

Guia Pràctica 5 – Segmentació Semàntica – Part1

Curs 2021-2022 – Dept. Ciències de la Computació, UAB.

1 Context

A les darreres classes hem vist diferents aproximacions per combinar múltiples vistes d'una escena i recuperar-ne l'estructura 3D, així com la posició i orientació relativa de les múltiples càmeres. Aquestes aproximacions són elements constituents del flux de procés que integren els programes de fotogrametria. La fotogrametria és la ciència i tecnologia per obtenir informació fiable sobre objectes físics, i pretén d'alguna manera invertir el procés fotogràfic: mapejar imatges planes 2D al món tridimensional real.

Un dels àmbits d'aplicació principal de la fotogrametria és el procés d'imatges aèries. La fotogrametria permet combinar les vistes captades d'un territori, per generar-ne una ortoimatge i un mapa de superfície. L'ortoimatge (o ortomapa) aporta una vista ortogràfica de l'escena on l'escala és uniforme, fet que simplifica molt el seu procés i anàlisi.

Una de les àrees d'aplicació de les ortoimatges és l'agricultura de precisió. Mitjançant drons es capturen imatges de finques agrícoles usant càmeres multiespectrals, amb l'objectiu de caracteritzar l'estat dels cultius i les seves necessitats. La reflectància captada en bandes espectrals concretes (Red, RedEdge, NIR, ...) està correlacionada amb aspectes de vigor i salut de les plantes, i permet gestionar el cultiu de manera sectoritzada.

En aquesta pràctica es proposa desenvolupar un sistema de visió per computador per localitzar males herbes en cultius. A partir del conjunt de dades anotat del [Remote Sensing 2018 Weed Map Dataset](#) (Fig.1) s'implementaran diferents algorismes de segmentació semàntica per classificar els píxels de diferents ortoimatges segons siguin part part del cultiu, de males herbes, o del sòl.

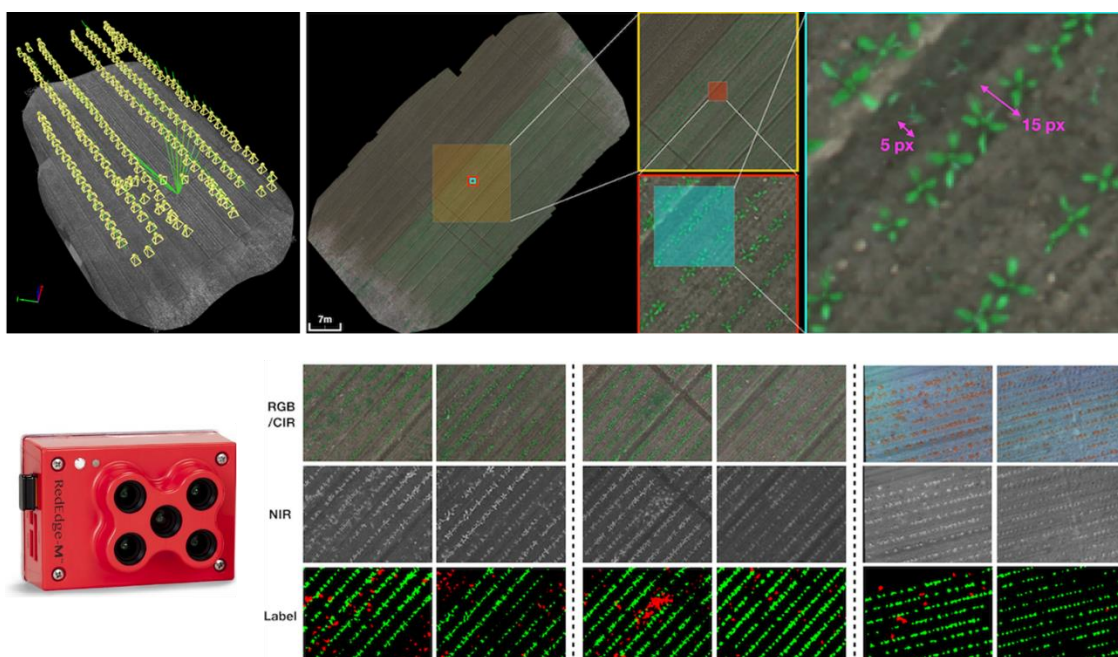


FIGURA 1 WEEDMAP DATASET: EN AQUESTA PRÀCTICA ES TREBALLA AMB LES ORTOIMATGES GENERADES USANT LA CÀMERA MULTIESPECTRAL MICA SENSE REDEGE-M

2 Enunciat

En aquesta pràctica es comença a treballar a partir del Jupyter notebook facilitat en el campus virtual. Aquest notebook aporta algunes classes i funcions per facilitar la interacció amb el conjunt de dades a processar.

L'objectiu és analitzar les característiques del conjunt de dades a processar, i desenvolupar un primer sistema de segmentació semàntica basat en la informació individual de cada píxel, emprant classificadors clàssics (és a dir, no basats en deep learning). El rendiment obtingut es mesurarà amb diferents mètriques, per així establir el rendiment base o de referència que s'haurà de millorar a la segona part de la pràctica.

2.1 Descàrrega del dataset i visualització de dades.

Descarrega i descomprimeix l'arxiu RedEdgeCompact.zip del campus virtual. Veuràs que és força voluminós (205 Mb). Conté les dades de 5 ortoimatges captades en 5 finques diferents, fragmentades en diferents retalls (*tiles*). No obstant, no cal que dediquis temps en veure com el dataset està organitzat perquè en el notebook es donen mecanismes per obtenir-ne informació de manera senzilla.

En el notebook, inicialitza la variable 'dataset_dir' amb el path al directori on has descarregat el dataset i comprova que visualitzes dades del dataset sense problema.



FIGURA 2 IMATGE COLOR A PARTIR DELS CANALS R,G,B, I MÀSCARES DELS 4 TIPUS DE PÍXELS DEL DATASET.

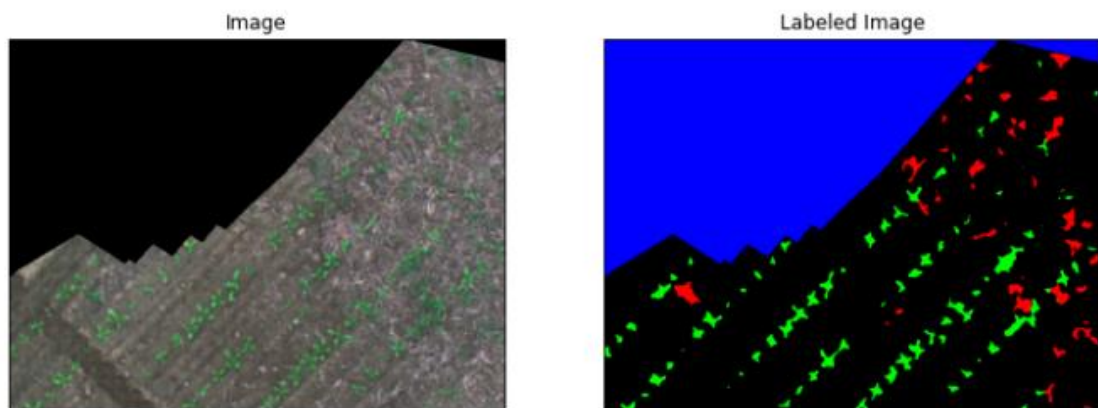


FIGURA 3 IMATGE COLOR A PARTIR DELS CANALS R,G,B I ETIQUETA DE CADA PÍXEL MOSTRADA EN UNA IMATGE COLOR

2.2 Repte 1: Anàlisi dels estadístics del dataset

Abans d'implementar el sistema de segmentació semàntica objectiu convé conèixer bé el conjunt de dades a processar. Per exemple, és molt rellevant conèixer la proporció de píxels de cada classe (Fig.4). Completa la secció 1.1 del notebook per analitzar aquesta informació. Per facilitar aquest anàlisi el notebook aportar una funció per registrar un mostreig de píxels d'un dataset en un DataFrame de Pandas.

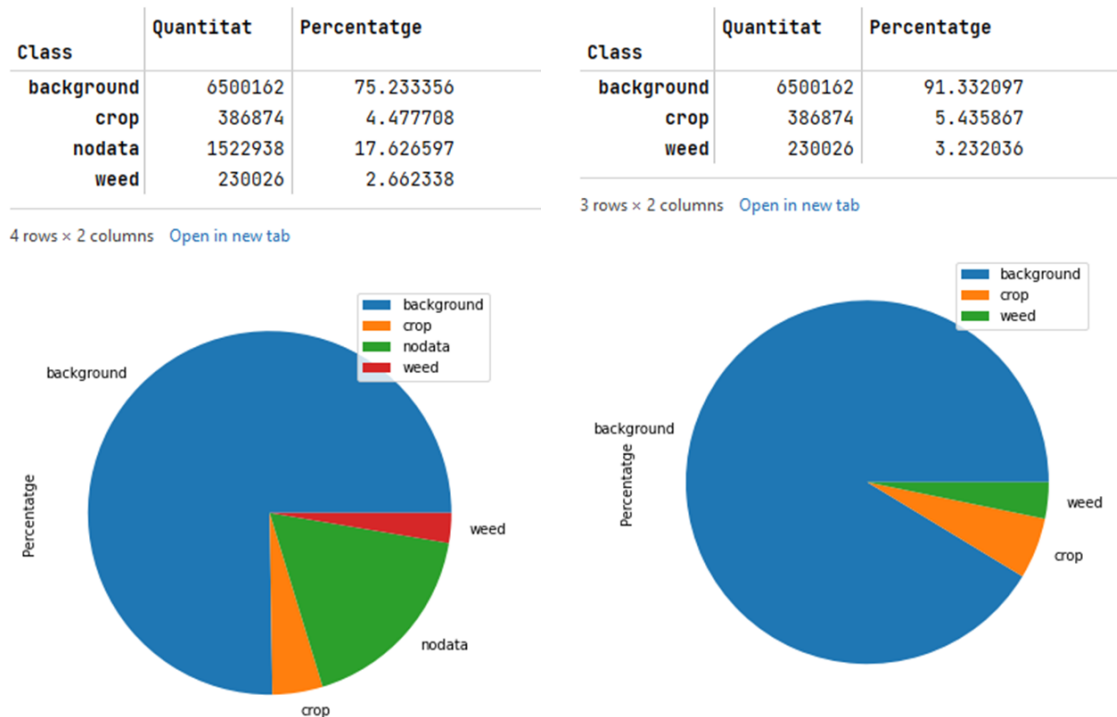


FIGURA 4 PROPORCIÓ DELS PÍXELS DEL DATAFRAME GENERAT A PARTIR DEL CONJUNT D'ENTRENAMENT, AMB I SENSE LA CLASSE 'NODATA'

Seguidament analitza a la secció 1.2 del notebook la informació disponible per cada píxel (5 bandes espectrals). Calcula i visualitza l'histograma a cada banda, pels píxels de les diferents classes (Fig.5). Utilitza el DataFrame sense els píxels de la classe 'nodata' (aquests píxels s'identifiquen de manera trivial, i a més ja se'n disposa de la seva màscara, donada pel software que ha generat les ortoimatges dels cultius).. L'objectiu és tenir una primera visió de la separabilitat de les diferents classes en base únicament a la informació espectral.

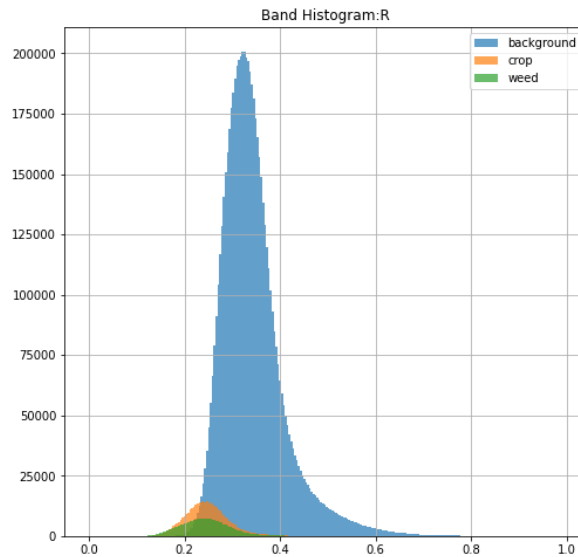


FIGURA 5 HISTOGRAMA DELS VALORS DE LA BANDA 'R' PELS PÍXELS DE LES 3 CLASSES A CONSIDERAR.

2.3 Repte 2: Segmentació semàntica basada en píxels

Per tenir una primera referència de la dificultat de segmentar les males herbes en les imatges del Weedmap dataset, analitza el rendiment de diferents classificadors *clàssics* en aquest problema. Segueix les indicacions de l'apartat 2 del notebook i mesura el rendiment dels classificadors que tens disponibles a scikit-learn. Utilitza també les funcions que hi ha en aquesta llibreria per analitzar el rendiment obtingut (Fig.6).

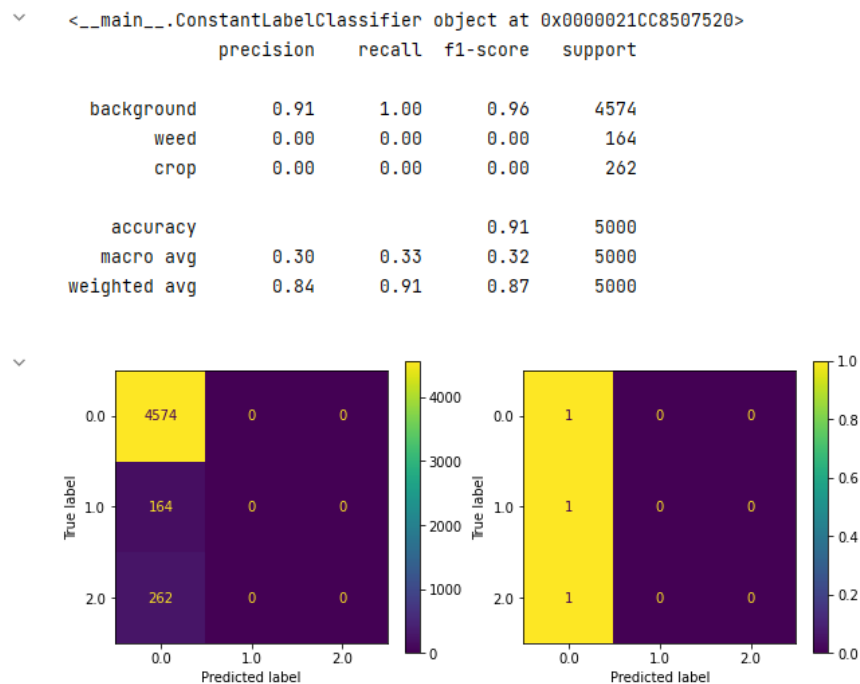


FIGURA 6 INDICADORS DE RENDIMENT I MATRIUS DE CONFUSIÓ CORRESPONENTS A UN CLASSIFICADOR CEC, EL QUAL SEMPRE PREDIU L'ETIQUETA 'BACKGROUND'. ADONA'T QUE TOT I SER UN CLASSIFICADOR INÚTIL, EL SEU VALOR D'ACCURACY ÉS MOLT ALT.

A banda, implementa també les mètriques d'avaluació recollides en aquest article <https://arxiv.org/abs/1704.06857>

Un cop identifiquis quins classificadors van millor, mira si pots millorar-ne el rendiment jugant amb els seus paràmetres.

2.4 Repte 3: Anàlisi del rendiment a nivell d'imatge.

Segueix les indicacions de la secció 3 del notebook i aplica el millor classificador identificat al repte 2 per realitzar la segmentació semàntica de les imatges del dataset. Processa totes les imatges del conjunt de test, calculant el rendiment observat a cada imatge amb les mètriques que consideris més adients. Identifica i visualitza les 5 imatges on el classificador funciona millor i les 5 imatge on funciona pitjor (Fig.7).

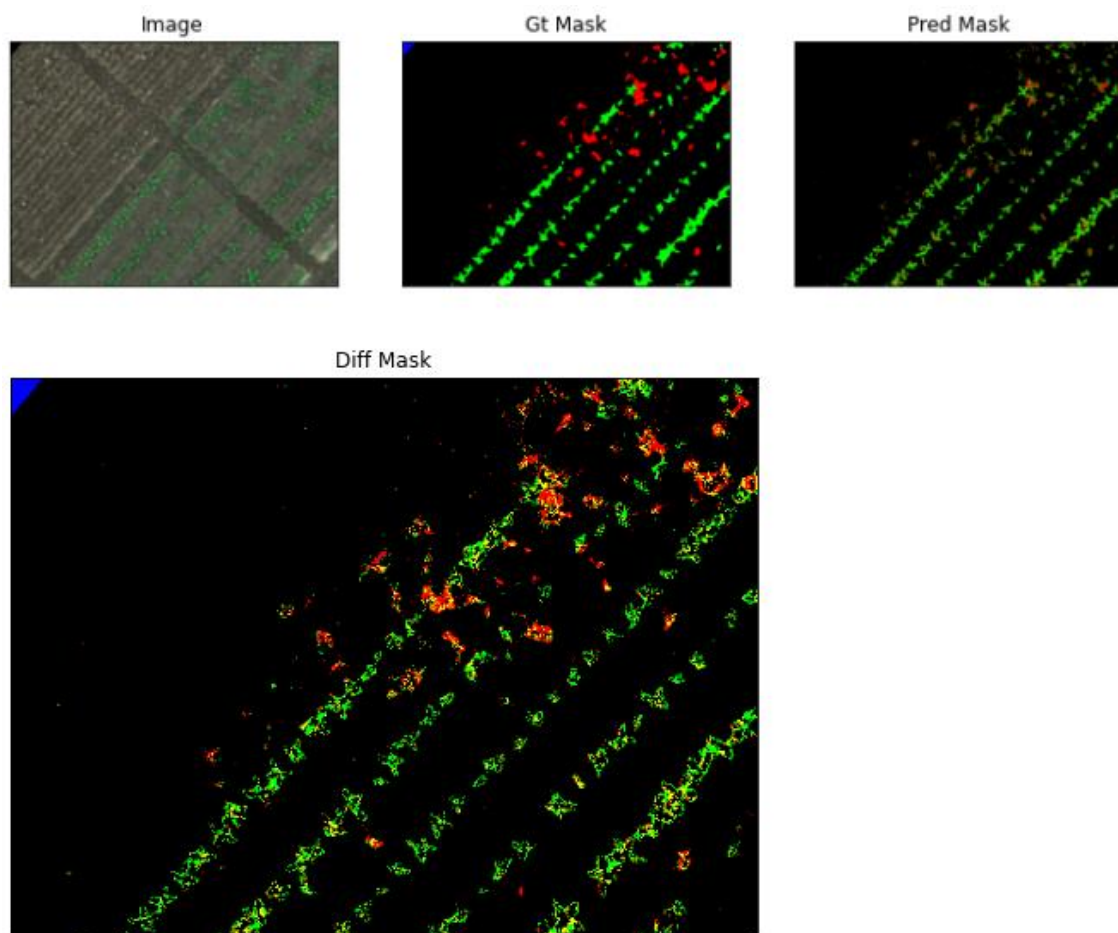


FIGURA 7 POSSIBLE VISUALITZACIÓ DELS RESULTATS OBTINGUTS PER UN CLASSIFICADOR, EN UNA DE LES IMATGES.