

Detecció de plantes no autòctones a les conques dels rius Miño i Sil a través d'IA



CIDAI-POC-2025-02

Drets reservats. Aquest treball està disponible sota la llicència Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International ([CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)).

Segons els termes d'aquesta llicència, podeu copiar, redistribuir i adaptar l'obra amb fins no comercials, sempre que l'obra sigui citada adequadament, tal com s'indica a continuació.

En qualsevol ús d'aquest treball, no s'ha de suggerir que el CIDAI doni suport a cap organització, producte o servei específic. No es permet l'ús del logotip CIDAI.

Si adapteu l'obra, heu de llicenciar-la amb la mateixa llicència Creative Commons o equivalent.

Si creeu una traducció d'aquest treball, heu d'afegir la següent exempció de responsabilitat juntament amb la cita suggerida: "Aquesta traducció no la va crear el Centre of Innovation for Data tech and Artificial Intelligence (CIDAI). CIDAI no es fa responsable del contingut ni de l'exactitud d'aquesta traducció. L'edició original en català serà l'edició autèntica i vinculant".

Qualsevol mediació relacionada amb disputes derivades de la llicència es durà a terme d'acord amb les normes de mediació de la [World Intellectual Property Organization](https://www.wipo.int/mediation/).

Cita suggerida. CIDAI-POC-2025-02 // Detecció de plantes no autòctones a les conques dels rius Miño i Sil a través d'IA per Coen Antens, Joan Lara Formoso, Alberto Rubio Pérez, Edgar Gracia Larrosa i Juan Ramón Jiménez Minguell, Centre de Visió per Computador CVC, CIDAI, 2025. Llicència: [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)



1.	ABSTRACT	4
2.	DESCRIPCIÓ DEL PROBLEMA QUE RESOL	6
3.	IMPLEMENTACIÓ DE LA SOLUCIÓ	15
4.	POTENCIAL IMPACTE DE LA SOLUCIÓ	20



1. Abstract

1.1. Objectius

L'objectiu de l'informe és determinar si, mitjançant les tècniques més recents d'intel·ligència artificial i visió per computador, existeix una tecnologia capaç d'aprofitar la informació proporcionada per les imatges aèries d'alta resolució del Pla Nacional d'Ortofotografia Aèria (PNOA) per a la detecció de plantes invasores a les conques dels rius Miño i Sil.

En el cas de les conques dels rius Miño i Sil les plantes que els interessa controlar són el bambú i la canya.

1.2. Visió

La Confederació Hidrogràfica del Miño-Sil (CHMS) és l'organisme encarregat de gestionar els recursos hídrics de les conques dels rius Miño i Sil, situades al nord-oest d'Espanya. Dintre de les seves principals activitats s'inclou la conservació i protecció del medi ambient, és en aquesta branca, i per tal de controlar espècies invasores, on es va realitzar l'estudi per a la detecció de plantes invasores tals com la canya i bambú.

1.3. Resum del problema

La presència de plantes de bambú i canya a les conques dels rius Miño i Sil genera diverses problemàtiques ambientals i ecològiques. El bambú i la canya han estat introduïts a diferents regions del món per usos ornamentals i funcionals, però la seva proliferació en eco-

sistemes no nadius pot tenir conseqüències negatives, especialment en ecosistemes fràgils com els fluvials d'aquesta zona d'Espanya.

Les principals problemàtiques associades al bambú són:

- **Competència amb espècies autòctones:** El bambú i la canya competeixen amb les plantes natives per recursos com la llum solar, l'aigua i els nutrients del sòl, desplaçant les espècies locals i alterant els equilibris ecològics.
- **Alteracions en l'estructura de l'ecosistema:** El seu creixement dens pot transformar els boscos de ribera i modificar hàbitats aquàtics com els aiguamolls, amb conseqüències negatives per a la biodiversitat i les espècies que depenen d'aquests ecosistemes.
- **Propagació invasiva:** La ràpida expansió del bambú i la canya dificulten la restauració dels ecosistemes degradats, sovint requerint mesures de control costoses i de difícil aplicació.
- **Impactes en infraestructures:** Les arrels invasives poden danyar estructures com ponts i canonades, incrementant els costos de manteniment i reparació.
- **Reptes en la gestió i control:** La seva erradicació és complexa per la capacitat de propagació i resistència. Les mesures habituals, com l'ús d'herbicides, poden generar impactes ambientals no desitjats.



Figura 1. Bambú a la conca del riu

En aquest context, la Confederació Hidrogràfica del Miño-Sil (CHMS) ha col·laborat amb el Centre de Visió per Computador (CVC) amb un estudi per analitzar la viabilitat de detectar la presència de bambú i canya a les conques dels rius Miño i Sil utilitzant imatges aèries d'alta resolució. Aquesta metodologia basada en intel·ligència artificial i visió per computador permetria desenvolupar una eina innovadora per a la gestió i el control d'aquest problema ambiental.

1.4. Solució

S'ha plantejat el desenvolupament d'un sistema capaç d'identificar la presència de les plantes invasores esmentades utilitzant imatges aèries d'alta resolució, xarxes neuronals convolucionals (CNN) i tècniques avançades de *Machine Learning*. Aquest sistema estaria dissenyat per detectar i discriminar amb precisió la presència d'aquestes plantes, i els resultats es geolocalitzarien en un format compatible amb sistemes d'informació geogràfica (SIG o GIS). L'objectiu és generar un arxiu geoespacial que permeti la visualització i anàlisi de les zones afectades, facilitant així la presa de decisions i les actuacions que el regulador consideri oportunes en cada cas.

2. Descripció del problema que resol

2.1. Captura de dades

La Confederació Hidrogràfica del Miño-Sil (CHMinoSil) ha iniciat un projecte basat en l'ús de [Quantum Geographic Information System \(QGIS\)](#), un SIG de codi obert. Aquest sistema permet integrar dades geoespaciales, incloent-hi mapes, localitzacions, capes geoespaciales i zones etiquetades prèviament per CHMinoSil amb presència de bambú.

Les imatges utilitzades en aquest projecte són proporcionades pel [Pla Nacional d'Ortofotografia Aèria \(PNOA\)](#), un programa governamental espanyol que genera ortofotografies georeferenciades amb una resolució de fins a 25 cm per píxel. Aquestes imatges són essencials per a aplicacions com la gestió del territori, l'agricultura, la conservació ambiental i la planificació urbana (vegeu Figura 2).

Figura 2. Imatge del PNOA



La regió d'anàlisi inclou les conques dels rius Miño i Sil, juntament amb els seus afluents (vegeu Figura 3), on inicialment s'han etiquetat 16 zones amb presència de bambú i 32 de canya (vegeu Figura 4).

Figura 3. Àrea a analitzar



Figura 4. Zones etiquetades de bambú.



Per a l'anàlisi, s'utilitzen imatges en format [TIFF georeferenciat](#) amb dimensions de 1000 x 1000 píxels, cadascuna abastant una superfície de 0,01 km². La regió total a inspeccionar cobreix aproximadament 16.76 km², fet que implica la necessitat de gestionar [1.7 milions d'imatges](#) per cobrir l'àrea completa. Amb un pes inicial de 3.82 MB per imatge, l'espai d'emmagatzematge requerit seria de 6.5 TB.

Per fer front a la gran quantitat d'imatges necessàries, s'ha desenvolupat una aplicació que millora dràsticament el procés de descàrrega. Aquest programari incorpora tècniques com l'ús de [threads](#) i la delimitació de zones d'interès basades en formes geogràfiques (*shapefiles*). A més, s'ha implementat una modificació per descarregar les imatges [en format JPEG](#) sense perdre la informació de georeferenciació, reduint significativament la mida de les imatges.

2.2. Detecció de les plantes invasores

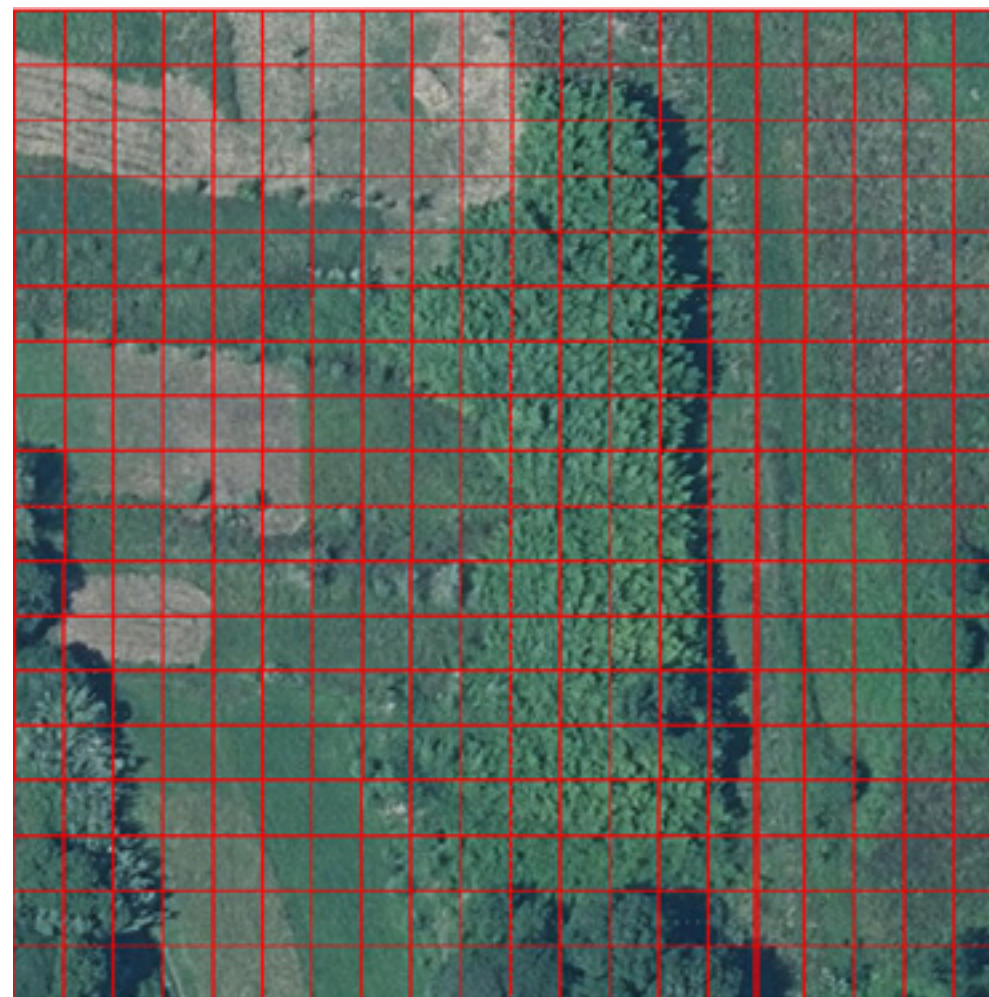
Degut a que el nombre de plantes invasores de bambú i canya és molt diferent (la presència actual de canya és molt més gran que la de bambú) els mètodes de detecció que s'han hagut d'implementar han estat diferents. A continuació s'analitzen els diferents mètodes que s'han treballat durant aquest estudi.

2.2.1 Detecció del Bambú

Preparació del Dataset i Data Augmentation

Un cop descarregades les imatges corresponents a les zones amb presència etiquetada de bambú, és necessari dividir-les en petites quadrícules o patches (vegeu Figura 5). Aquesta divisió es realitza perquè els sistemes d'aprenentatge automàtic, com les xarxes neuronals, no poden gestionar imatges de gran mida de manera eficient.

Figura 5. Patches de la zona.



Després de la divisió, es classifica cada *patch* en dues categories principals:

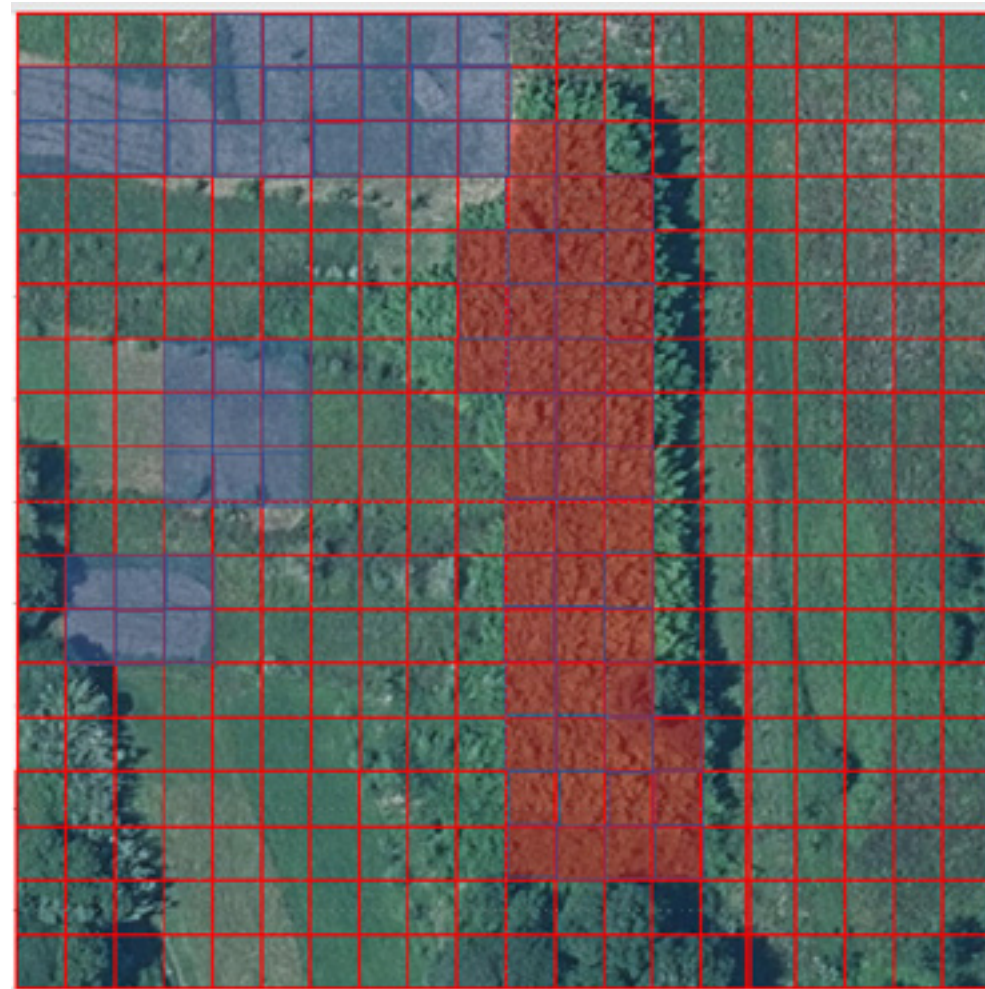
- **Background (fons):** Zones sense presència de bambú.
- **Bambú:** Zones on es detecta la presència de la planta invasora.

Per determinar la classificació, s'utilitza una aplicació que combina la imatge georeferenciada descarregada i una màscara específica de bambú. Aquesta màscara identifica àrees amb presència de bambú, de manera que:

- Els *patches* que col·lideixen amb la màscara es classifiquen com a "bambú".
- Els *patches* que no hi col·lideixen es classifiquen com a "background".

Per optimitzar el procés i reduir el volum de dades irrellevants, s'aplica un filtre de color previ que elimina àrees que no són verdes. Aquesta tècnica permet descartar patches que clarament no contenen vegetació, reduint el nombre de quadrícules a analitzar (vegeu Figura 6).

Figura 6. Patches de la zona.



En el procés de classificació, s'empren els següents codis de color:

- **Vermell:** Patches classificats com a exemples de bambú.
- **Blau:** Patches descartats per no ser verds.
- **Altres colors:** Patches classificats com a "background".

Aquest enfocament no només millora l'eficiència del sistema, sinó que també proporciona una entrada de dades precisa per a l'entrenament dels sistemes de xarxes neuronals en la detecció de plantes invasores.

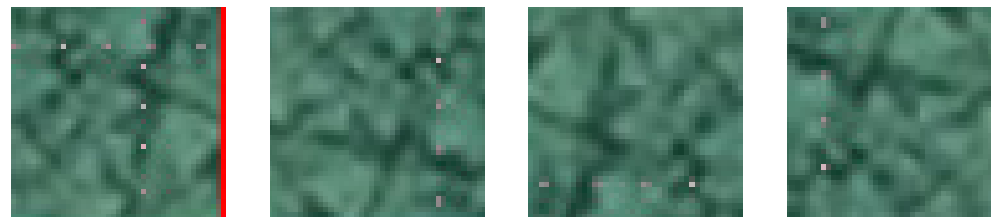
Degut a la limitada disponibilitat d'imatges de bambú, és necessari realitzar un procés de *Data Augmentation*. Aquesta tècnica, també coneguda com a augment de dades, és fonamental en el context de visió per computador i entrenament de models d'aprenentatge automàtic, especialment en tasques de reconeixement d'imatges.

El *Data Augmentation* consisteix a generar noves mostres per al conjunt d'entrenament mitjançant l'aplicació de diverses transformacions a les imatges originals. Aquestes transformacions poden incloure:

- Rotacions
- Translacions
- Miralls horitzontals i verticals
- Canvis en el brillantor i el contrast, entre altres.

Aquestes operacions no alteren la classe o etiqueta de la imatge original, però introdueixen variacions que enriqueixen el conjunt de dades (vegeu Figura 7).

Figura 7. Data Augmentation aplicant rotacions.



En el nostre cas, s'ha implementat una tècnica de *Data Augmentation* específica per a les imatges relacionades amb el bambú. Això es deu al desequilibri en el conjunt de dades: mentre que es disposa d'una gran quantitat d'informació sobre el background (fons), les dades relatives al bambú són molt limitades.

Models de detecció

El Random Forest és un algorisme d'aprenentatge automàtic basat en la construcció de múltiples arbres de decisió durant el procés d'entrenament. Destaca per la seva capacitat per avaluar la importància de les característiques utilitzades per al model.

Es van considerar les següents característiques:

- Mitjana dels nivells RGB (vermell, verd, blau) i la seva desviació estàndard.
- Informació de textura obtinguda a partir de les *Gray-Level Co-occurrence Matrices (GLCM)*.

També es va realitzar una comparativa per determinar si era més beneficiós treballar amb aquestes característiques a nivell de píxel o dividir les imatges en petites zones (patches). Els resultats es presenten a les figures Figura 8 i Figura 9.

Figura 8. Resultats a nivell de píxel.

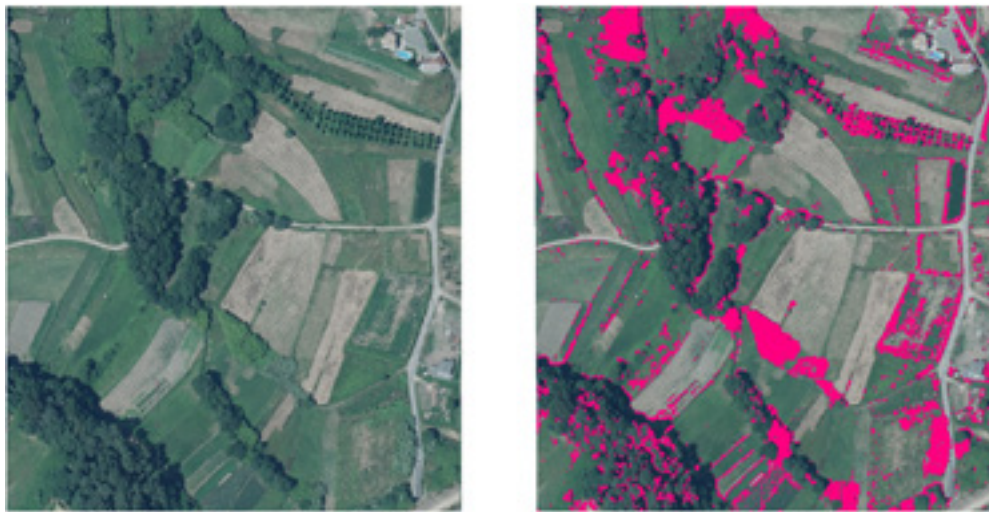
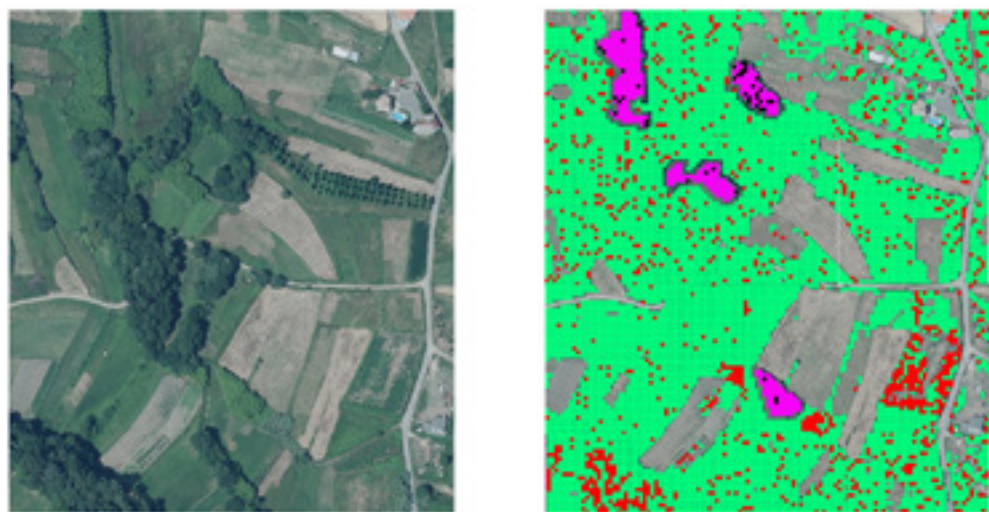


Figura 9. Resultats a nivell de patch.



Tot i que els resultats a nivell de píxel eren més precisos, els resultats a nivell de patch van demostrar:

- Una reducció dels falsos positius.
- Una velocitat d'inferència superior.

Posteriorment, es va dur a terme una anàlisi per identificar quines característiques tenien un major pes en la predicció de la presència de bambú i l'absència del mateix. Els resultats van indicar que la informació de color tenia una influència més significativa que la textura.

2.2.2 Detecció de Canya

Models de detecció

El procediment amb el qual treballarà la nostra Intel·ligència Artificial és el d'Oriented Bounding Box (OBB). Funciona de manera que el model, al passar-li una imatge, ens indicarà en un requadre on es troba la canya.

El model a utilitzar com a base serà YOLO (You Only Look Once) versió 11. Aquest model base no coneix com detectar la canya des de les ortoimatges, per tant l'haurem de entrenar per a tal. El procés amb el qual l'entrenarem serà la anotació, que consisteix en marcar amb un rectangle el àrea on es troba la canya. Entrenant el model amb aquestes anotacions aprendrà els patrons amb els quals la canya es presenta a les imatges, de forma que el model podrà detectar-les (vegeu Figura 10).

Figura 10. Anotacions de canya.



A l'anotar un nombre considerable d'aquestes imatges amb canya, localitzacions de les quals havien sigut prèviament confirmades pel client, llancem la IA a un àrea aleatòria on sabem que trobarem canya. Així obtindrem les primeres deteccions.

Obtenció de Resultats

Les deteccions obtingudes se'ns retornaran dins d'un arxiu .shp (capa), que ens indicarà a QGIS les Bounding Boxes on el model ha trobat canya. Dins de QGIS convertirem aquestes boxes als seus centroides, que ens indicaran el centre de cada detecció. Amb un arxiu .csv amb el valors de tots aquests centroides podem fàcilment descarregar totes les imatges on l'algoritme pensa que ha trobat canya.

Figura 11. Deteccions inicials



Figura 12. Falsos positius vs positius reals



Dins d'aquestes imatges, especialment a les primeres iteracions del model, ens trobarem un gran nombre de falsos positius. Per tant la nostra feina serà filtrar les deteccions correctes, anotar-les i tornar-les a passar al model per entrenar-lo. D'aquesta forma, anirem fent prediccions a àrees cada vegada més grans i allunyades, filtrant les deteccions i passar-les com a Ground Truth (imatges d'entrenament) al model.

A cada àrea que el model prediu, hem d'esperar moltes equivocacions degut a que a cada zona la il·luminació, orientació, època de l'any i fins i tot resolució pot variar. Les imatges de PNOA es recopilen d'entre multitud de vols fets a moments diferents. Aleshores, al tenir la canya noves característiques a cada zona, el model necessita més reentrenament amb els nous patrons.

Per aquest motiu hem dividit el àrea a analitzar en 5 subzones, cada una amb unes característiques a les imatges lleugerament diferenciades (Figura 13).

Figura 13. Divisió per zones.

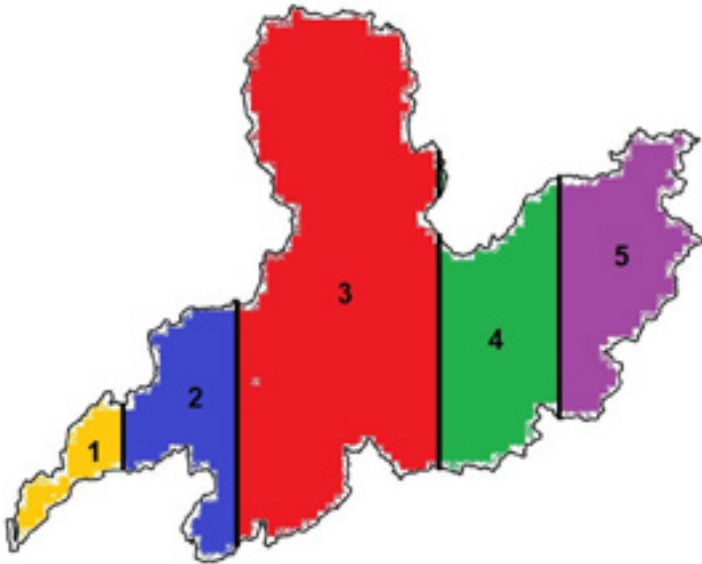


Figura 14. Canya zona 1 vs Canya zona 3



El model acabarà sent cada vegada més precís segons li afegim imatges d'entrenament, i aprendrà els patrons amb els que la canya apareix a les diferents àrees. Repetint aquest model iterativament a cada zona i filtrant els resultats, trobarem de forma definitiva la localització de la canya.

3. Implementació de la solució

3.1. Arquitectura, tecnologia i dades utilitzades

Després d'haver implementat cadascun d'aquests algorismes, s'ha desenvolupat el pipeline que s'aprecia a la Figura 15 i la Figura 16.

Figura 15. Pipeline de detecció de bambú.

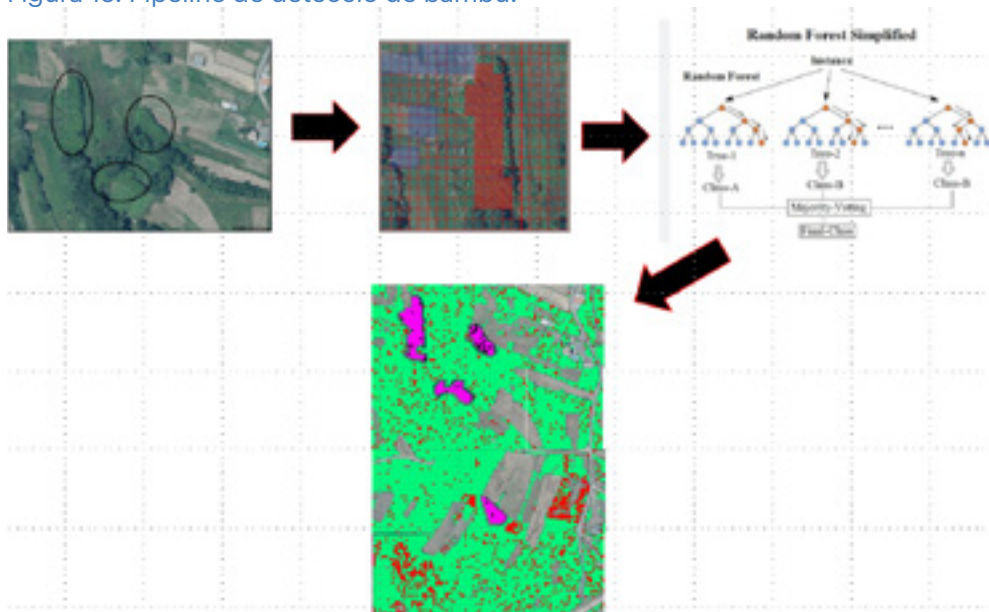
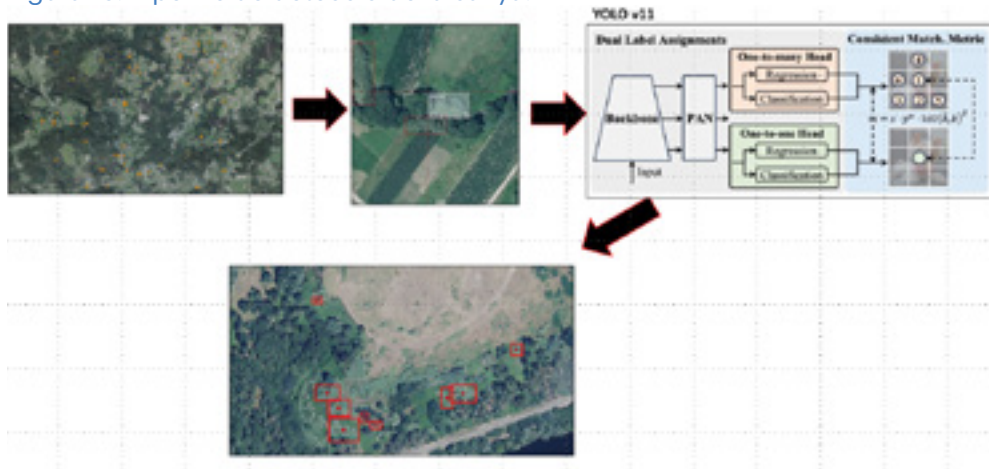


Figura 16. Pipeline de detecció de la canya.



3.2. Reptes resoltos i resultats obtinguts

La complexitat d'aquest projecte és la dificultat per a diferenciar tant la canya com el bambú d'altres tipus de plantes o arbres a les imatges. L'ull humà pot tenir dubtes en alguns casos, ja que hi ha plantes molt similars a la canya o el bambú des d'una vista aèria. Per tant els patrons amb què tant la canya i el bambú es presentarà són de difícil distinció per a l'algorisme, causant un nombre massa gran de falsos positius.

Figura 17. Fals positiu.



Un problema afegit és la manca d'exemples en el cas del bambú, que es soluciona mitjançant el mètode de patches i el data augmentation. Un algorisme que abans d'entrenar el model, augmenta la quantitat d'imatges d'entrenament agafant les actuals i distorsionant-les lleugerament, rotant-les, canviant el color, etc. La idea d'això era que el model aprengués els patrons diferenciadors a base d'aprendre de una quantitat molt gran de imatges.

Els resultats no van ser favorables, portant a 8.000 deteccions només a la zona 1, un nombre que no és acceptable per a fer una revisió manual.

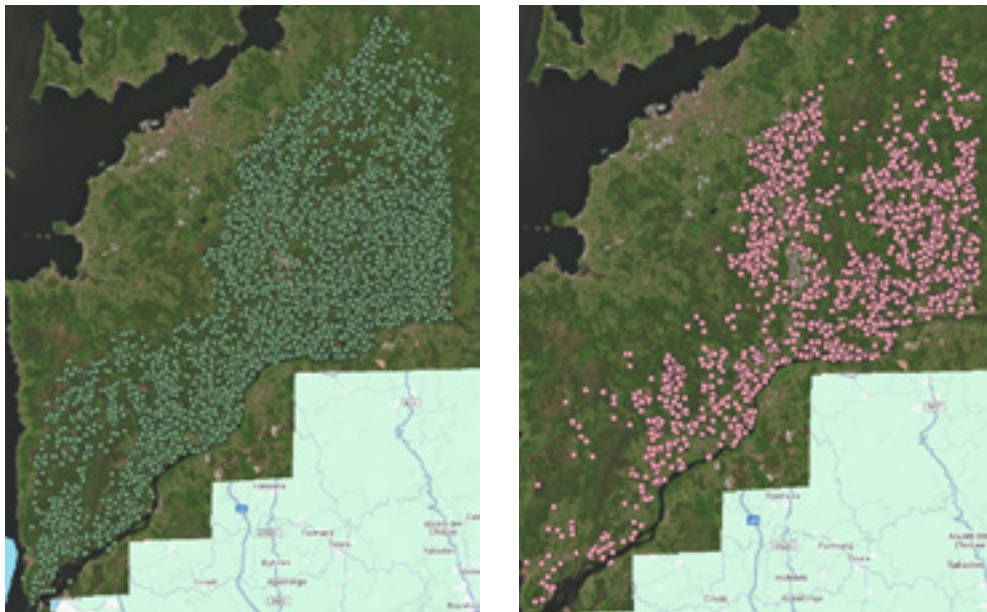
El mètode que sí que ha donat resultats molt favorables ha sigut el de la creació de noves classes. És a dir, el model no només predirà canya, si no els tipus de falsos positius que més es detecten erròniament com a canya o com bambú.

Així en el cas de la canya analitzant prediccions anteriors, ens vam adonar que de forma molt reiterada es detectaven palmeres en comptes de canya, així que també les vam afegir com a entrenament i el model va començar a detectar-les de forma separada a la canya.

Amb aquesta metodologia es van incorporar no només les palmeres, sinó els camps de cultiu, un tipus d'arbust similar a la canya, boscs de la muntanya i els arbres morts.

L'objectiu amb aquestes noves classes no és la detecció precisa, simplement alleugerar la càrrega de falsos positius del que realment és canya. De manera similar es va haver de fer amb el cas de bambú, però en comptes de a nivell de bounding box a nivell de patch.

Figura 18. Resultats de prediccions amb data augmentation vs resultats amb creació de noves classes



En el cas del bambú la problemàtica associada a l'aspecte de les imatges s'ha abordat mitjançant una tècnica anomenada CycleGAN. La tècnica CycleGAN és un tipus d'algorisme d'aprenentatge automàtic utilitzat en visió per computador per al processament i la generació d'imatges. Es basa en xarxes neuronals generatives adversàries (*Generative Adversarial Networks*, GAN), que consten de dues xarxes principals:

- **Generador:** Crea noves imatges a partir d'un conjunt de dades. En el context de CycleGAN, el generador pren imatges d'un domini A i intenta transformar-les perquè semblin imatges del domini B.
- **Discriminador:** Avalua les imatges generades pel generador i determina si són autèntiques o falses en comparació amb un conjunt de dades real del domini B. L'objectiu és millorar la capacitat del generador per produir imatges creïbles.

La particularitat de CycleGAN és la seva capacitat d'aprendre la correspondència entre dos dominis d'imatges sense necessitat de parelles de dades etiquetades. En lloc de requerir imatges corresponents una a una entre els dos dominis, CycleGAN pot aprendre les relacions entre els dominis A i B de manera no supervisada.

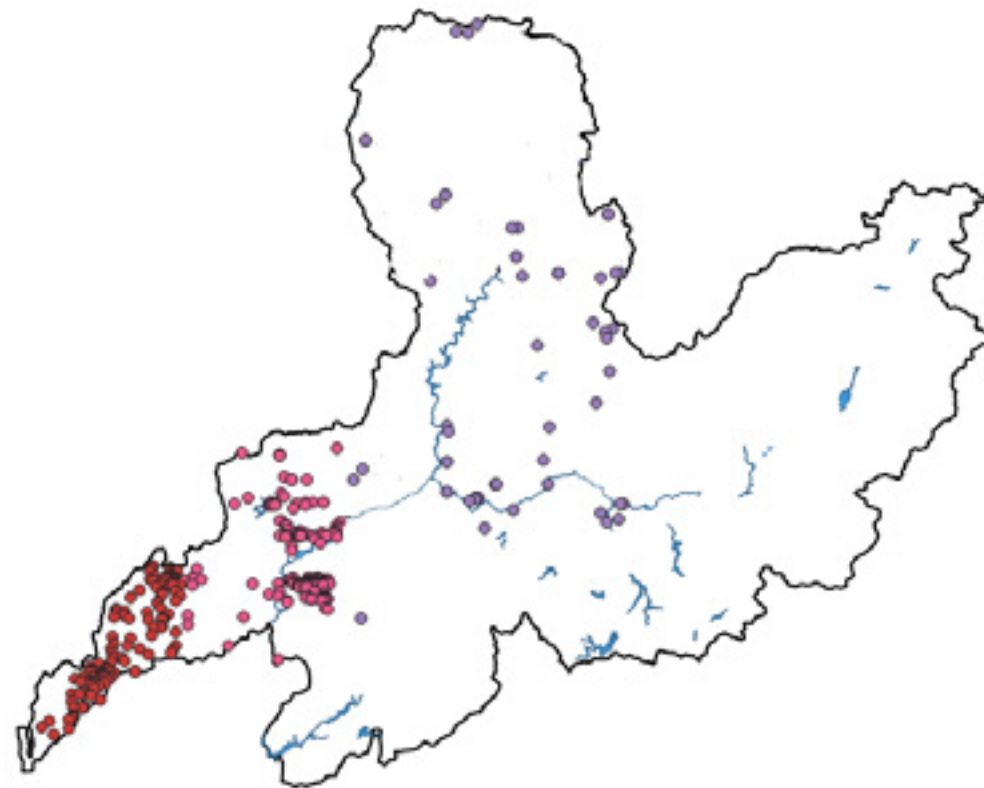
En el nostre context, CycleGAN s'ha utilitzat per convertir l'aspecte de les imatges de la zona 1 a les zones 2 i 3 (vegeu Figura 19).

Figura 19. A la esquerra tenim les imatges reals de les zones 2 i 3, i a la dreta tenim les imatges transformades per tindre un aspecte de la zona 1.



Aplicant aquestes tècniques en el cas del bambú obtenim els següents resultats (vegeu Figura 20):

Figura 20. Xinxetes de predicció presència de bambú.



Es pot observar que s'ha reduït ostensiblement el nombre de falsos positius. Tot i així, es continua observant que el bambú es detecta principalment a les zones 1 i 2, mentre que a la zona 3 gairebé no es detecta. No sabem si això és així realment o si patim d'un baix índex de detecció.

En el cas de la canya tenim els següents resultats:

Figura 21. Localització de les canyes detectades



Els resultats finals han sigut 2806 canyes totals detectades. Predominen als afluent occidentals del riu Miño i a la seva vora, i s'està expandint ràpidament cap al riu Sil i Ourense. La canya es reforça a les zones més humides, i ja ha envaït pràcticament el riu Miño, però si s'actua ràpidament encara es pot frenar als voltants del riu Sil. Aquest projecte pot servir com a aprenentatge per un millor control d'espècies a qualsevol indret d'Espanya.

3.3. Limitacions actuals de la tecnologia

La principal problemàtica trobada ha estat causa de les imatges descarregades del PNOA. Aquestes imatges s'actualitzen cada 3 anys aproximadament, i la versió descarregada a les zones més orientals és anterior a la actual. Això comporta una desactualització de la evolució de la invasió de canya o de bambú.

Les imatges al PNOA visualitzades a QGIS són actuals, llavors tenim una disparitat respecte a què visualitzem i què visualitza el model. La canya s'ha estès significativament des de fa 3 anys, i fa que al PNOA visualitzem canya que no existeix a les imatges del model. Això vol dir que el que es pensava que eren falsos negatius, realment no ho eren, ja que a aquesta àrea el model no tenia res a predir.

Figura 22. Prediccions visualitzades a PNOA actual.



Figura 23. Imatges desactualitzades utilitzades pel model.



Aquestes imatges demostren clarament l'evolució de la espècie invasora i la seva agressió a espècies autòctones.

El objectiu d'aquest tipus de projecte és la prevenció de invasió de espècies no autòctones. Veiem com esperar 3 anys ja ha dificultat la detecció i la lluita contra aquesta. La ràpida actuació i detecció de perills al ecosistema autòcton és essencial per preservar-lo.

4. Potencial impacte de la solució

El projecte va demostrar la viabilitat d'utilitzar ortofotos del Pla Nacional d'Ortofotografia Aèria (PNOA) en combinació amb tècniques de visió per computador, específicament Deep Learning, per a la detecció de la planta invasora bambú i canya a les conques dels rius Miño-Sil.

La implementació de xarxes neuronals convolucionals (CNN) i Machine Learning va resultar exitosa en la detecció precisa de zones amb presència de bambú i canya a les imatges satel·litals, fet que suggereix un potencial significatiu per a futures aplicacions en la monitorització i detecció d'altres espècies invasores.

Es va observar una distribució espacial heterogènia tant del bambú com la canya, amb una concentració significativa a les zones occidentals, especialment a les desembocadures dels rius, mentre que la presència era menor a la part central de la regió estudiada.

Aquesta variabilitat subratlla la importància de considerar factors geogràfics i ambientals en la gestió d'espècies invasores.

Malgrat l'èxit en la detecció del bambú, es van identificar limitacions en la qualitat de les imatges satel·litals, especialment a la part oriental de la regió d'estudi. La baixa resolució d'aquestes imatges va afectar la precisió de l'anàlisi en aquestes zones, fet que posa de manifest la necessitat d'utilitzar fonts de dades més actualitzades per a futures investigacions, així com fer una monitorització més continua, a través d'altres fonts d'imatges.

Els resultats obtinguts emfatitzen la importància de la monitorització contínua de la distribució d'espècies invasores mitjançant tecnologies de teledetecció i Deep Learning. Aquesta informació és essencial per implementar mesures de control i gestió efectives dels ecosistemes afectats, així com per a la conservació de la biodiversitat.



Socis



Membres

