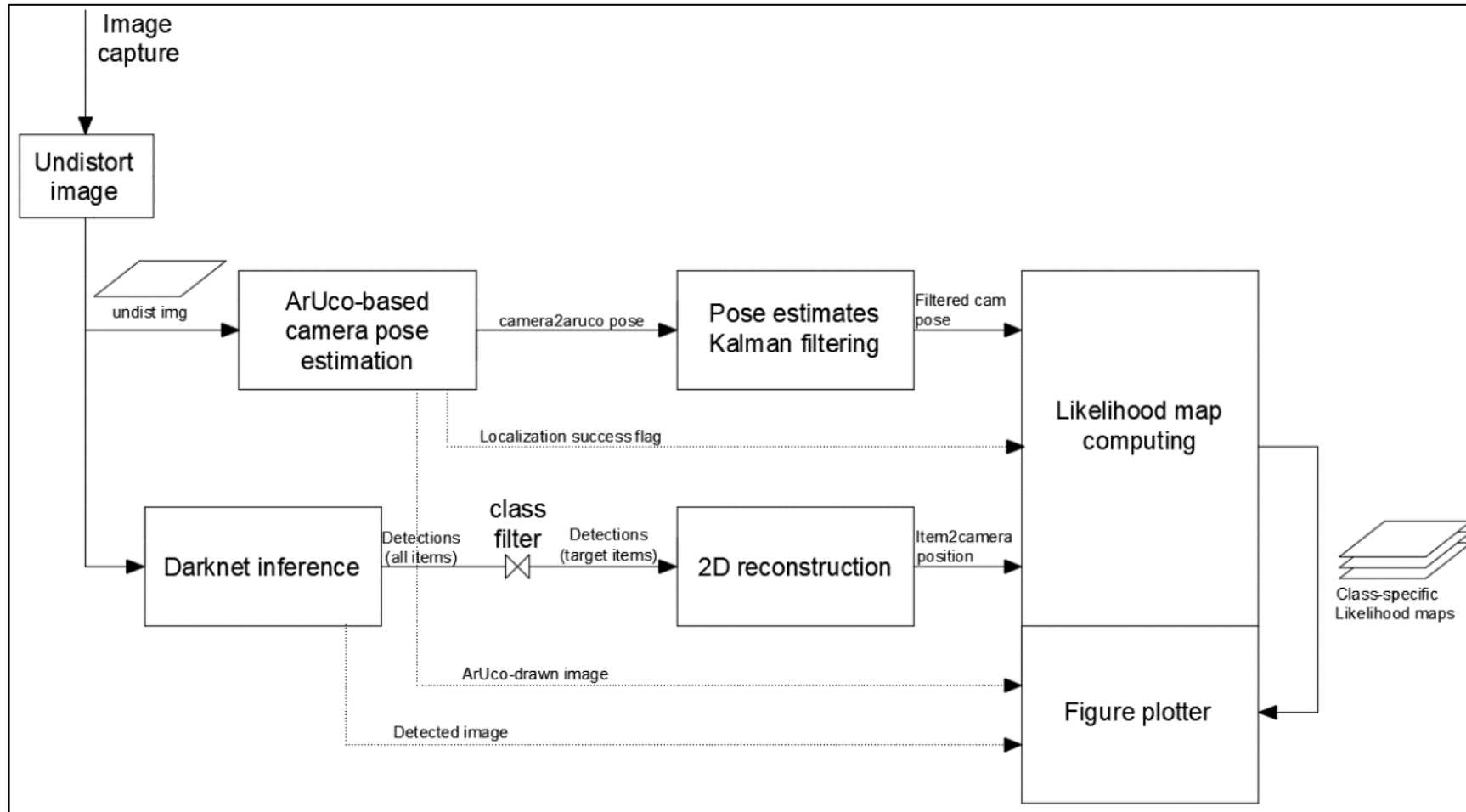


Mapas de verosimilitud

- Conceptos fundamentales:
 - Campo de visión
 - Memoria
- Ecuación matricial en diferencias que representa a posición dos items:
 - + Kernel gaussiano no punto de detección
 - – Distribución uniforme sobre FoV



Capa de percepción





Demostración

- Setup experimental
- Vídeo-demostración
- Comentarios

Setup experimental

- Obxectos DARPA “de mentira”
- Detector YOLOv4-CSP
- Coche teledirixido cunha cámara
- Mapa de 30x6m
- 9 marcadores ArUco

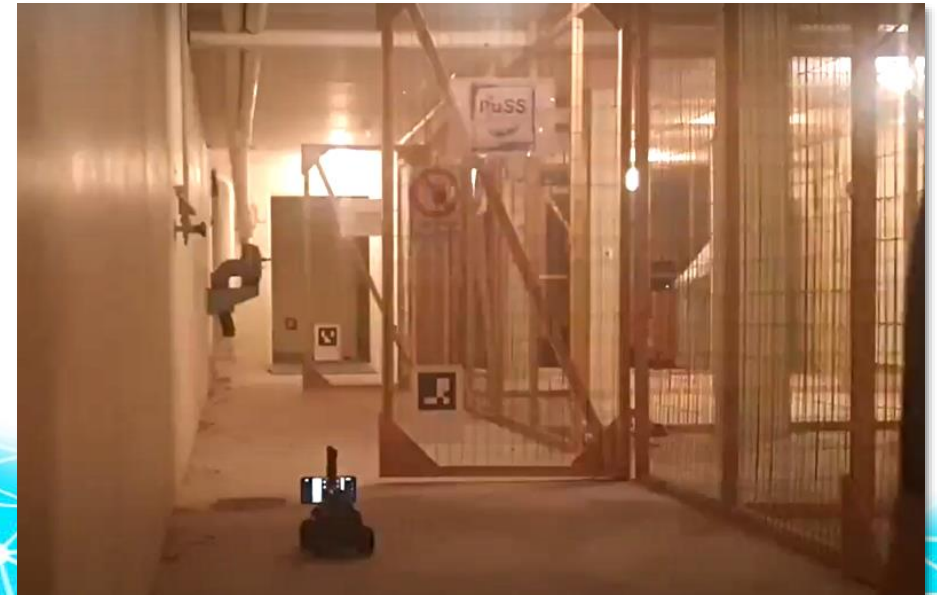


Vídeo-demostración



Observacións

- Dúbida de clase co survivor
- Inestabilidade caixas de detección
- Xenérase ruído que se limpa
- Intermitencia detección taladro

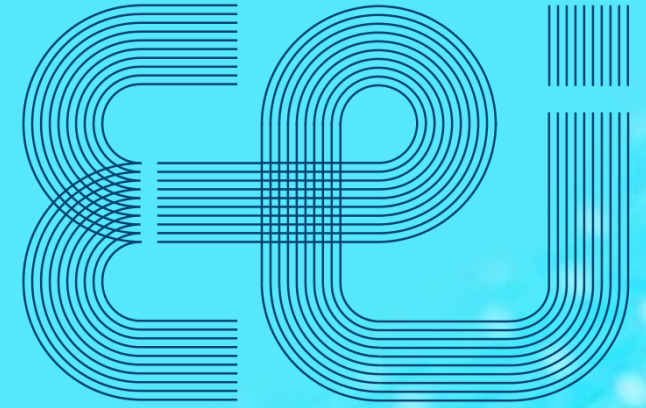


Conclusións

- Estúdase o problema de detección de obxectos para os obxectos DARPA:
 - Obtéñense datos para este caso específico
 - Benchmarking de modelos e.d.a. obtendo conclusións relevantes
- Constrúese unha capa de percepción monocular en tempo real baseada en detección de obxectos:
 - Robusta a erros de medida
 - Agnóstica ao modelo de detección de obxectos
 - Simple incorporación a dispositivos autónomos

`github.com/pabsan-0/sub-t`

Redes neurais convolucionais para detección de obxectos en entornos subterráneos



Universidade de Vigo *, Luleå Tekniska Universitet †
Mestrado en Enxeñaría Industrial
Traballo de fin de Mestrado

Pablo Santana González

Enrique Paz Domonte*

George Nikolakopoulost†

Christoforos Kanellakist†