



# ENTREGABLE E-10

## Informe sobre bondad de ajuste y validación de algoritmos

### SUBSANACIÓN

PROYECTOS EN EL ÁMBITO DE LAS  
TECNOLOGÍAS HABILITADORAS DIGITALES

<b>Número de expediente:</b>	TSI-100909-2019-62
<b>Programa:</b>	Impulso Tecnologías Habilitadoras Digitales (THD)
<b>Título del proyecto:</b>	TWINFOREST: GEMELO DIGITAL FORESTAL PARA CLASIFICAR USOS DEL SUELO MEDIANTE LA COMBINACIÓN DE VISIÓN ARTIFICIAL, APRENDIZAJE PROFUNDO Y SENSORES REMOTOS
<b>Beneficiario:</b>	FORA FOREST TECHNOLOGIES, S.L.L.

**DOCUMENTO:**  
ENTREGABLE E-10 'INFORME SOBRE BONDAD DE AJUSTE Y VALIDACIÓN DE  
ALGORITMOS' – SUBSANACIÓN  
PROYECTO 'TWINFOREST: GEMELO DIGITAL FORESTAL PARA CLASIFICAR USOS DEL  
SUELO MEDIANTE LA COMBINACIÓN DE VISIÓN ARTIFICIAL, APRENDIZAJE PROFUNDO  
Y SENSORES REMOTOS'

**CÓDIGO PROYECTO:**  
TSI-100909-2019-62

**FECHA:**  
16/05/2022



## LISTA DE ACRÓNIMOS Y ABREVIATURAS

ACRÓNIMO ABREVIATURA	DESCRIPCIÓN
<b>ALS</b>	LiDAR aerotransportado
<b>API</b>	Application Programming Interface
<b>CNIG</b>	Centro Nacional de Información Geográfica
<b>E</b>	Exactitud
<b>F1</b>	Precisión global
<b>FN</b>	Falsos Negativos
<b>FP</b>	Falsos Positivos
<b>föra</b>	Fora forest technologies, S.L.L.
<b>GIS</b>	Sistemas de Información Geográfica
<b>GSV</b>	Google Street View
<b>IGN</b>	Instituto Geográfico Nacional
<b>IR</b>	Infrarrojo
<b>ITD</b>	Individual Tree Detection
<b>LiDAR</b>	Light Detection and Ranging
<b>MDAV</b>	Modelo Digital de Altura de Vegetación
<b>MDS</b>	Modelo Digital de Superficie
<b>MDT</b>	Modelo Digital del Terreno
<b>ML</b>	Machine Learning
<b>p</b>	Precisión
<b>PNOA</b>	Plan Nacional de Ortofotografía Aérea
<b>r</b>	Exhaustividad
<b>SIGPAC</b>	Sistema de Información Geográfica de la Política Agraria Común
<b>SIOSE</b>	Sistema de Información sobre Ocupación del Suelo de España
<b>TWINFOREST</b>	Gemelo digital forestal para clasificar usos del suelo mediante la combinación de visión artificial, aprendizaje profundo y sensores remotos
<b>VC</b>	Visión Computacional
<b>VLS</b>	LiDAR transportado en un vehículo
<b>VP</b>	Verdaderos Positivos

## ÍNDICE

LISTA DE ACRÓNIMOS Y ABREVIATURAS.....	1
1. INTRODUCCIÓN .....	3
2. DATOS DE PARTIDA.....	4
3. METODOLOGÍA.....	6
4. ZONA DE ESTUDIO Y VALIDACIÓN DE RESULTADOS.....	8
5. ESTADÍSTICOS DE CONTRASTE PARA ANALIZAR LOS RESULTADOS.....	10
6. PILOTO 1. POSICIONADO DE ÁRBOLES PARTIENDO DE IMÁGENES A PIE DE CALLE DE GSV .....	11
6.1 Obtención de Imágenes de GSV .....	11
6.2 Detección de objetos.....	12
6.3 Geolocalización .....	13
6.4 Agrupación de puntos .....	14
6.5 Resultados.....	15
7. PILOTO 2. Posicionado de árboles partiendo de nube de puntos LiDAR.....	16
7.1 Procesado de puntos LiDAR.....	16
7.2 Detección de copas y ápices .....	16
7.3 Agrupamiento de puntos .....	17
7.4 Resultados obtenidos .....	18
8. FILTRADO DE PUNTOS.....	18
8.1 Filtrado por GSV.....	19
8.2 Filtrado por Ortofoto (ML).....	22
9. RESUMEN DE LOS RESULTADOS OBTENIDOS.....	23
10. DISCUSIÓN Y CONCLUSIONES Y PROPUESTAS DE MEJORA .....	24
11. BIBLIOGRAFÍA.....	27

## 1. INTRODUCCIÓN

El objetivo principal del proyecto TWINFOREST es desarrollar una metodología capaz de detectar y clasificar usos del suelo a partir de imágenes procedentes de Google Street View (GSV), basada en técnicas de visión artificial, segmentación de instancias y aprendizaje profundo, y que sea capaz de combinarse con datos procedentes de sensores remotos para mejorar los productos de evaluación de recursos forestales ya existentes en föra.

Durante el transcurso del proyecto nos hemos encontrado con que en espacios abiertos no existe mucha información que se pueda obtener de GSV, ya que sus imágenes actualmente están vinculadas a las carreteras y en espacios naturales hay muy poca densidad de carreteras. En cambio, en entornos urbanos sí que se puede obtener todo el potencial de las imágenes de GSV, al tener muchas imágenes incluso para poder abordar la idea de realizar un inventariado del arbolado de los viarios de los núcleos urbanos.

Tomando como válido este cambio de entorno, nos vemos obligados a desechar la herramienta de clasificación de usos de suelo ya que se han encontrado fuentes de datos oficiales más precisas para hacer esta clasificación de usos del suelo dentro de núcleos urbanos (Catastro, SIOSE, SIGPAC, etc.).

Dentro del proyecto TWINFOREST se han desarrollado herramientas para obtener estos datos de manera automática y adaptarlos así a nuestras necesidades.

En el entregable E9-Base datos arbolado urbano con verdad terreno, se ha creado una base de datos de los árboles urbanos validados de manera semiautomática con información obtenida de las imágenes de GSV partiendo de listados de árboles georreferenciados *opensource*. Esta base de datos está en formato *shapefile* y los datos se puede visualizar con cualquier programa GIS.

En este informe se detalla la metodología empleada y la validación de los resultados obtenidos.

## 2. DATOS DE PARTIDA

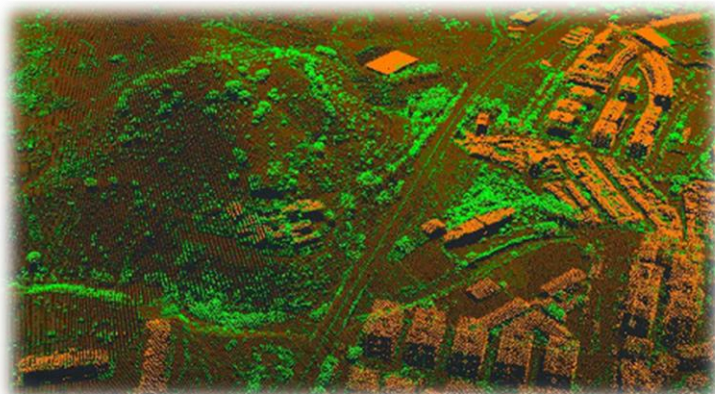
En este caso, nuestro objetivo será la detección y posicionado de los árboles que se encuentran dentro de un entorno urbano.

Nuestra metodología está limitada por la calidad de los datos procedentes de sensores remotos que utilizamos como datos de partida. En nuestro caso vamos a utilizar las siguientes fuentes de datos:

- Ortofotografías aéreas: Son fotografías geo-referenciadas que nos podemos descargar del PNOA (Plan Nacional de Ortofotografías Aérea) del Instituto Geográfico Nacional (IGN). Son imágenes con resolución de 25 a 50 cm con un período de actualización de 2 a 3 años, según las zonas.



- Datos LiDAR (*Laser Imaging Detection and Ranging*) es un dispositivo que permite determinar la distancia desde un emisor a un objeto o superficie utilizando un haz láser pulsado. Con estos dispositivos que van montados en drones o aviones en nuestro caso, obtenemos una nube de puntos X,Y,Z de cada punto del terreno. La densidad de puntos es de 0,5-4 puntos/m<sup>2</sup>, existiendo excepciones en las que la densidad es aún mayor. Los datos se distribuyen de manera gratuita a través del Centro de Descarga del CNIG en ficheros digitales de 2x2 km de extensión.



- Imágenes a pie de calle: Son imágenes que podemos obtener nosotros mismos o podemos descargar de GSV, OpenStreets, Bing maps, etc. Estas imágenes además de estar geo-referenciadas y tendremos que conocer la orientación e inclinación. En nuestro caso estamos utilizando la fuente de GSV con su API correspondiente.



- Imágenes multi-espectrales. En este caso dentro del espectro del infrarrojo. El PNOA también captura imágenes en infrarrojo que utiliza para clasificar los mapas de nubes de puntos LiDAR. Estas imágenes tienen un coste asociado.

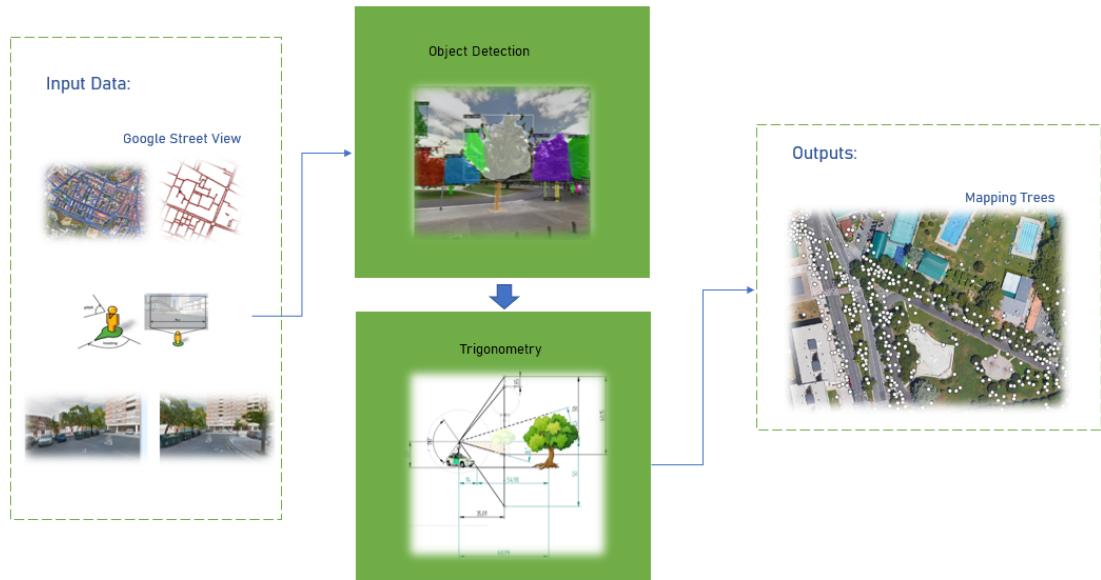




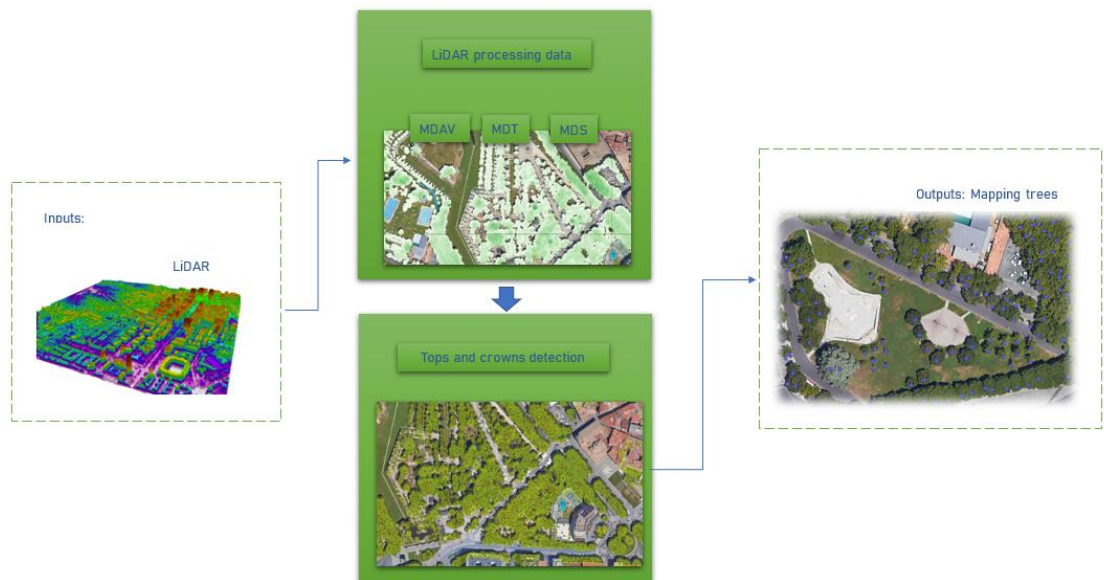
### 3. METODOLOGÍA

Para llevar a cabo nuestro objetivo hemos utilizado dos caminos diferentes en función de la fuente de datos empleada como punto de partida:

- Posicionado de árboles partiendo de imágenes a pie de calle de GSV:



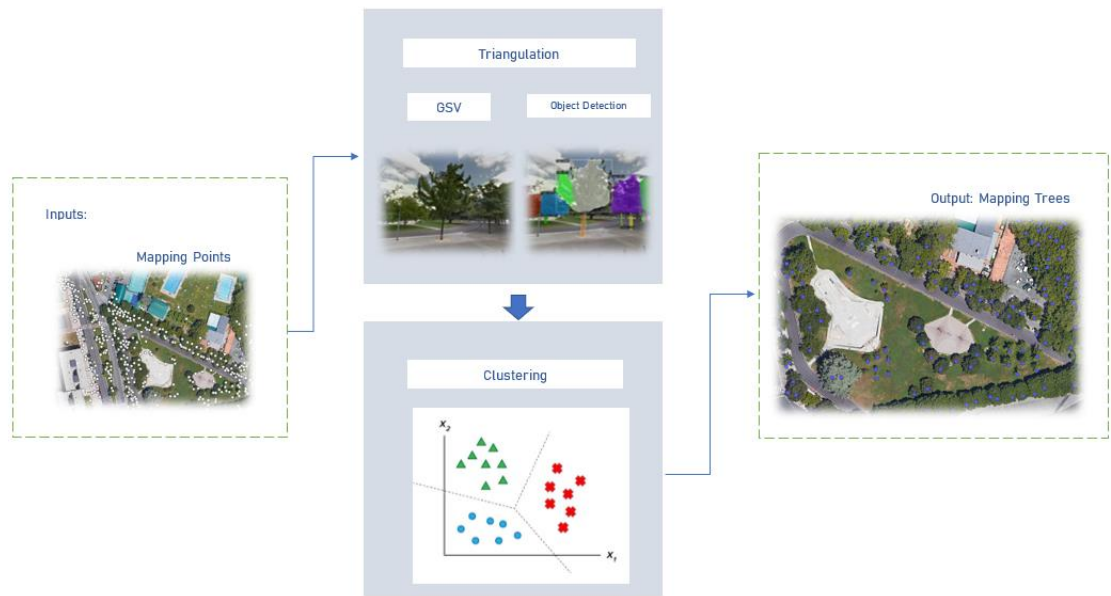
- Posicionado de árboles partiendo de nube de puntos LiDAR:



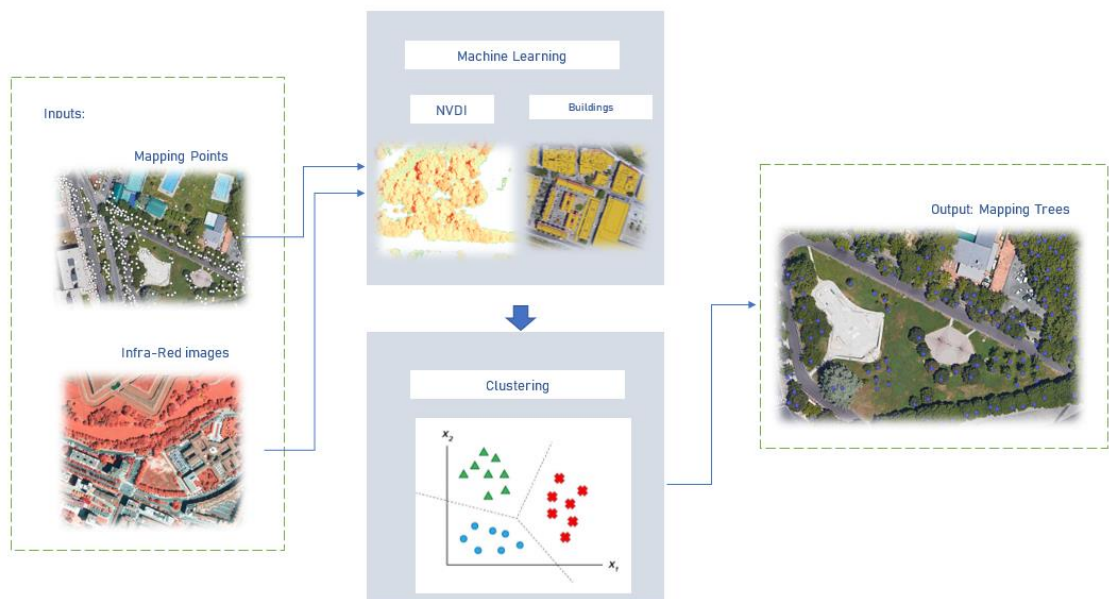


Una vez obtenida la posición de los árboles obtenemos muchos falsos positivos que filtraremos usando los siguientes métodos en función de los datos que tengamos disponibles:

- **Filtrado de árboles partiendo de imágenes de GSV**, realizando detección de troncos en las imágenes y filtrando por desviaciones.



- **Filtrado de árboles partiendo de ortofotos de imágenes aéreas y LiDAR**, realizando un *machine learning* entrenando un modelo a nivel de píxel.



#### 4. ZONA DE ESTUDIO Y VALIDACIÓN DE RESULTADOS

La zona de estudio donde apliquemos nuestra metodología será la ciudad de Pamplona.

Esta ciudad ha sido seleccionada porque dispone de un LiDAR con una precisión de 2 puntos por metro cuadrado, también dispone de cobertura de GSV y de orto imágenes del PNOA de varios años.

La validación de los resultados la hemos realizado con la base de datos del inventario municipal realizado a pie de calle. Analizando la base de datos del arbolado de Pamplona hemos comprobado que no tienen posicionados todos los árboles correctamente y hemos tenido que reposicionar y en ocasiones crear nuevos árboles. Este trabajo lo hemos realizado fotointerpretando manualmente ortofotografías del PNOA.

Hemos creado varios círculos de validación de 300 m de radio con los que podremos comparar los resultados obtenidos. Estos círculos los hemos elegido de manera que sean representativos de distintas topologías de zonas urbanas.



Hacemos una comprobación de los resultados obtenidos con una capa de árboles de control donde previamente se han revisado y añadido /quitado manualmente los árboles. Comparando y sacando distancias obtenemos unos resultados con una métrica numérica.



Para la validación hemos seguido la siguiente nomenclatura:

- Puntos blancos (Arboles de control), son los árboles de control que vamos a utilizar como verdad-terreno.
- Puntos verdes (OK), son los árboles detectados que hemos posicionado correctamente
- Puntos naranjas (Lejanos), son árboles que hemos detectado pero que se encuentran a una distancia significativa del árbol de control más cercano.
- Puntos rojos (False Positivos), son árboles que hemos detectado que no se corresponden con ningún árbol de control o ya hay un árbol detectado que se encuentra más cerca del árbol de control.
- Puntos salmón (False Negativos), son árboles de control que no hemos sido capaces de detectar.

El criterio para la validación para cada árbol detectado es el siguiente:

- Primero buscamos el árbol de control más cercano al árbol detectado
- Buscamos reciprocidad, es decir que el punto de control tiene como punto más cercano, nuestro punto inicial. Si no es recíproco nuestro punto es falso positivo
- Calculamos la distancia entre el punto inicial y el punto de control.
- Si la distancia es menor que un valor de control (3 m) punto OK, en caso de ser mayor punto (lejano).
- Los puntos de control que no tengan asociado un árbol serán nuestros False Negativos.

## 5. ESTADÍSTICOS DE CONTRASTE PARA ANALIZAR LOS RESULTADOS

Para explorar la influencia de los distintos métodos, se ha llevado a cabo una evaluación del rendimiento en términos de tasa de error relativa. Esta evaluación de la precisión se ha basado en la evaluación de medidas estadísticas basadas en los datos de Verdaderos Positivos (VP, detección correcta, árboles bien detectados), Falsos Negativos (FN, error de omisión o árboles no detectados) y Falsos Positivos (FP, error de comisión o árboles detectados que no son árboles). Con estos datos se han construido los siguientes estadísticos: exhaustividad ( $r$ ), precisión ( $p$ ), precisión global (F1) y exactitud ( $E$ ). Estos estadísticos son ampliamente usados para evaluar el error en detección de árboles individuales (Goutte and Gaussier 2005; Sokolova et al. 2006; Mohan et al. 2019). La exhaustividad nos da una medida de los árboles detectados y está inversamente relacionado con el error de omisión, la precisión implica una medida de los árboles correctamente detectados y está inversamente relacionada con el error de comisión, y el F1 nos permite combinar precisión y exhaustividad en un único valor a través de una modificación de su media. La formulación de los estadísticos se muestra a continuación:

$$r = \frac{VP}{VP + FN}$$
$$p = \frac{VP}{VP + FP}$$
$$F1 = \frac{2 \times r \times p}{r + p}$$

La precisión es una gran medida cuando los datos son simétricos (número parecido de FP y FN), y en donde ambos errores tienen la misma influencia. En nuestro caso los FP tiene menor influencia ya que en la fase de depurado se busca minimizarlos. En el caso en que la influencia sea diferente, es mejor evaluar la F. Por otra parte, la exhaustividad hace referencia a lo seguro que estás de que no perder VP, es decir al número de VP respecto a todos los árboles reales detectables. Así la exhaustividad es el estadístico más importante cuando no te importa tener más FP a costa de tener menos FN. Ese es nuestro caso antes de la depuración, ya que se busca tener el máximo de ápices y minimizar los FN aunque todo ello implique tener muchos FP que después tendremos que depurar. Para este trabajo se busca un equilibrio entre los tres estadísticos



## 6. PILOTO 1. POSICIONADO DE ÁRBOLES PARTIENDO DE IMÁGENES A PIE DE CALLE DE GSV

Se pretende obtener la posición de los árboles tomando como datos de partida las imágenes de GSV. Para cada imagen se detectarán y segmentarán los árboles, se estimará su posición en función de los ángulos de la cámara y se eliminarán repetidos por triangulación y se agruparán los puntos cercanos.

Esta metodología solamente funciona para el posicionado de árboles en viales que es de donde podemos obtener imágenes. Para poder posicionar árboles lejanos a estos viales tendríamos que obtener las imágenes por otros medios.

### 6.1 Obtención de Imágenes de GSV

Partimos de las imágenes de GSV. Estas imágenes están asociadas a un identificador llamado PanoID. Cada PanoID tiene asociada una imagen panorámica de 360°. Para cada calle de un entorno urbano hay varios PanoID separados cada 10-15 m. Lo que vamos a hacer será descargarnos las imágenes a ambos lados de la calle y detectar los árboles que aparezcan en las imágenes para más adelante posicionar los árboles detectados.

Los pasos que se seguirán serán los siguientes:

1. Seleccionamos el entorno urbano donde queremos obtener los árboles.
2. Obtenemos el mapa de cobertura de GSV.
3. Obtenemos los puntos a lo largo de las líneas del mapa anterior



4. Para cada uno de los puntos del mapa obtenemos el PanoID más cercano. El listado de puntos con su PanoID lo guardamos en un archivo csv.

5. Para cada punto del listado descargamos las imágenes de GSV a ambos lados (0+90, 0-90).



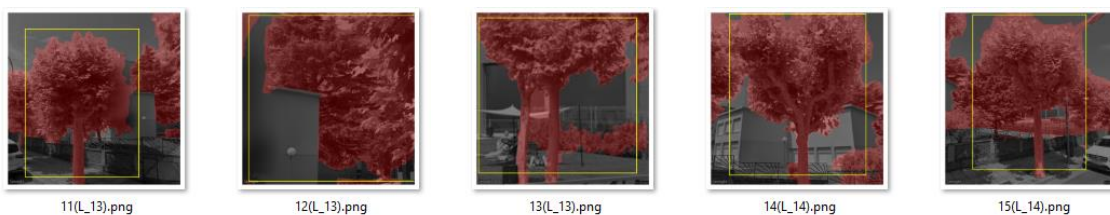
Cada imagen lleva asociada como metadatos, las coordenadas de la cámara donde se realizó la imagen, el ángulo respecto del Norte (heding), el ángulo de inclinación (pich) y el ángulo de captura de la imagen (fov: field of view).



## 6.2 Detección de objetos

Para la detección de objetos utilizamos redes neuronales que previamente hemos entrenado con un montón de imágenes segmentadas manualmente. (Entregable E5-Base de datos de imágenes GSV anotadas y etiquetadas).

Partimos de un modelo ya entrenado que es capaz de segmentar vegetación en imágenes y hemos hecho un *fine tuning* en las últimas capas para reentrenarlo a segmentar árboles con nuestras imágenes.



También hemos creado un modelo capaz de segmentar las partes de un árbol (Tronco y Copa) Este modelo está pensado para poder sacar métricas de los árboles y añadir esos datos a nuestra base de datos.



Para cada una de las imágenes obtenidas en el paso anterior vamos a pasar el modelo de detección de troncos y copas. Cada tronco detectado en la imagen será un punto que añadiremos a nuestro listado de árboles.

Las coordenadas de cada punto las obtenemos gracias a los metadatos de las imágenes de GSV, ya que para cada imagen sabemos, las coordenadas del punto donde se ha tomado la imagen, el ángulo de giro horizontal y vertical y el zoom de la imagen. Por trigonometría podemos obtener la dirección a la que se encuentra el árbol, pero no sabemos la profundidad (distancia) a la que se encuentra

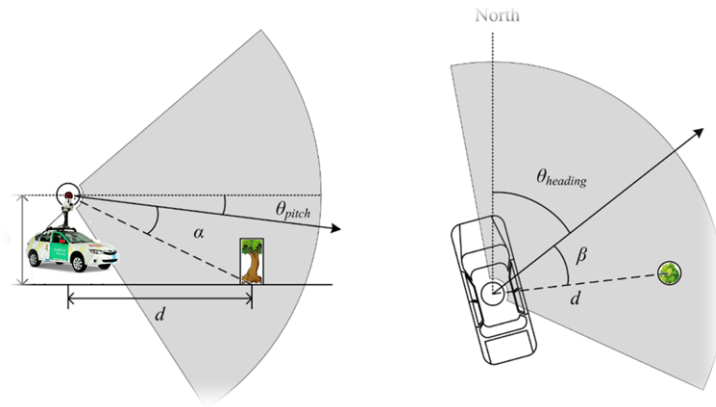
### 6.3 Geolocalización

Sabiendo el ángulo y punto de partida si obtenemos la distancia respecto a la posición de partida, podemos calcular la posición de cada tronco detectado.

Para ello tenemos que localizar el punto de la imagen donde el tronco está en contacto con el suelo, por trigonometría podremos saber a qué distancia se encuentra ese punto si sabemos los siguientes parámetros que nos proporciona la aplicación de GSV:

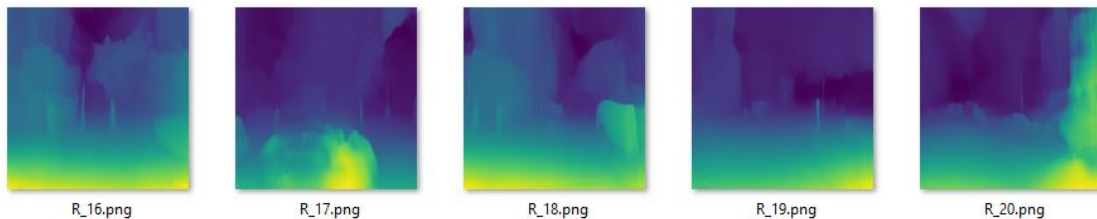
- Altura de la cámara
- Campo de visión [Grados]
- Inclinación [Grados]





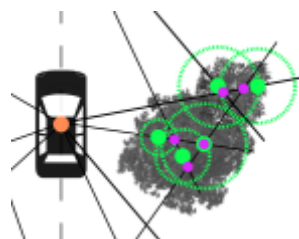
Se ha observado que hay veces que hemos obtenido resultados no satisfactorios debido a que en la imagen no se pudo ver el comienzo del árbol con el suelo debido a que hay otros objetos que lo ocultan. Aunque se ha entrenado a la red neuronal para que haga una predicción en caso de que el tronco esté oculto, no funciona correctamente para todos los casos.

Hemos probado otras técnicas de geoposicionamiento basado en redes neuronales que calculan la profundidad en imágenes monoculares, pero los resultados no son muy válidos porque dan unos resultados de profundidad relativos para cada imagen y no podemos sacar un valor promedio general para estimar la distancia. Entendemos que es un proceso que está en evolución constante y posiblemente en un futuro se pueda utilizar y mejoraremos los resultados.



## 6.4 Agrupación de puntos

Como tenemos muchos árboles marcados donde no hay (Falsos Positivos), vamos a filtrar esos puntos para eliminar puntos duplicados



Para ello aplicaremos dos un filtro que elimina los puntos que parten desde el mismo coche de Google y tienen el mismo ángulo  $\pm 10^\circ$ .

También agruparemos los puntos que se encuentran en un radio de 3 m dejando solamente el principal.

## 6.5 Resultados

Nos hemos dado cuenta de que con este método podemos posicionar el 85% de los árboles viarios de una zona determinada. Pero tenemos un error muy grande.

Los resultados los hemos validado para zona viaria. No tiene sentido utilizar este método para posicionar árboles fuera de zona viaria.

PILOTO 1. GSV (ZONA VIARIO)	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1	Accuracy
Zona 1	700	635	911	51	41,07%	92,57%	56,90%	39,76%
zona 2	215	112	351	73	24,19%	60,54%	34,57%	20,90%
Zona 3	652	461	420	177	52,33%	72,26%	60,70%	43,57%
Zona 6	366	306	345	44	47,00%	87,43%	61,14%	44,03%
Zona 8	839	759	1228	54	38,20%	93,36%	54,21%	37,19%
<b>Todas las zonas</b>	<b>2.772</b>	<b>2.273</b>	<b>3.255</b>	<b>399</b>	<b>41,12%</b>	<b>85,07%</b>	<b>55,44%</b>	<b>38,35%</b>

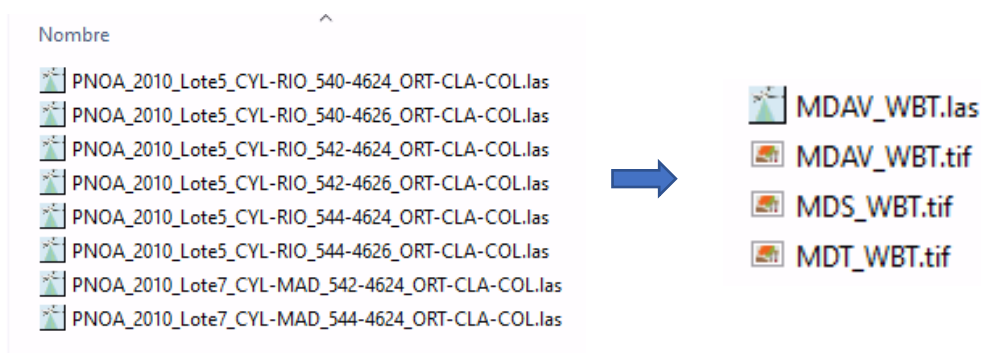


## 7. PILOTO 2. Posicionamiento de árboles partiendo de nube de puntos LiDAR

Se pretende posicionar los árboles urbanos de una zona partiendo del mapa de puntos LiDAR. Se procesará la nube de puntos para obtener el Modelo de altura de vegetación (MDAV), Modelo de altura del terreno (MDT), Modelo de altura de superficie (MDS). Procesando estos modelos con el algoritmo de ForestTools detectamos las copas y ápices de los árboles geolocalizados en formato GIS.

### 7.1 Procesado de puntos LiDAR

Descargamos los datos de partida desde el centro de descargas de CNIG. Descomprimos los datos y obtenemos los distintos modelos de superficies utilizando herramientas de procesamiento de datos LiDAR (Fusion, whiteBoxTools)



### 7.2 Detección de copas y ápices

Aplicamos el algoritmo para encontrar ápices y copas en los puntos LiDAR. Este algoritmo se basa en la búsqueda de máximos zonales y análisis de las pendientes de la nube de puntos.



Una vez procesado el algoritmo obtenemos el siguiente resultado:



Aunque parece que el resultado es bueno, analizando los datos de validación hemos observado que hay bastantes False Positivos, es decir, estamos marcando árboles donde no los hay. Esto depende mucho de la calidad y de los datos LiDAR, para ciudades con más densidad de puntos obtenemos mucho mejores resultados. Esto es una cosa que va a ir mejorando con el tiempo, cada día van mejorando la tecnología con la que se obtienen estos datos y suponemos que la densidad de puntos y la periodicidad de los datos mejorará exponencialmente y solo es cuestión de tiempo poder obtener un inventariado del arbolado urbano con unos errores muy bajos.

Nos hemos dado cuenta de que para no perder puntos válidos es mejor poner unos parámetros que limiten poco el algoritmo y obtener más puntos, aunque tengamos más False Positivos y utilizar otros métodos de filtrado para eliminarlos más adelante.

### 7.3 Agrupamiento de puntos

Nos hemos descargado la capa de edificios del catastro y hemos eliminado los puntos que se encontraban dentro de éstos.

También hemos agrupado los árboles que estaban a menos distancia de 3 metros, para bajar la tasa de error.

## 7.4 Resultados obtenidos

Nos hemos dado cuenta de que con este método podemos posicionar el 80% de los árboles viarios de una zona determinada. Pero tenemos un error elevado.

PILOTO 2. LiDAR	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
Zona 1	734	638	703	67	47,58%	90,50%	62,37%
zona 2	2.098	1633	1746	356	48,33%	82,10%	60,84%
Zona 3	1.049	851	1586	107	34,92%	88,83%	50,13%
Zona 6	1.610	1388	1444	172	49,01%	88,97%	63,21%
Zona 8	1.491	1329	1168	132	53,22%	90,97%	67,16%
<b>Todas las zonas</b>	<b>5.491</b>	<b>4.510</b>	<b>5.479</b>	<b>702</b>	<b>45,15%</b>	<b>86,53%</b>	<b>59,34%</b>

PILOTO 2. LiDAR (VIARIO)	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
<u>Zona 1</u>	700	608	655	64	48,14%	90,48%	62,84%
zona 2	215	164	205	36	44,44%	82,00%	57,64%
Zona 3	652	490	367	132	57,18%	78,78%	66,26%
Zona 6	366	341	314	20	52,06%	94,46%	67,13%
Zona 8	839	772	718	50	51,81%	93,92%	66,78%
<b>Todas las zonas</b>	<b>1.933</b>	<b>1.603</b>	<b>1.541</b>	<b>252</b>	<b>50,99%</b>	<b>86,42%</b>	<b>64,13%</b>

PILOTO 2. LiDAR (NO VIARIO)	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
Zona 1	10	7	38	1	15,56%	87,50%	26,42%
zona 2	1.714	1352	1169	283	53,63%	82,69%	65,06%
Zona 3	378	314	291	49	51,90%	86,50%	64,88%
Zona 6	1.195	995	577	169	63,30%	85,48%	72,73%
Zona 8	655	530	480	106	52,48%	83,33%	64,40%
<b>Todas las zonas</b>	<b>3.297</b>	<b>2.668</b>	<b>2.075</b>	<b>502</b>	<b>56,25%</b>	<b>84,16%</b>	<b>67,43%</b>

## 8. FILTRADO DE PUNTOS

Vamos a mostrar distintos métodos para eliminar puntos no deseados y bajar el error obtenido sin perder porcentaje de árboles detectados.

Para poder comparar estos métodos de filtrado, vamos a comparar solamente los resultados obtenidos para zonas viarias, ya que los métodos que se basan en imágenes de GSV solamente se pueden aplicar para estas zonas.

Realizaremos un filtrado por GSV y otro por ortofoto (ML) y analizaremos los resultados.

## 8.1 Filtrado por GSV

Hemos pensado que, si nos descargamos tres imágenes de GSV para cada punto y pasamos el modelo de detección de árboles para cada imagen, nuestro árbol se debería encontrar en el centro de cada una de las imágenes.

- Buscamos los PanoID más cercanos de cada punto  
Previamente nos hemos descargado todos los PanoID de la zona. Para cada punto del listado de resultados, buscamos los 3 PanoID más cercanos.



- Obtenemos las imágenes de GSV  
Nos descargamos las imágenes de cada punto del listado desde cada PanoID obteniendo 3 imágenes del árbol desde 3 posiciones distintas.



1.jpg



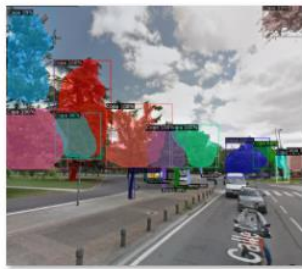
2.jpg



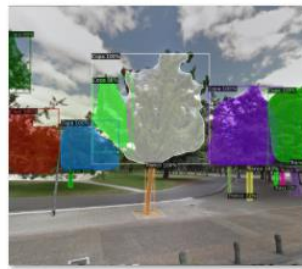
3.jpg

- Detectamos troncos y copas

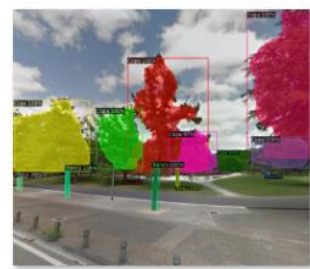
Detectamos troncos y copas de cada una de las imágenes y nos obtenemos con las coordenadas de cada recuadro.



Detectron\_1.jpg



Detectron\_2.jpg



Detectron\_3.jpg

Estimamos la distancia a la que se encuentra cada objeto, por trigonometría en función de los ángulos de la imagen

Para cada imagen sabemos la distancia real a la que se encuentra el punto respecto al PanoID, y la distancia estimada que hemos calculado para cada objeto de la imagen.

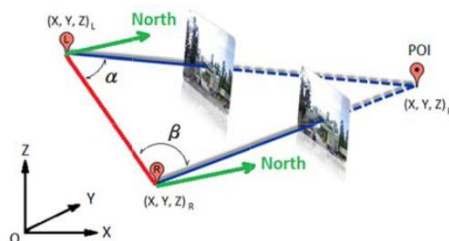
Asignamos un peso en función de la diferencia entre distancias y la posición del objeto detectado respecto del centro de la imagen.

- Filtrado por OffCenter:

Establecemos un valor offcenter que es la distancia desde el centro del cuadro donde se ha detectado un tronco respecto al centro de la imagen. Si el árbol está bien posicionado el offcenter tendría que ser próximo a cero. Hacemos un filtrado de los puntos eliminando todos los puntos que no tienen al menos una imagen con un offcenter  $< 0.05$  o dos imágenes con un offcenter  $< 0.1$

- Filtrado por intersección:

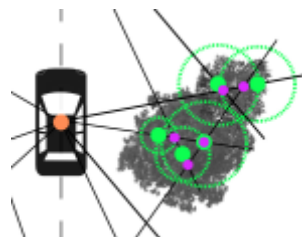
Para cada conjunto de 3 imágenes nos quedamos con la que tiene mejor offcenter, antes modificábamos las coordenadas por la intersección en función de desplazar el ángulo del punto por el desplaz\_x





- Filtrado por paralelos:

Para cada punto sabemos cuál es el origen de la imagen (Coche de Google) y por tanto podemos agrupar todos los puntos con el mismo origen y un ángulo cercano ( $\pm 2^\circ$ ). Además, establecemos como condición que la distancia respecto al origen sea parecida ( $\pm 2\text{m}$ )



Para ello aplicaremos dos un filtro que elimina los puntos que parten desde el mismo coche de Google y tienen el mismo ángulo  $\pm 10^\circ$

También agruparemos los puntos que se encuentran en un radio de 3 m dejando solamente el principal.

- Agrupado de puntos cercanos (4 m)

Todos los puntos que se encuentran a una distancia menor de 4 m los agrupamos. Entendemos que no puede haber árboles tan juntos. Utilizaremos de punto de referencia el punto que tenga mejor valoración en los algoritmos anteriores.

PILOTO 1. GSV (Filtrado por GSV)	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
Zona 1	700	581	271	95	68,19%	85,95%	76,05%
zona 2	215	91	104	104	46,67%	46,67%	46,67%
Zona 3	652	409	137	227	74,91%	64,31%	69,20%
Zona 6	366	282	97	75	74,41%	78,99%	76,63%
Zona 8	839	707	358	105	66,38%	87,07%	75,33%
<b>Todas las zonas</b>	<b>2.772</b>	<b>2.070</b>	<b>967</b>	<b>606</b>	<b>68,16%</b>	<b>77,35%</b>	<b>72,47%</b>

PILOTO 2. LiDAR (Filtrado por GSV)	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
Zona 1	700	542	146	141	78,78%	79,36%	79,07%
zona 2	215	102	66	103	60,71%	49,76%	54,69%
Zona 3	652	421	87	219	82,87%	65,78%	73,34%
Zona 6	366	308	82	52	78,97%	85,56%	82,13%
Zona 8	839	707	270	113	72,36%	86,22%	78,69%
<b>Todas las zonas</b>	<b>2.772</b>	<b>2.080</b>	<b>651</b>	<b>628</b>	<b>76,16%</b>	<b>76,81%</b>	<b>76,48%</b>

## 8.2 Filtrado por Ortofoto (ML)

Tomamos como fuente de datos, las imágenes georreferenciadas que hemos obtenido del PNOA del último año y del anterior. También utilizaremos la imagen de infrarrojo (IR) que hemos adquirido del PNOA. Para el piloto 2 utilizaremos además los datos LiDAR procesados, estos datos no los vamos a usar para el piloto 1 porque entendemos que el piloto 1 lo emplearemos para sitios donde no dispongamos de una fuente de datos LiDAR.

Para poder realizar este filtrado, previamente manualmente hemos seleccionado un conjunto de puntos representativos (Verdad-Terreno) y con ayuda de las imágenes georreferenciadas hemos clasificado manualmente cada punto como árbol / no árbol.

Con todos estos puntos de verdad terreno vamos a componer una tabla de datos a la que añadiremos columnas con la información a nivel de pixel de cada una de las capas que componen las imágenes (Red, Green, Blue, IR), con toda esta información entrenaremos un modelo de machine Learning de clasificación (árbol / no árbol) en función de la información obtenida a nivel de pixel de cada una de las capas.

Posteriormente hemos aplicado el modelo a los puntos de cada de los resultados de los dos pilotos y eliminaremos aquellos puntos clasificados como “no árbol” y hemos obtenido los siguientes resultados:

PILOTO 1. GSV (Filtrado por ML)	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
Zona 1	700	581	304	95	65,65%	85,95%	74,44%
zona 2	215	98	120	92	44,95%	51,58%	48,04%
Zona 3	652	421	142	221	74,78%	65,58%	69,88%
Zona 6	366	278	102	77	73,16%	78,31%	75,65%
Zona 8	839	711	389	98	64,64%	87,89%	74,49%
<b>Todas las zonas</b>	<b>2.772</b>	<b>2.089</b>	<b>1.057</b>	<b>583</b>	<b>66,40%</b>	<b>78,18%</b>	<b>71,81%</b>

PILOTO 2. LiDAR (Filtrado por ML)	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
Zona 1	700	509	51	186	90,89%	73,24%	81,12%
zona 2	215	141	68	66	67,46%	68,12%	67,79%
Zona 3	652	393	55	251	87,72%	61,02%	71,98%
Zona 6	366	298	76	61	79,68%	83,01%	81,31%
Zona 8	839	612	207	217	74,73%	73,82%	74,27%
<b>Todas las zonas</b>	<b>2.772</b>	<b>1.953</b>	<b>457</b>	<b>781</b>	<b>81,04%</b>	<b>71,43%</b>	<b>75,93%</b>

## 9. RESUMEN DE LOS RESULTADOS OBTENIDOS

Insistimos en que para poder hacer comparables los resultados de los métodos de detección de árboles descritos, centraremos los resultados solo en las zonas viarias, ya que no disponemos de cobertura de GSV en zonas no transitadas. Los resultados obtenidos son muy satisfactorios, identificando más del 80% de los árboles viales. Una vez realizada la depuración de FP, esta identificación baja al 75% al perder también VP al depurar FP.

En la siguiente tabla se muestran los resultados obtenidos en cada uno de los métodos evaluados. Se muestran por separado los resultados de cada una de las zonas y los resultados globales (denominado Zona “Todos”).

### Filtrado GSV

PILOTO 1. GSV	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
Todas las zonas	2.772	2070	967	606	68,16%	77,35%	72,47%

PILOTO 2. LiDAR	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
Todas las zonas	2.772	2080	651	628	76,16%	76,81%	76,48%

### Filtrado Machine Learning (ML)

PILOTO 1. GSV	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
Todas las zonas	2.772	2089	1057	583	66,40%	78,18%	71,81%

PILOTO 2. LiDAR	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
Todas las zonas	2.772	1953	457	781	81,04%	71,43%	75,93%

### Filtrado GSV y ML

PILOTO 1. GSV	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
Todas las zonas	2.772	2001	619	711	76,37%	73,78%	75,06%

PILOTO 2. LiDAR	Ground True	True Pos. (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1
Todas las zonas	2.772	2049	557	668	78,63%	75,41%	76,99%

Si nos fijamos en la exhaustividad, un mayor valor indica que tenemos más verdaderos positivos, sin importar los falsos positivos que encontramos. En

general hemos encontrado un valor muy parecido tanto si abordamos la ITD a través de GSV como a través de LiDAR, aunque generalmente encontramos más FP al utilizar GSV. Por tanto, si solo realizamos la primera etapa, siempre detectamos más VP que combinando etapas. Mientras que los resultados obtenidos solo con GSV o con ALS son similares, pero ligeramente superiores al usar solo ALS.

Teniendo en cuenta la segunda etapa, para depurar FP son tan válidos los métodos basados en GSV como en ML. La combinación de etapas, (ITC y depuración de FP) es algo mejor cuando partimos de una ITD realizada con ALS. No obstante, las diferencias no son demasiado significativas. Realizando la ITD con GSV mantenemos unos pocos más VP que cuando usamos ALS.

Finalmente, una cosa que no se ha evaluado en este trabajo es que si nos interesa medir variables relacionadas con la copa del árbol es imprescindible realizar una tercera etapa que consistirá en realizar la segmentación del límite de la copa, que generalmente se basa en algoritmos de región de crecimiento.

## 10. DISCUSIÓN, CONCLUSIONES Y PROPUESTAS DE MEJORA

Los avances en Visión Computacional (VC) están permitiendo obtener de forma automática información robusta basada en datos a nivel de calle (He et al. 2017). Estos algoritmos han sido desarrollados para la investigación de ciudades inteligentes utilizando imágenes a nivel de calle, y se han aplicado para evaluar demografía (Gebru et al. 2017), cambio urbano (Naik et al. 2015), riqueza (Glaeser et al. 2018), seguridad urbana percibida (Naik et al. 2017), para clasificar tipos de edificios (Kang et al. 2018) y para analizar la morfología urbana (Middel et al. 2019). En el campo de la silvicultura urbana inteligente, las imágenes a pie de calle en combinación con la VC se han aplicado en tres áreas clave: 1) estimación de la provisión de sombra para árboles urbanos (Li et al. 2018; Li and Ratti 2019), 2) cuantificación de la percepción urbana cobertura del dosel (Li et al. 2015; Seiferling et al. 2017; Duarte and Ratti 2018; Cai et al. 2018; Stubbings et al. 2019), y 3) mapeo de la ubicación de los árboles urbanos (Wegner et al. 2016; Branson et al. 2018). Berland and Lange (2017) usaron GSV y obtuvieron un 93% de acierto sobre arbolado urbano y descubrieron que era posible evaluar el género, la especie, la ubicación, el diámetro a la altura del pecho y la salud del árbol. Rousselet et al. (2013) fueron capaces de identificar, con un 96% de éxito, árboles atacados por

procesionaria del pino. Sin embargo, estos estudios y muchos otros no están automatizados y están limitados por costosas labores manuales. Li et al. (2018) calcularon un factor que indica el nivel de cerramiento de los cañones de las calles, para cuantificar la provisión de sombra de los árboles. Seiferling et al. (2017) a partir de GSV y ML cuantificaron la cobertura de dosel de copas urbano. Li et al. (2015) evaluaron el porcentaje de vegetación en las calles cuantificando la cantidad de píxeles verdes vistos en una escena de GSV. Estas metodologías ayudaron a dar forma al llamado *Green View Index* (Duarte and Ratti 2018), que indica cómo de verdes los peatones perciben las calles. Wegner et al. (2016) diseñaron un flujo de trabajo para la detección automática de árboles de calles y la geolocalización a partir de imágenes de GSV y Google Maps, basado en el modelo *Faster Region-based Convolutional Neural Network*.

Las metodologías que basadas en de datos procedentes de datos LiDAR son menos abundantes, pero aun así han sido implementadas y automatizadas en algunas ciudades importantes. Tanhuanpää et al. (2014) fueron capaces de detectar mediante un procedimiento de mapeo automático en la ciudad de Helsinki (Finlandia) el 88,8% de los árboles actuales del inventario de la ciudad, en donde además midieron su altura (con un RMSE de 1,27 m) y su diámetro normal (con un RMSE de 6,9 cm). Holopainen et al. (2011), en una pequeña muestra de 44 árboles, también en Helsinki, determinan que es más preciso identificar esos árboles a partir de un LiDAR transportado en un vehículo (VLS) que con datos LiDAR aerotransportados (ALS). Holopainen et al. (2013), esta vez sobre 438 árboles situados en parques y bosques urbanos vuelven a encontrar que el VLS obtiene mayor eficacia que el ALS (79.22% vs 68.04% de acierto, respectivamente). No obstante, en ambos casos los resultados se apoyan en una metodología manual. En el caso en que se automatiza la identificación de árboles, los aciertos caen hasta el 26.94% y 65.53% respectivamente. Matasci et al. (2018) evalúan el inventario de arbolado en Vancouver con ALS, sobre una muestra de 22.211 árboles, obteniendo una exhaustividad del 76.6% para posteriormente sobre una submuestra de árboles estimar su altura (con RMSE de 2.6m) y el diámetro de su copa (con un RMSE de 3.85m). En Múnich (Alemania), Wu et al. (2018) comparan VLS y ALS, obteniendo un mejor porcentaje de árboles detectados (83.36%) con el ALS, frente a un 77.2% con VLS. Finalmente, Hanssen et al. (2019) en un análisis exhaustivo sobre arbolado urbano de Oslo (Noruega) a través de ALS son capaces de detectar el 73.6% de los árboles. Además, en este mismo estudio

se analiza la cobertura de dichos árboles y se relaciona con la información espectral de Sentinel-2.

En nuestro caso sobre una metodología completamente automatizada conseguimos posicionar en torno al 75% de los árboles, obteniendo unos resultados completamente comparables a los resultados obtenidos por otros autores. Además, en nuestro caso se obtiene una relación entre VP, FN y FP bastante equilibrada. Los resultados anteriores reflejan que ambos métodos muestran unas métricas de precisión similares. En función de la fuente de datos que se disponga (ALS, GSV y/o ortofotos) podemos utilizar una u otra combinación de métodos para realizar un inventario urbano de bajo coste y automatizado. El siguiente paso a la geolocalización del arbolado urbano será la identificación de su familia y especie a través de las propias imágenes de GSV.

A continuación, indicamos algunas propuestas de mejora para situar el error por debajo del 15%:

- **Mejorando la calidad de los datos de partida**, en el transcurso de este proyecto se ha constatado que los resultados obtenidos dependen mucho de la calidad de los datos de partida. En nuestro caso, los datos utilizados son imágenes de Google Street View y nubes de puntos LiDAR del PNOA.

Se ha comprobado que los resultados de ciudades con peores datos LiDAR como puede ser Soria (1 pto/m<sup>2</sup>) son mucho peores que los de ciudades como Pamplona (15 ptos/m<sup>2</sup>). Siguiendo esta lógica se puede pensar que si se utilizan fuentes de datos LiDAR con mejor precisión el error disminuirá de forma significativa. Lo mismo pasaría si se utilizasen imágenes a pie de calle de otras fuentes. Las imágenes de GSV, están tomadas cada 20 m de distancia, si esas imágenes se hicieran cada menos distancia o incluso fueran continuas (videos) se mejorarían los datos. Este sería replicar las técnicas de los vehículos autónomos o equipos de *mobile mapping* que obtienen resultados mucho más precisos.

- **Mejorando los algoritmos utilizados**, hay muchos algoritmos utilizados como pueden ser los de triangulación, los de detección de ápices, o incluso técnicas de visión computacional para

detectar y segmentar objetos por medio de redes neuronales que van mejorando a un ritmo importante. Utilizando las nuevas técnicas que se están desarrollando, los resultados mejorarán de forma significativa.

## 11. BIBLIOGRAFÍA

- Berland, A.; Lange, D.A. Google Street View Shows Promise for Virtual Street Tree Surveys. *Urban Forestry and Urban Greening* 2017, 21, 11–15, doi:10.1016/j.ufug.2016.11.006.
- Branson, S.; Wegner, J.D.; Hall, D.; Lang, N.; Schindler, K.; Perona, P. From Google Maps to a Fine-Grained Catalog of Street Trees. *ISPRS Journal of Photogrammetry and Remote Sensing* 2018, 135, 13–30, doi:10.1016/j.isprsjprs.2017.11.008.
- Cai, B.Y.; Li, X.; Seiferling, I.; Ratti, C. Treepedia 2.0: Applying Deep Learning for Large-Scale Quantification of Urban Tree Cover. In *Proceedings of the Proceedings - 2018 IEEE International Congress on Big Data, BigData Congress 2018 - Part of the 2018 IEEE World Congress on Services*; 2018.
- Duarte, F.; Ratti, C. What Big Data Tell Us about Trees and the Sky in the Cities. In *Humanizing Digital Reality*; Springer, 2018; pp. 59–62.
- Goutte, C.; Gaussier, E. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation.; Losada, D.E., Fernández-Luna, J.M., Eds.; Springer Berlin Heidelberg, 2005; pp. 345–359.
- Hanssen, F.; Barton, D.N.; Nowell, M.; Cimburova, Z. Mapping Urban Tree Canopy Cover Using Airborne Laser Scanning. *Applications to Urban Ecosystem Accounting for Oslo*. 2019.
- He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-Cnn.; 2017; pp. 2961–2969.
- Holopainen, M.; Vastaranta, M.; Kankare, V.; Hyypä, H.; Vaaja, M.; Hyypä, J.; Liang, X.; Litkey, P.; Yu, X.; Kaartinen, H.; et al. The Use of ALS, TLS and VLS Measurements in Mapping and Monitoring Urban Trees. In *Proceedings of the 2011 Joint Urban Remote Sensing Event*; 2011; pp. 29–32.
- Kang, J.; Körner, M.; Wang, Y.; Taubenböck, H.; Zhu, X.X. Building Instance Classification Using Street View Images. *ISPRS Journal of Photogrammetry and Remote Sensing* 2018, 145, doi:10.1016/j.isprsjprs.2018.02.006.
- Li, X.; Ratti, C. Using Google Street View for Street-Level Urban Form Analysis, a Case Study in Cambridge, Massachusetts. In *Modeling and Simulation in Science, Engineering and Technology*; 2019.
- Li, X.; Ratti, C.; Seiferling, I. Quantifying the Shade Provision of Street Trees in Urban Landscape: A Case Study in Boston, USA, Using Google Street View.



- Landscape and Urban Planning 2018, 169, 81–91, doi:10.1016/j.landurbplan.2017.08.011.
- Li, X.; Zhang, C.; Li, W.; Ricard, R.; Meng, Q.; Zhang, W. Assessing Street-Level Urban Greenery Using Google Street View and a Modified Green View Index. *Urban Forestry and Urban Greening* 2015, 14, doi:10.1016/j.ufug.2015.06.006.
- Matasci, G.; Coops, N.C.; Williams, D.A.R.; Page, N. Mapping Tree Canopies in Urban Environments Using Airborne Laser Scanning (ALS): A Vancouver Case Study. *Forest Ecosystems* 2018, 5, doi:10.1186/s40663-018-0146-y.
- Middel, A.; Lukasczyk, J.; Zakrzewski, S.; Arnold, M.; Maciejewski, R. Urban Form and Composition of Street Canyons: A Human-Centric Big Data and Deep Learning Approach. *Landscape and Urban Planning* 2019, 183, doi:10.1016/j.landurbplan.2018.12.001.
- Mohan, M.; Mendonça, B.A.F. de; Silva, C.A.; Klauberg, C.; de Saboya Ribeiro, A.S.; Araújo, E.J.G. de; Monte, M.A.; Cardil, A. Optimizing Individual Tree Detection Accuracy and Measuring Forest Uniformity in Coconut (*Cocos Nucifera* L.) Plantations Using Airborne Laser Scanning. *Ecological Modelling* 2019, 409, doi:10.1016/j.ecolmodel.2019.108736.
- Naik, N.; Kominers, S.D.; Raskar, R.; Glaeser, E.L.; Hidalgo, Cc. Do People Shape Cities, or Do Cities Shape People? The Co-Evolution of Physical, Social, and Economic Change in Five Major U.S. Cities. *SSRN Electronic Journal* 2015, doi:10.2139/ssrn.2698292.
- Rousselet, J.; Imbert, C.-E.; Dekri, A.; Garcia, J.; Goussard, F.; Vincent, B.; Denux, O.; Robinet, C.; Dorkeld, F.; Roques, A. Assessing Species Distribution Using Google Street View: A Pilot Study with the Pine Processionary Moth. *PLoS One* 2013, 8, e74918.
- Seiferling, I.; Naik, N.; Ratti, C.; Proulx, R. Green Streets – Quantifying and Mapping Urban Trees with Street-Level Imagery and Computer Vision. *Landscape and Urban Planning* 2017, 165, doi:10.1016/j.landurbplan.2017.05.010.
- Sokolova, M.; Japkowicz, N.; Szpakowicz, S. Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation.; Sattar, A., Kang, B., Eds.; Springer Berlin Heidelberg, 2006; pp. 1015–1021.
- Stubblings, P.; Peskett, J.; Rowe, F.; Arribas-Bel, D. A Hierarchical Urban Forest Index Using Street-Level Imagery and Deep Learning. *Remote Sensing* 2019, 11, doi:10.3390/rs11121395.
- Tanhuanpää, T.; Vastaranta, M.; Kankare, V.; Holopainen, M.; Hyyppä, J.; Hyyppä, H.; Alho, P.; Raisio, J. Mapping of Urban Roadside Trees—A Case Study in the Tree Register Update Process in Helsinki City. *Urban forestry & urban greening* 2014, 13, 562–570.
- Wegner, J.D.; Branson, S.; Hall, D.; Schindler, K.; Perona, P. Cataloging Public Objects Using Aerial and Street-Level Images—Urban Trees.; 2016; pp. 6014–6023.

Wu, J.; Yao, W.; Polewski, P. Mapping Individual Tree Species and Vitality along Urban Road Corridors with LiDAR and Imaging Sensors: Point Density versus View Perspective. *Remote Sensing* 2018, 10, doi:10.3390/rs10091403.