

Fully Convolutional Neural Networks (FCNN)

Caso de uso: Clasificación Automática de Ortofotografías Aéreas del PNOA

ML Sharing Group
28/09/2017

Clasificación Automática de Ortofotografías Aéreas del PNOA

Machine Learning en Computer Vision



CAT

Clasificación



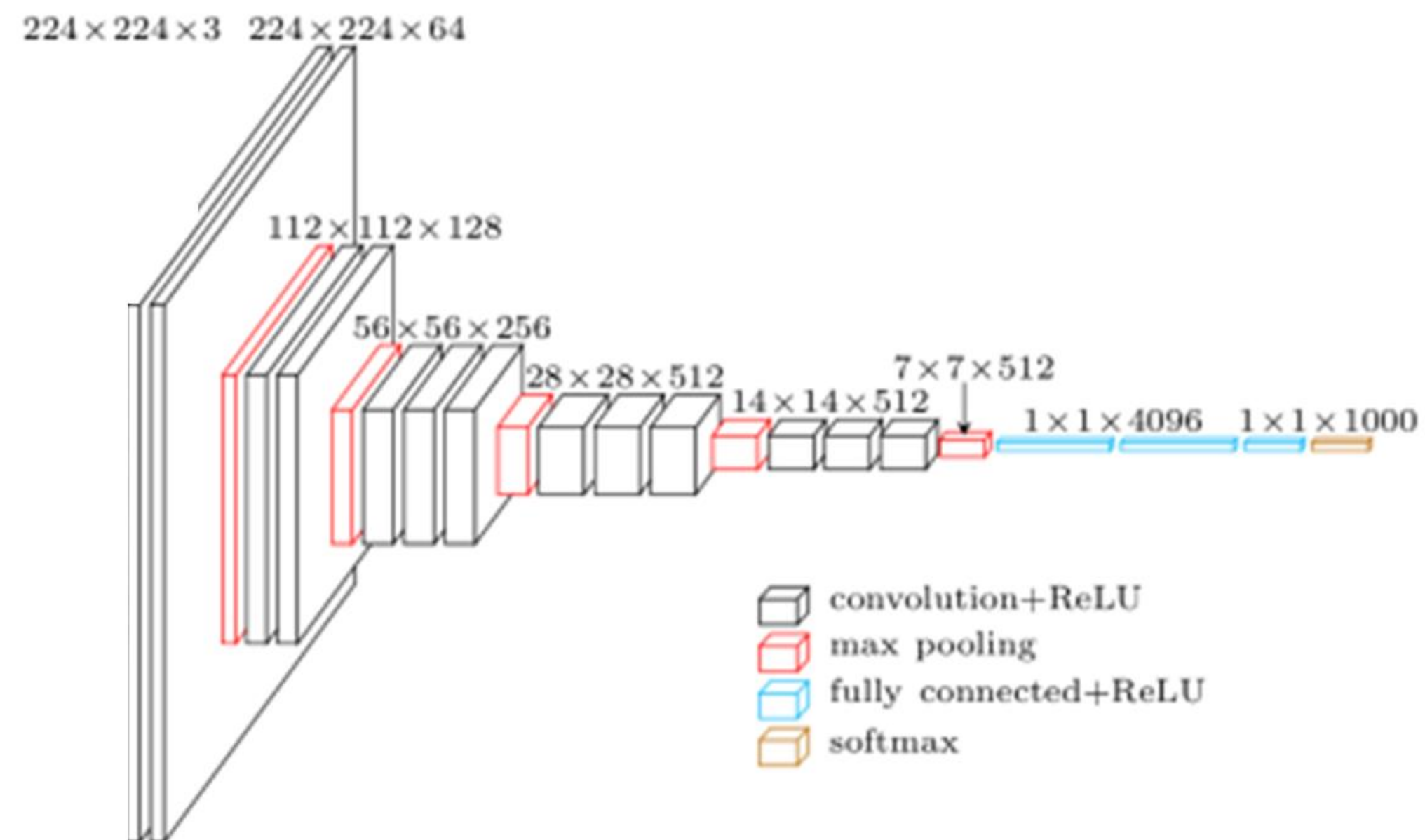
Clasificación Automática de Ortofotografías Aéreas del PNOA

Machine Learning en Computer Vision



CAT

Clasificación



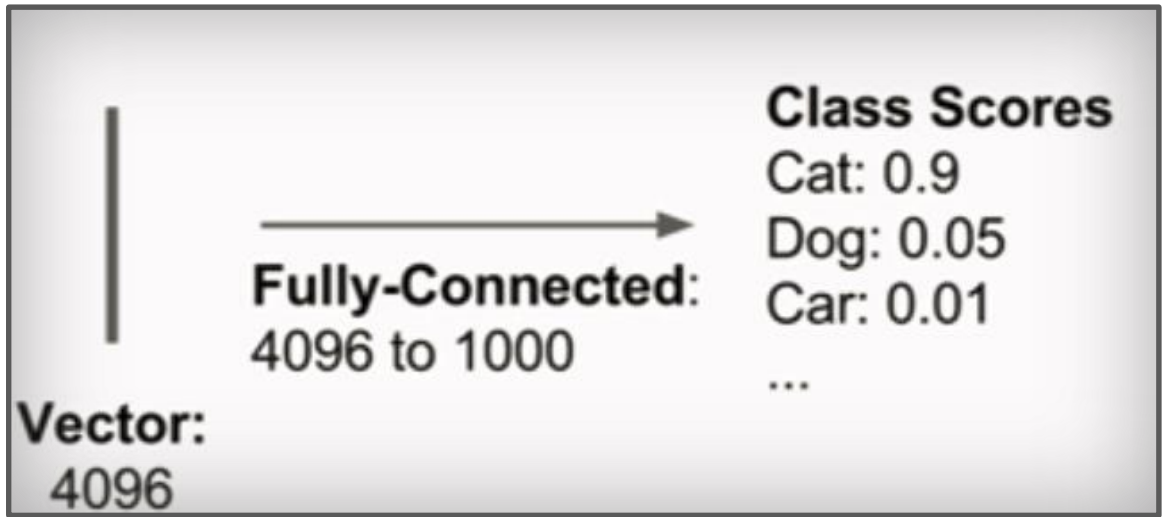
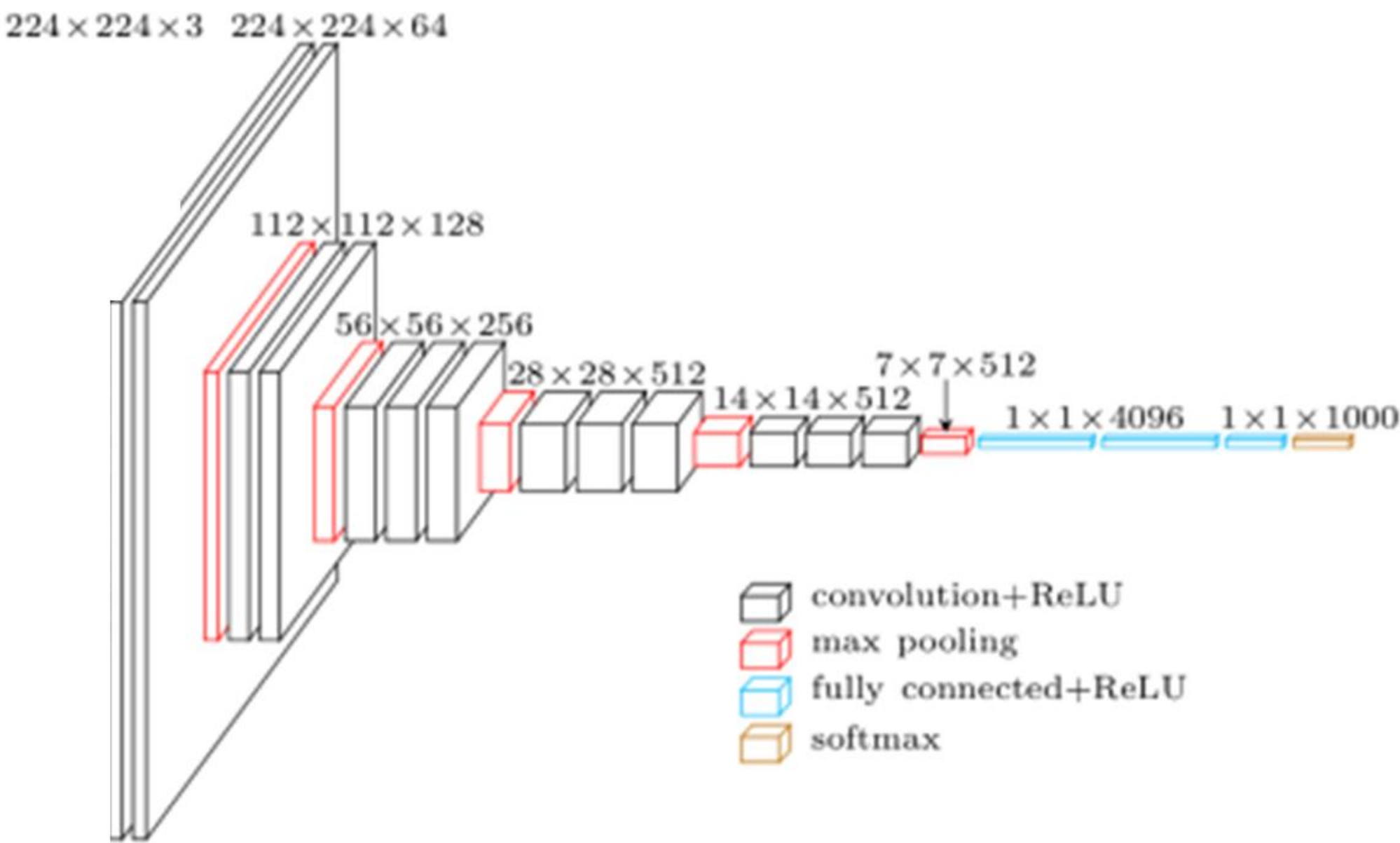
Clasificación Automática de Ortofotografías Aéreas del PNOA

Machine Learning en Computer Vision



CAT

Clasificación



Clasificación Automática de Ortofotografías Aéreas del PNOA

Machine Learning en Computer Vision



CAT

Clasificación



CAT

Detección

Clasificación Automática de Ortofotografías Aéreas del PNOA

Machine Learning en Computer Vision



CAT

Clasificación



CAT

Detección



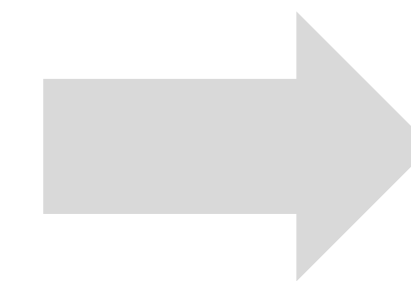
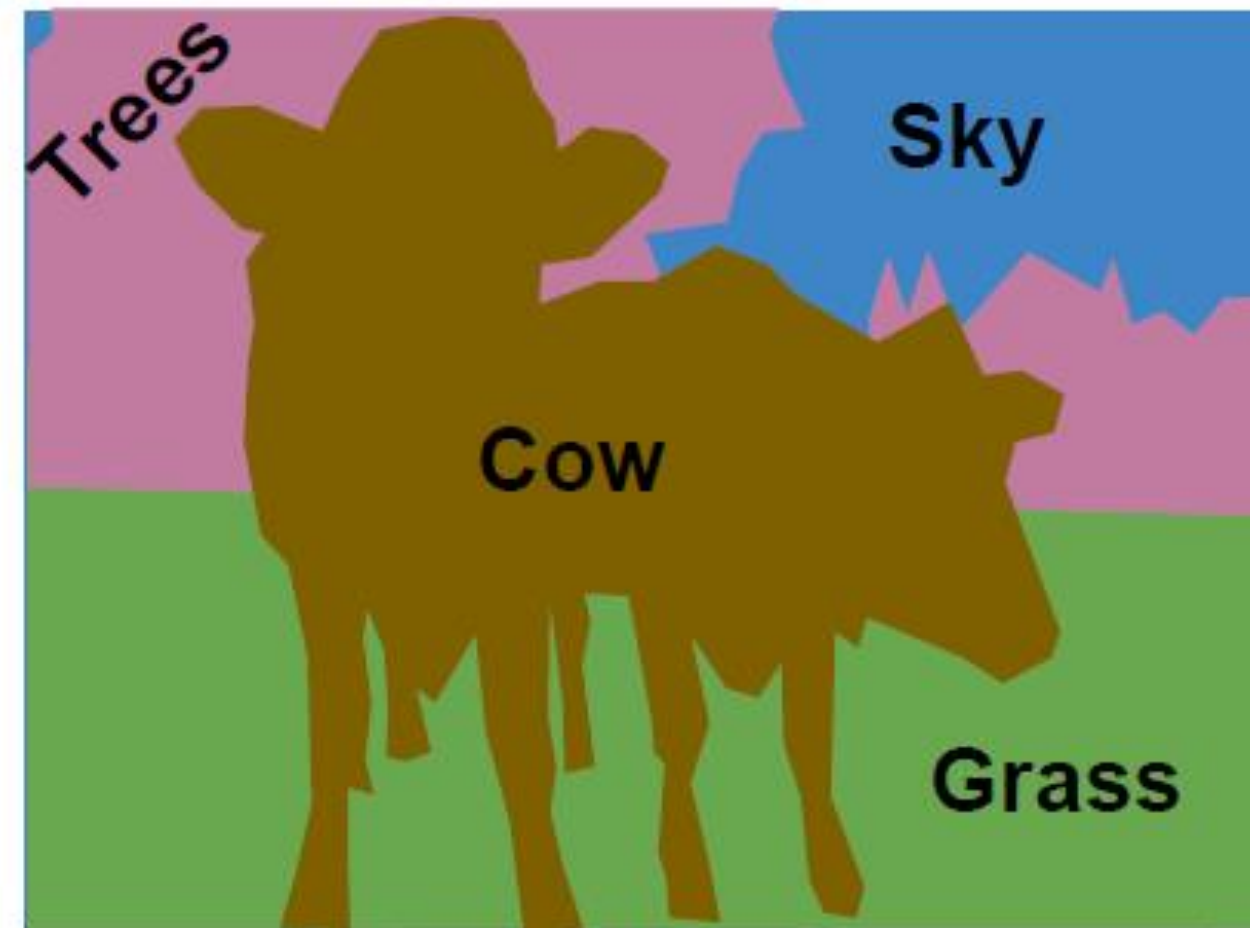
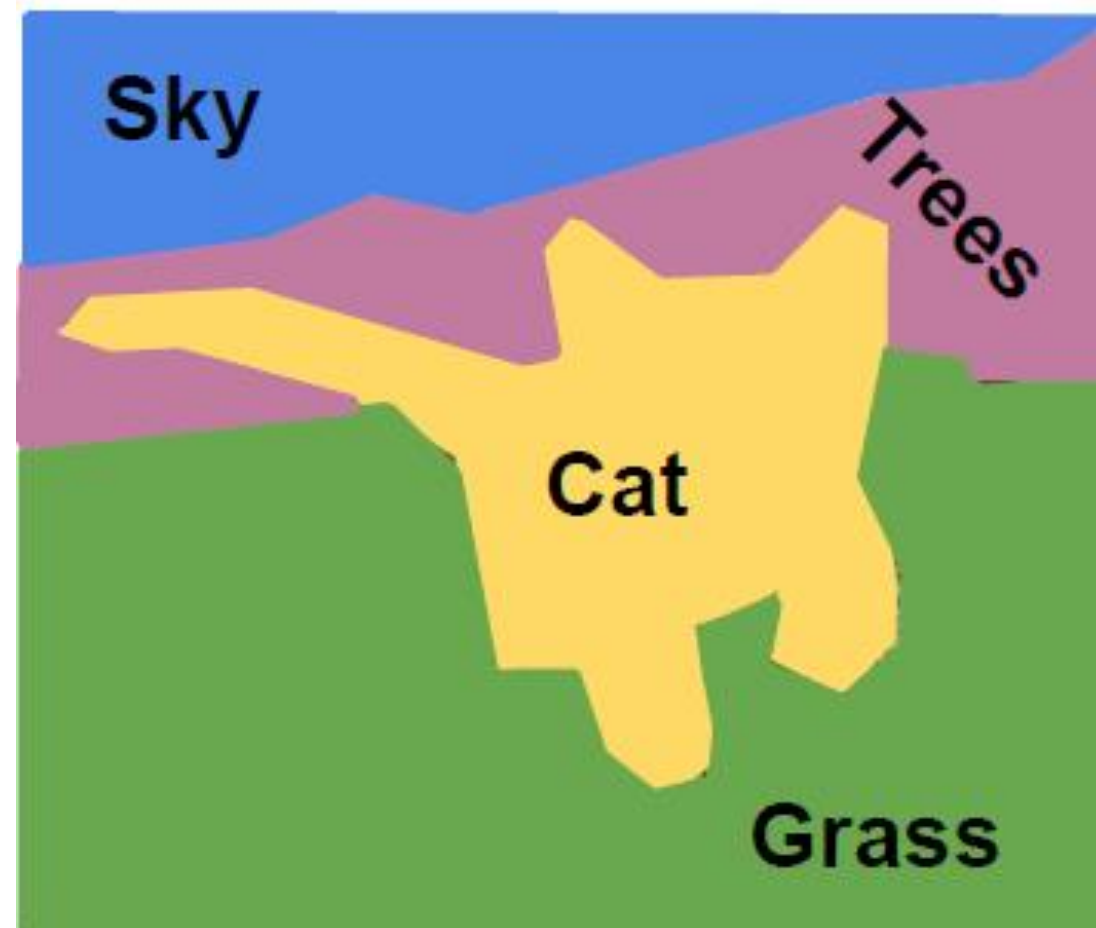
CAT, DOG, DUCK

Segmentación



Clasificación Automática de Ortofotografías Aéreas del PNOA

Machine Learning en Computer Vision



**Each pixel
independiently**



Clasificación Automática de Ortofotografías Aéreas del PNOA

Machine Learning en Computer Vision



CAT

Clasificación



CAT

Detección



CAT, DOG, DUCK

Segmentación



DOG, DOG, CAT

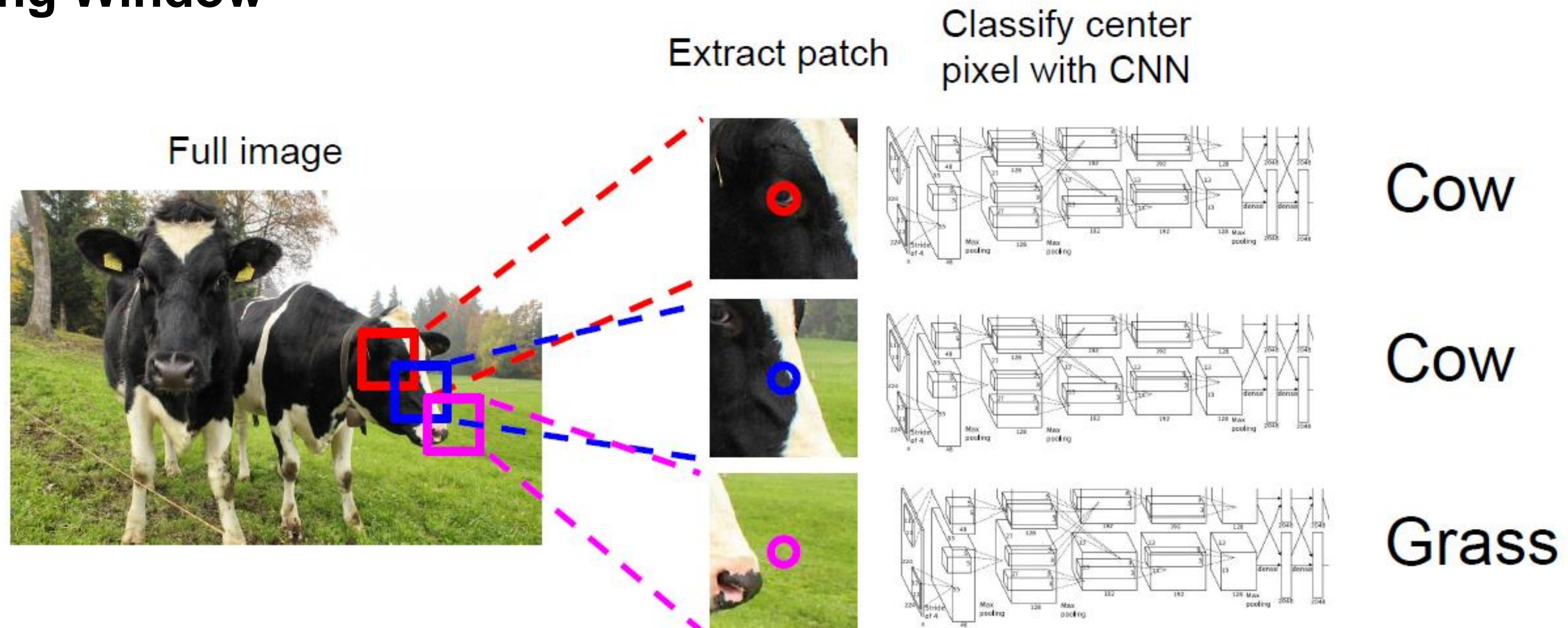
**Instance
Segmentation**



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- Sliding Window



Farabet et al, "Learning Hierarchical Features for Scene Labeling," TPAMI 2013

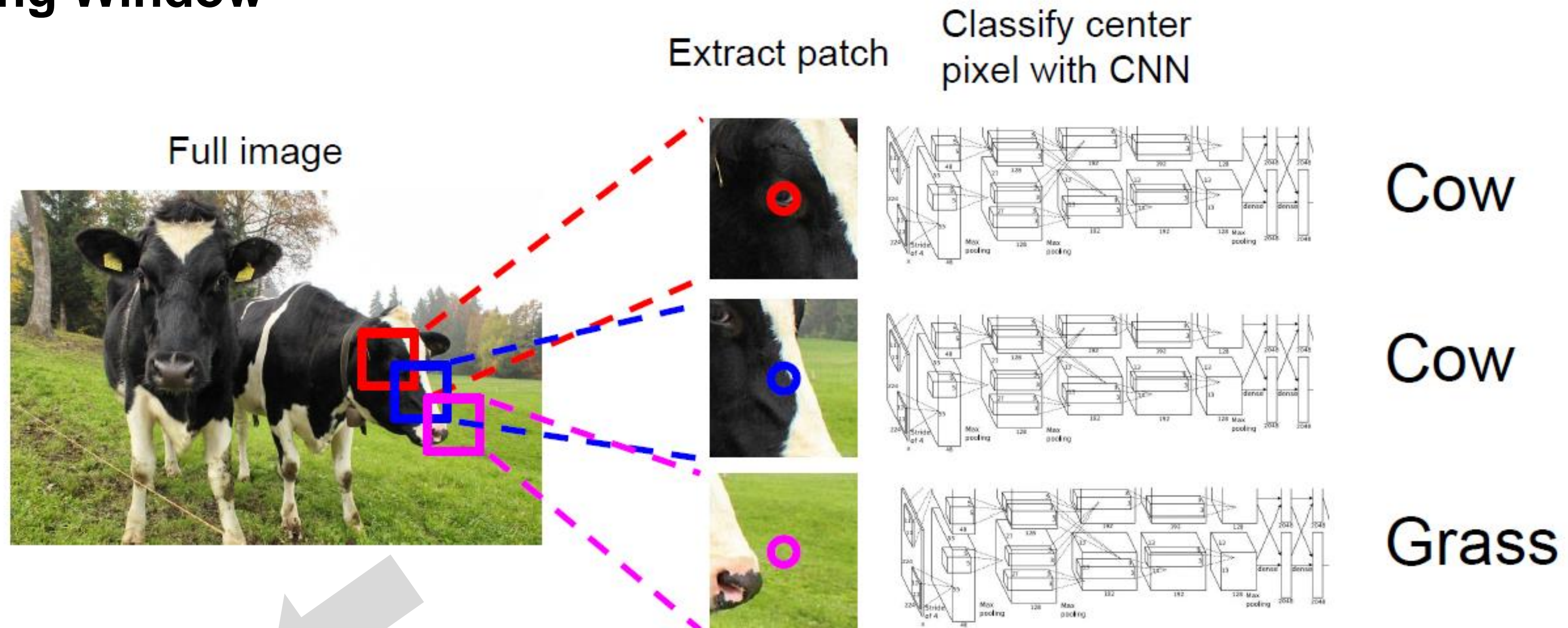
Pinheiro and Collobert, "Recurrent Convolutional Neural Networks for Scene Labeling", ICML 2014



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- Sliding Window



Altamente ineficiente!
Coste computacional
muy alto

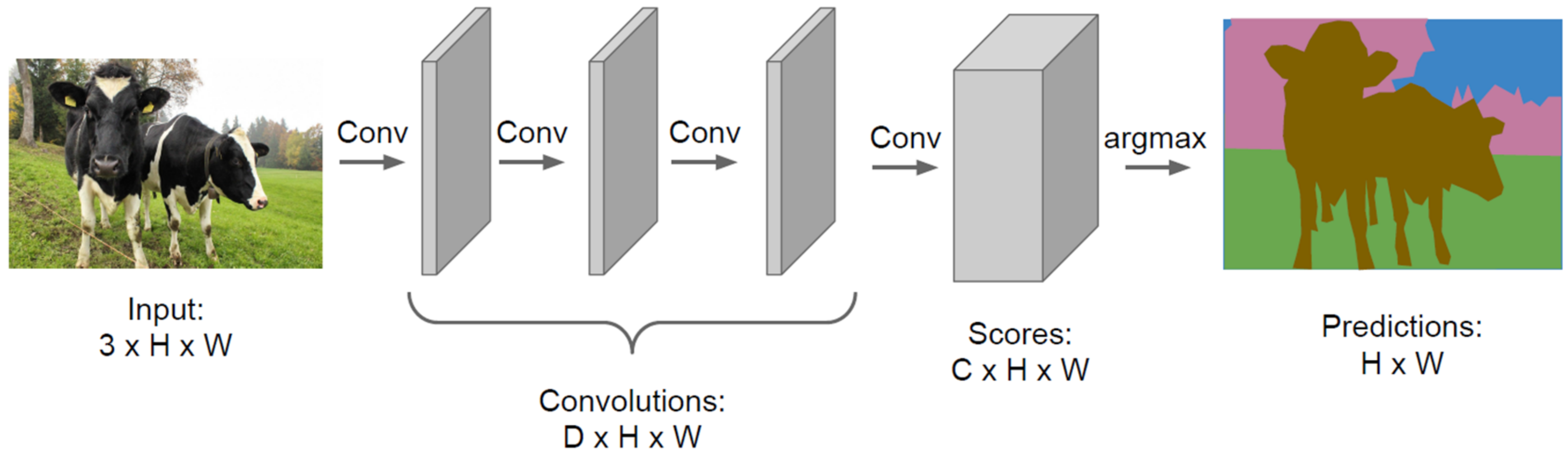
Farabet et al, "Learning Hierarchical Features for Scene Labeling," TPAMI 2013
Pinheiro and Collobert, "Recurrent Convolutional Neural Networks for Scene Labeling", ICML 2014



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

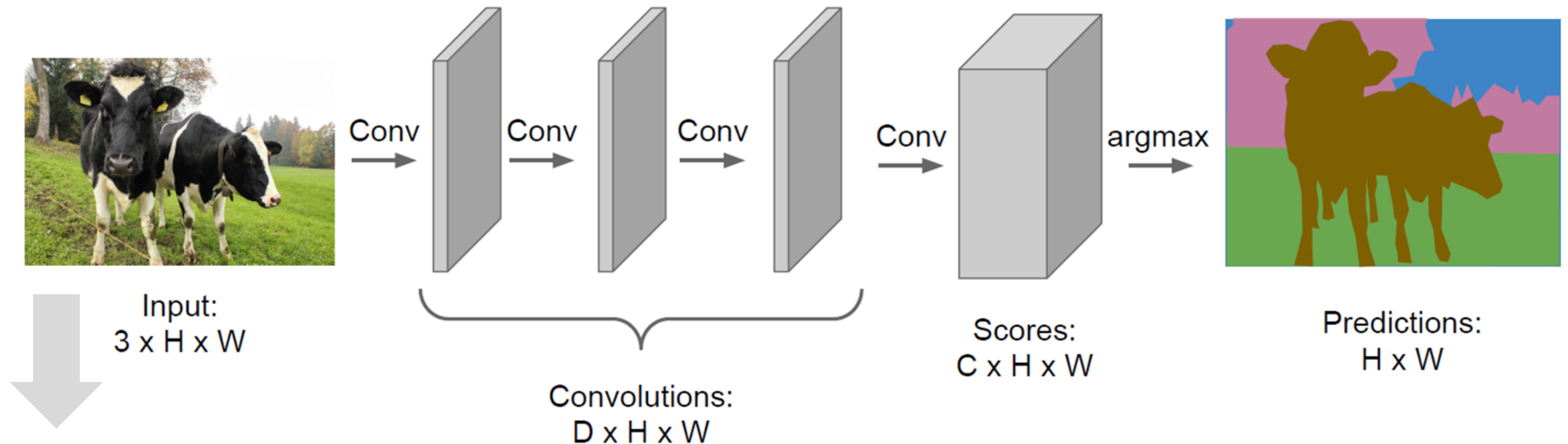
- Fully Convolutional



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- Fully Convolutional

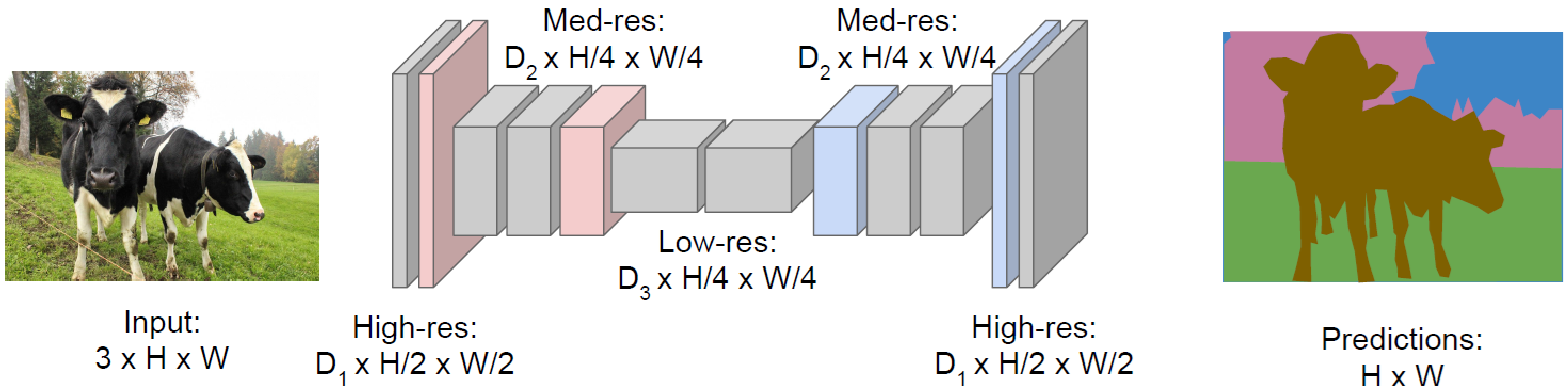


Altamente ineficiente!
Coste computacional
muy alto

Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- Fully Convolutional



Design network as a bunch of convolutional layers, with **downsampling** and **upsampling** inside the network!



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- How upsampling?
 - Pooling Layer → Unpooling Layer

Nearest Neighbor

1	2
3	4

Input: 2 x 2



1	1	2	2
1	1	2	2
3	3	4	4
3	3	4	4

Output: 4 x 4

“Bed of Nails”

1	2
3	4

Input: 2 x 2



1	0	2	0
0	0	0	0
3	0	4	0
0	0	0	0

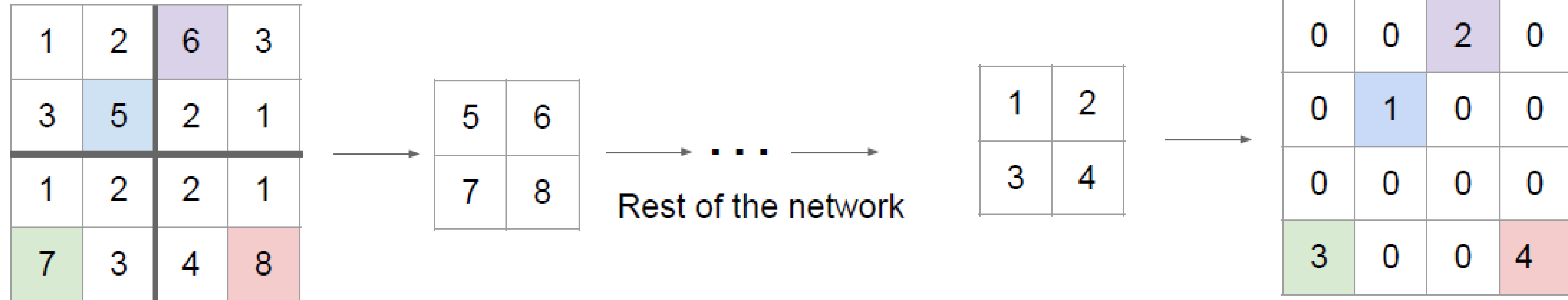
Output: 4 x 4



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

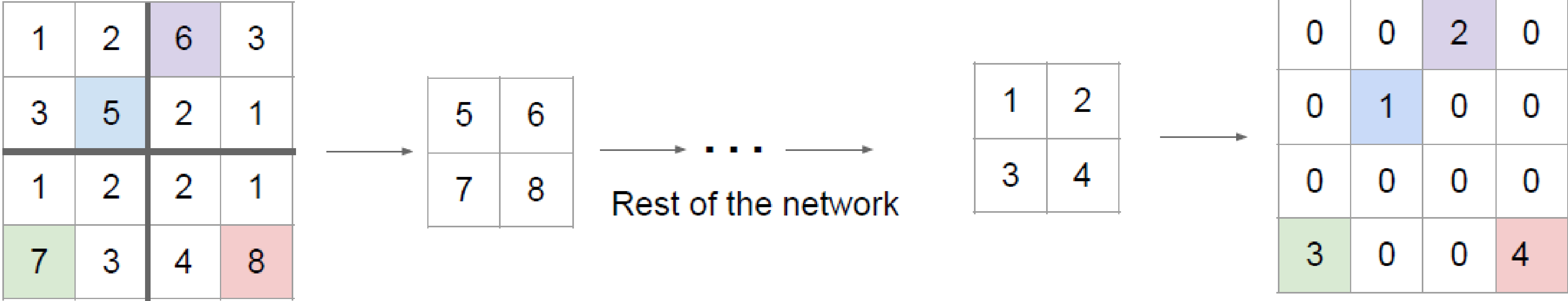
- How upsampling?
 - Pooling Layer → Unpooling Layer



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- **How upsampling?**
 - Pooling Layer → Unpooling Layer



Remember which
element was max

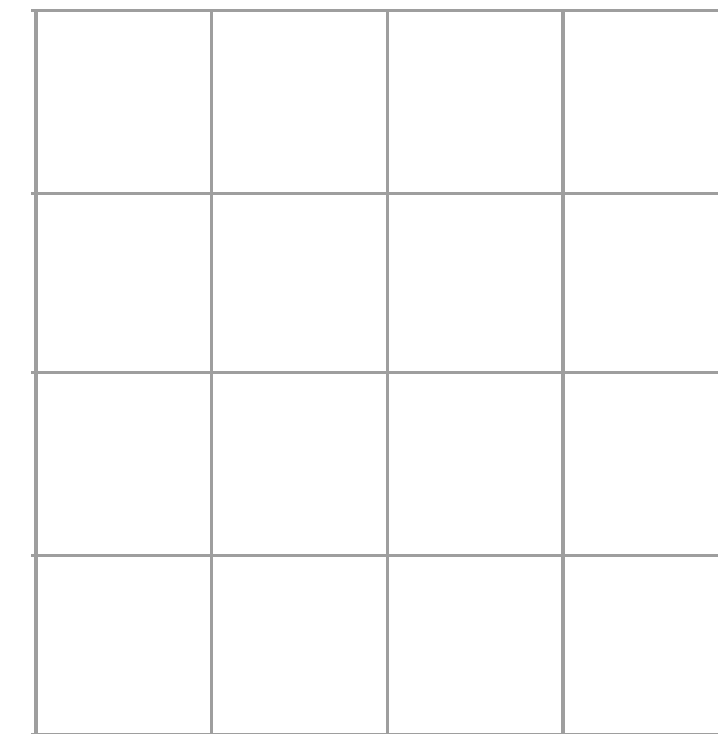
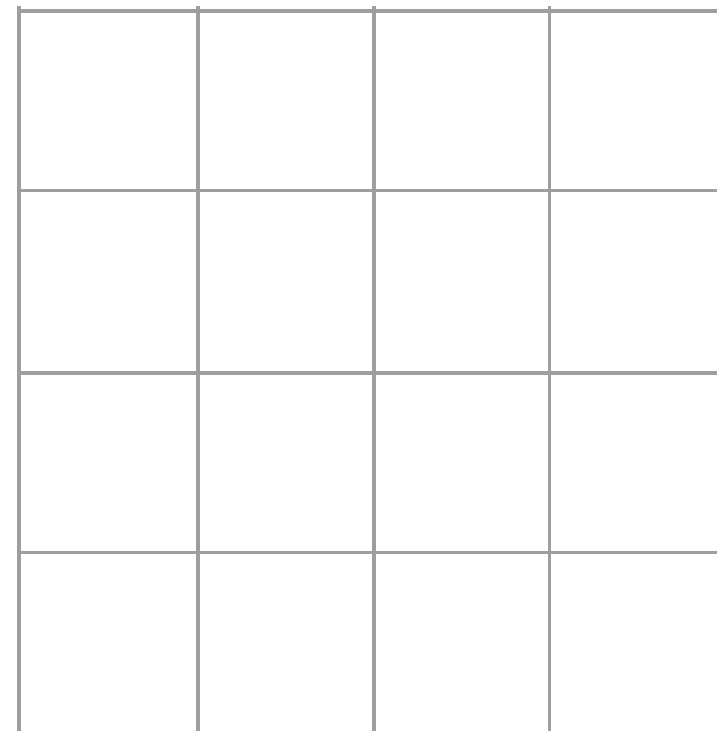
Use positions from
pooling layer



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- **How upsampling?**
 - Convolutional Layer → Deconvolutional Layer



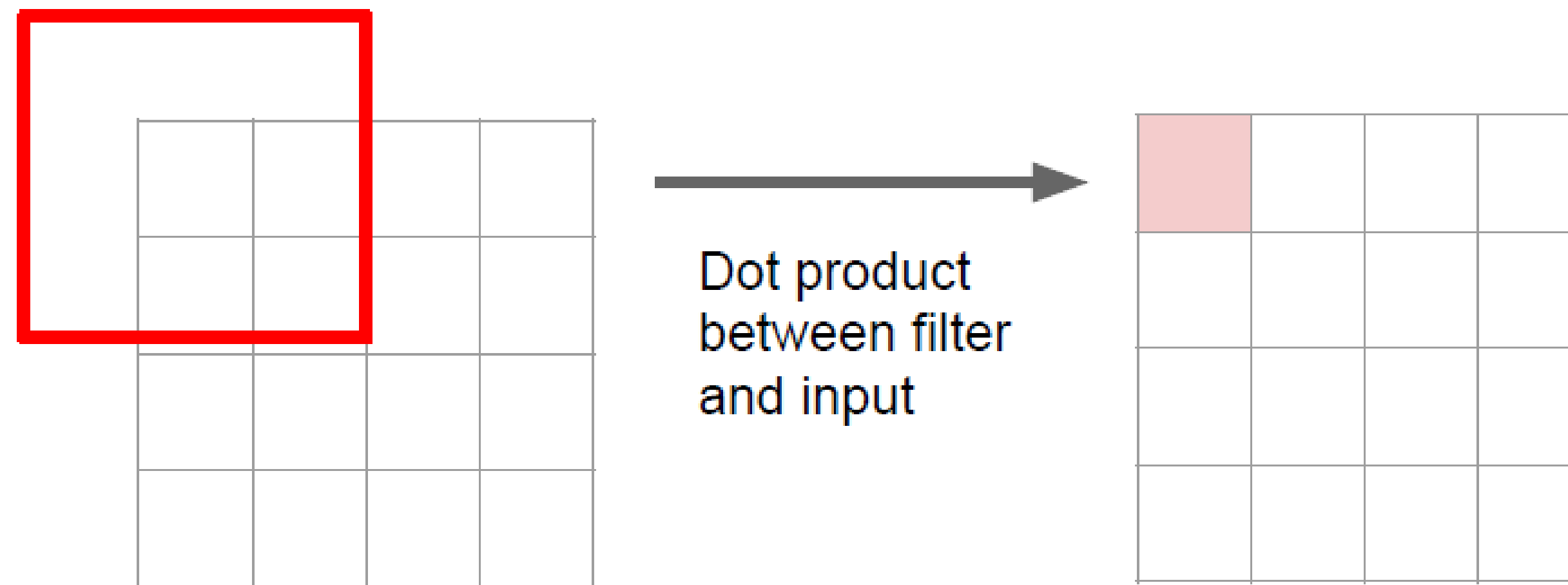
Typical 3 x 3 convolution, stride 1 pad 1



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- How upsampling?
 - Convolutional Layer → Deconvolutional Layer



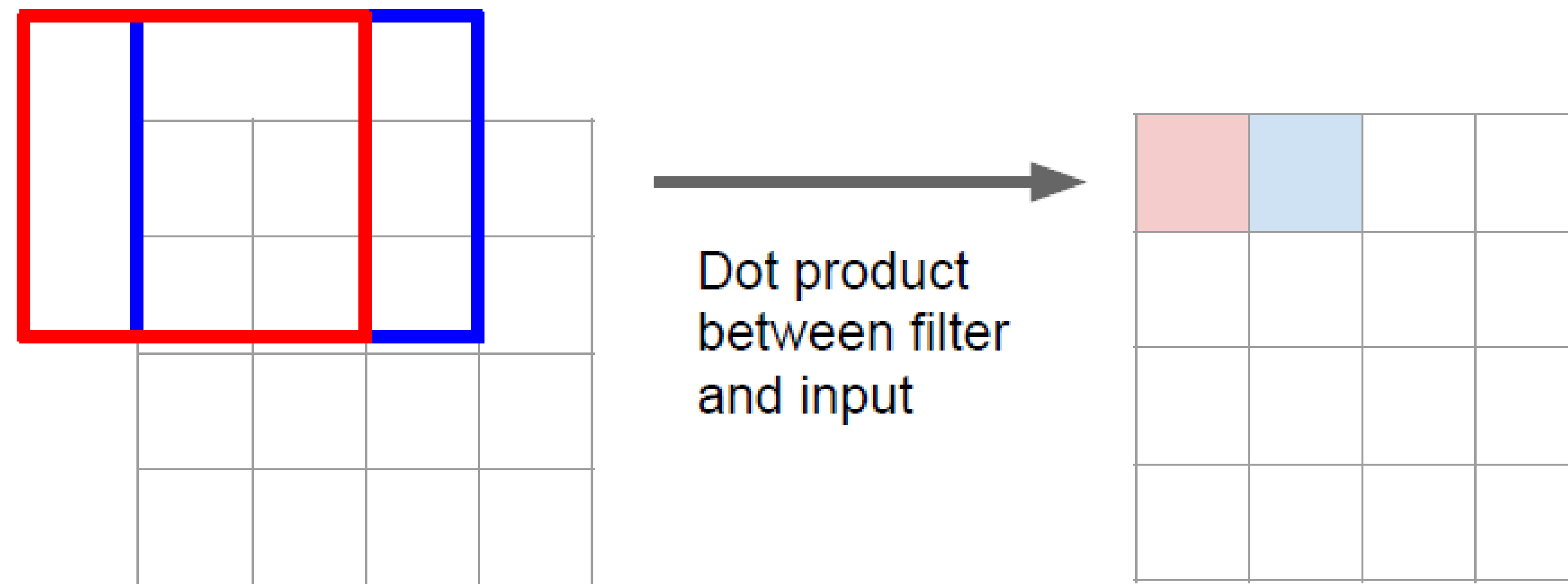
Typical 3 x 3 convolution, stride 1 pad 1



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- How upsampling?
 - Convolutional Layer → Deconvolutional Layer



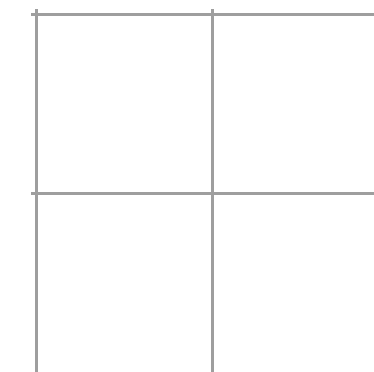
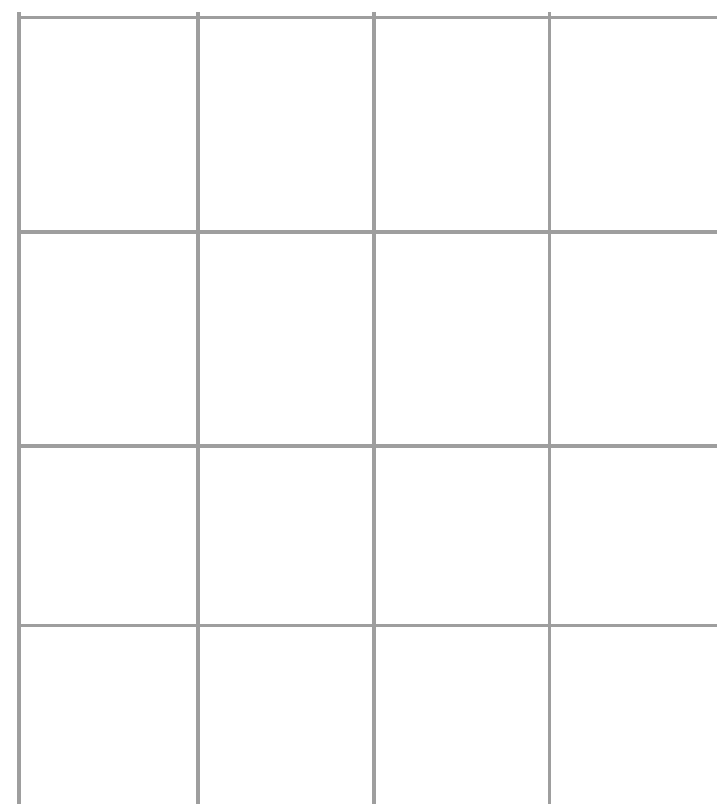
Typical 3 x 3 convolution, stride 1 pad 1



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- **How upsampling?**
 - Convolutional Layer → Deconvolutional Layer



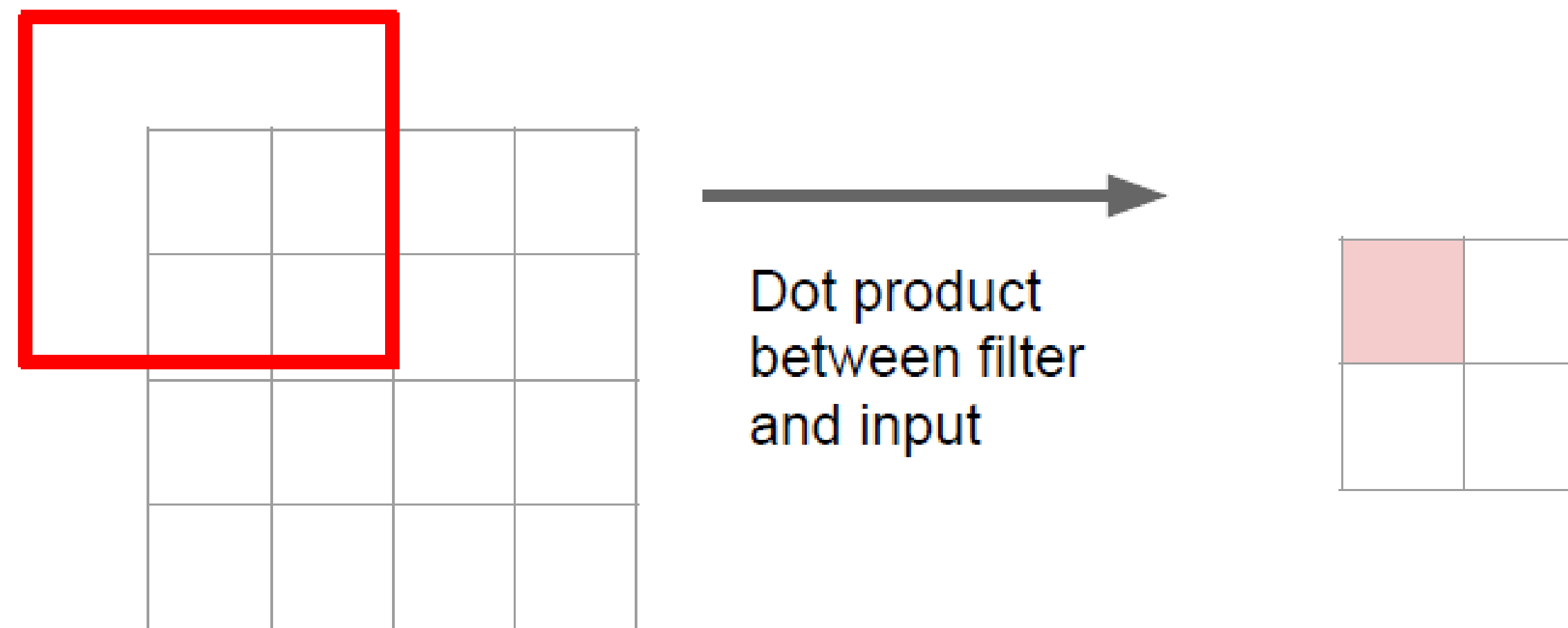
Normal 3 x 3 convolution, **stride 2** pad 1



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- How upsampling?
 - Convolutional Layer → Deconvolutional Layer



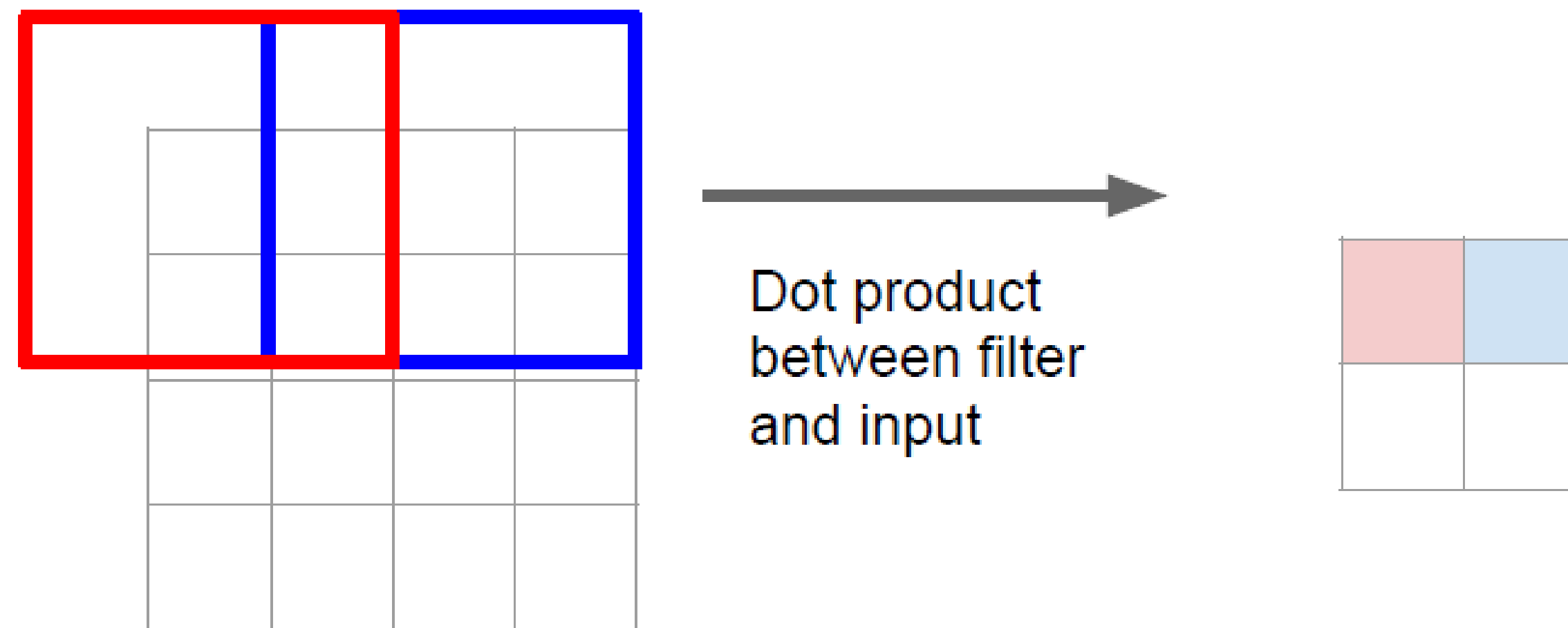
Normal 3 x 3 convolution, **stride 2** pad 1



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- How upsampling?
 - Convolutional Layer → Deconvolutional Layer



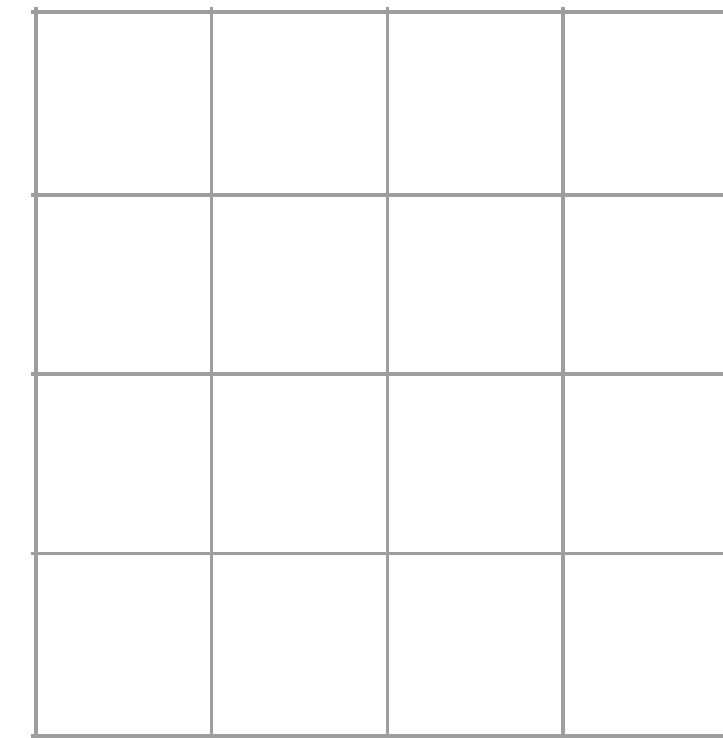
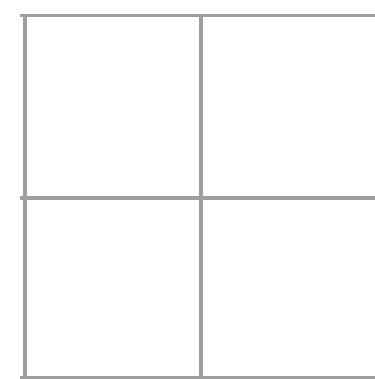
Normal 3 x 3 convolution, **stride 2** pad 1



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- **How upsampling?**
 - Convolutional Layer → Deconvolutional Layer



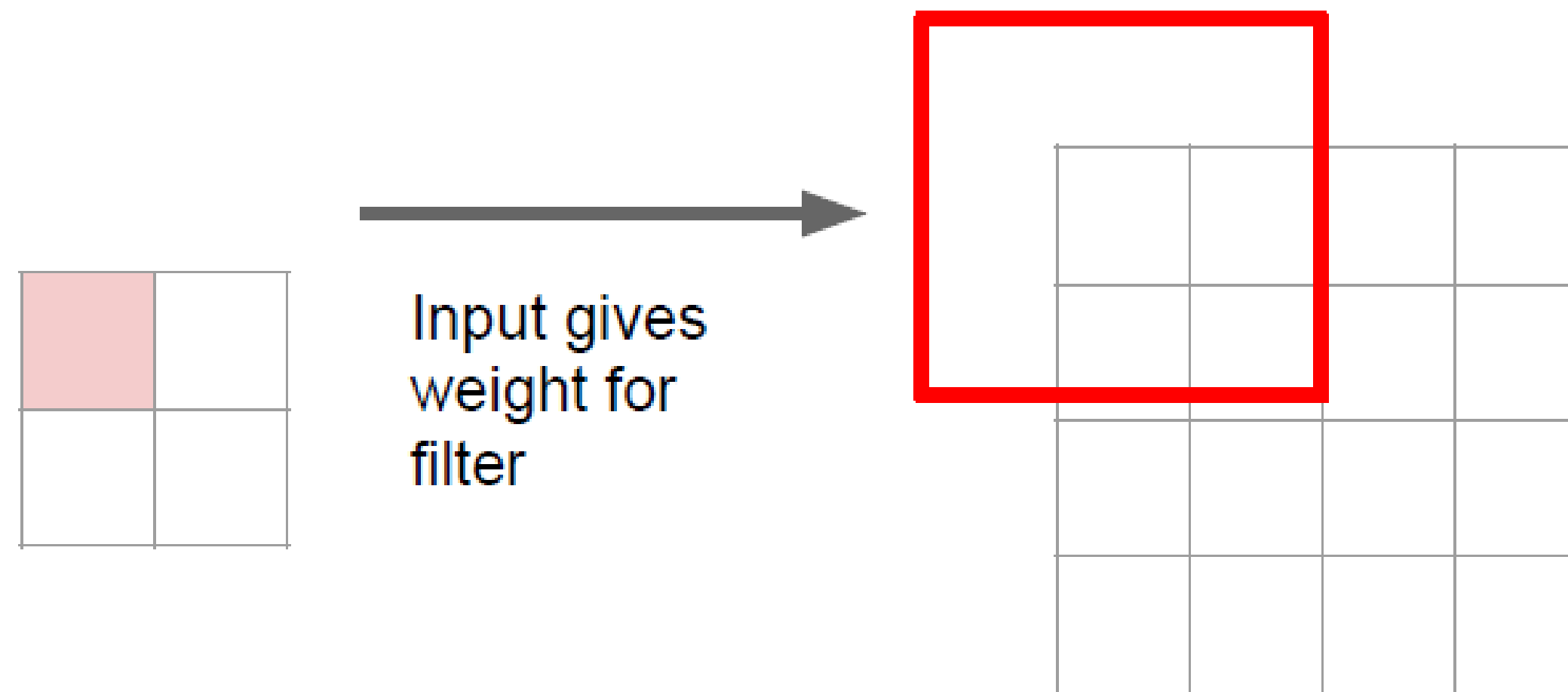
3 x 3 deconvolution, stride 2 pad 1



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- How upsampling?
 - Convolutional Layer → Deconvolutional Layer



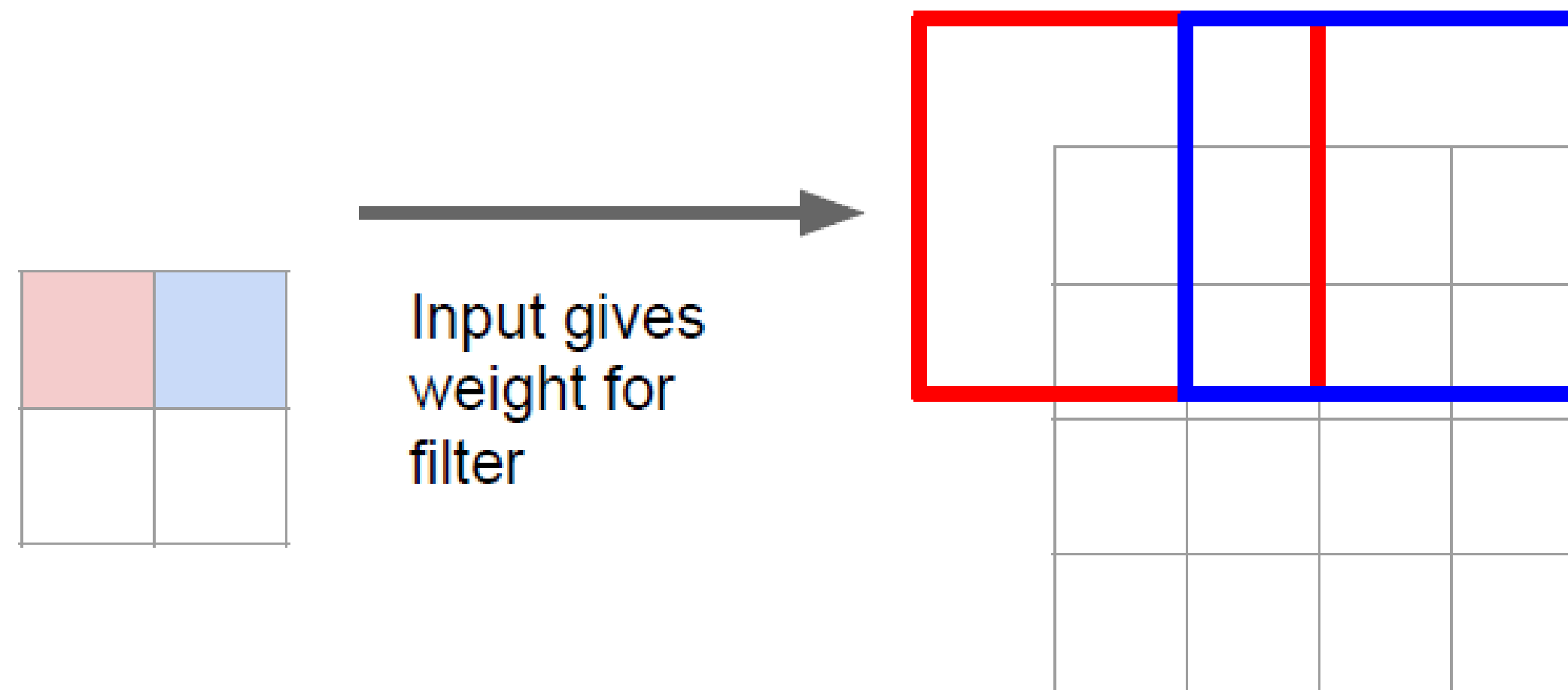
3 x 3 deconvolution, stride 2 pad 1



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- How upsampling?
 - Convolutional Layer → Deconvolutional Layer



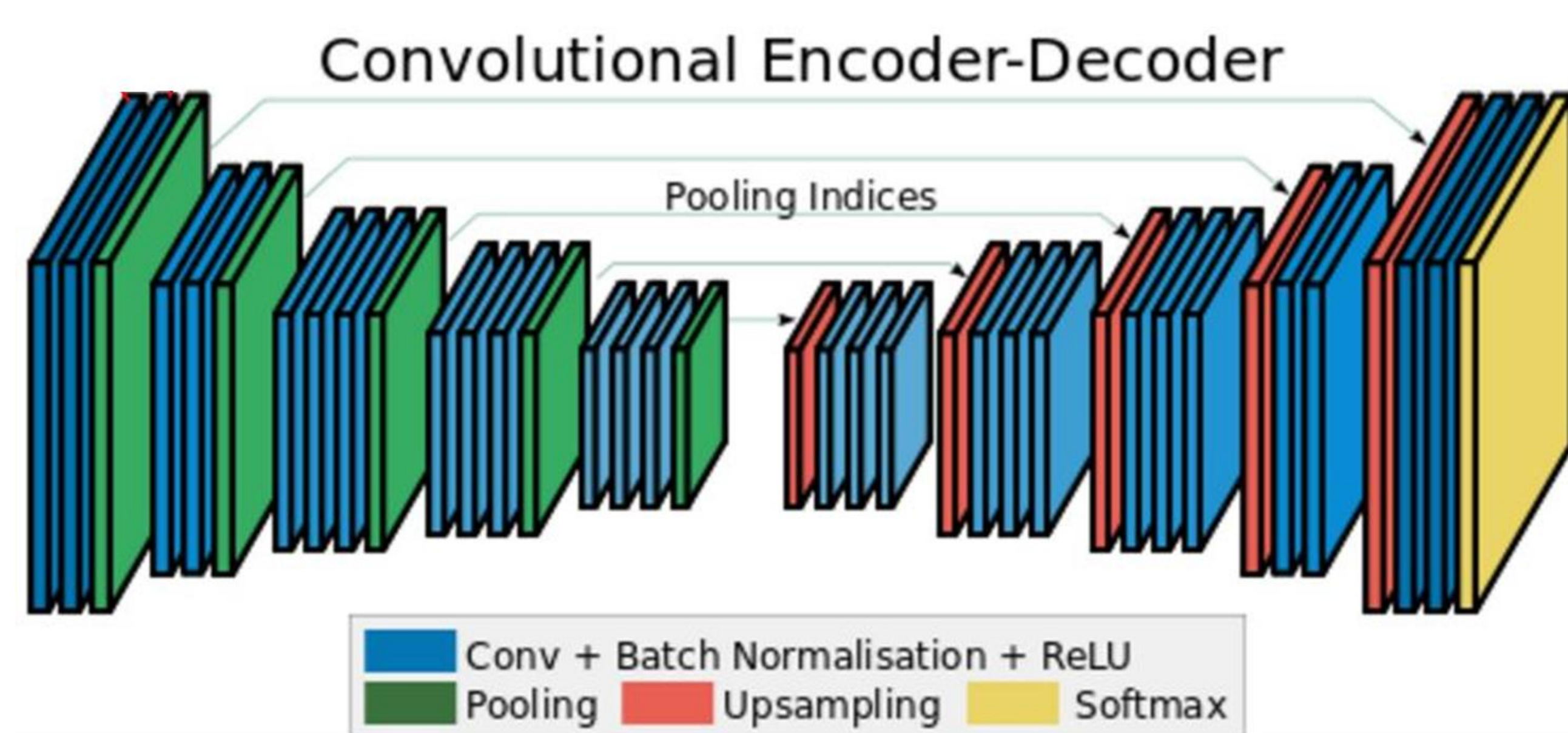
3 x 3 deconvolution, stride 2 pad 1



Clasificación Automática de Ortofotografías Aéreas del PNOA

Semantic Segmentation

- **SegNet Architecture**
 - Symmetric
 - Unpooling: store pooling indices for reconstruction



Caso de Uso: Clasificación Automática de Ortofotografías Aéreas del PNOA

Clasificación Automática de Ortofotografías Aéreas del PNOA

Contenidos

- Introducción
- Caso de Uso
- Metodología
- Dataset
- Resultados
- Recursos Computacionales
- Conclusiones



Clasificación Automática de Ortofotografías Aéreas del PNOA

Introducción

Objetivo dual

- Desarrollo de una herramienta para la segmentación automática de imágenes basada en CNN's.
- Análisis de los recursos computacionales del CESGA para este tipo de aplicaciones.



Clasificación Automática de Ortofotografías Aéreas del PNOA

Caso de Uso

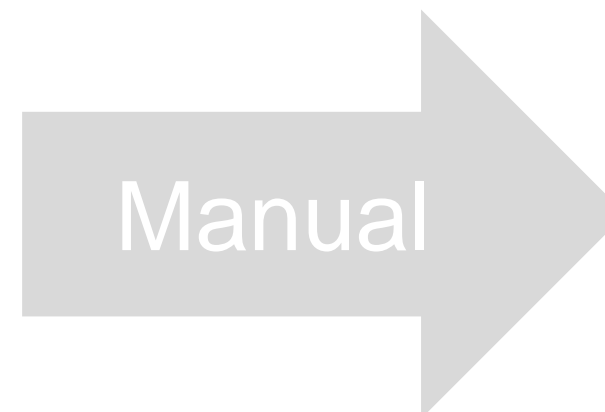
Elaboración automática de la capa SIOSE.

Actualidad:

- Procesado manual con diversas fuentes de referencia (satélite, PNOA,...)



Diversas fuentes



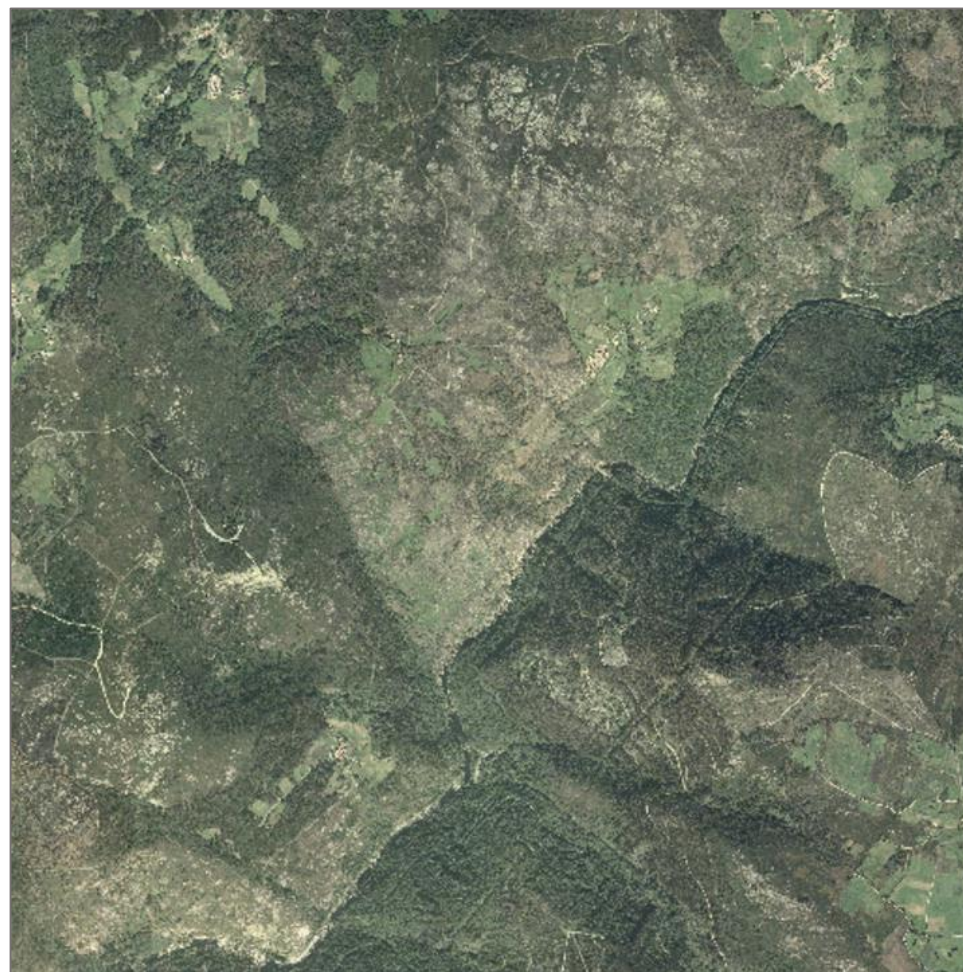
SIOSE



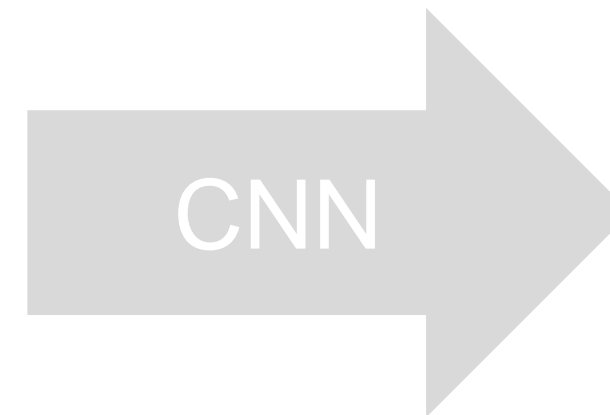
Clasificación Automática de Ortofotografías Aéreas del PNOA

Caso de Uso

Elaboración automática de la capa SIOSE.



PNOA



Predicción (arbolado forestal)

Ventajas:

- Automatización del proceso.
- Ahorro de tiempo.



Clasificación Automática de Ortofotografías Aéreas del PNOA

Metodología y Desarrollo del Proyecto

- Preparación de los datos
 - Re-etiquetado del ground-truth.
 - Extracción de Imágenes.
 - Creación de ficheros LMDB.
 - Entrenamiento del modelo.
- Generación de predicciones
- Evaluación de los resultados



Clasificación Automática de Ortofotografías Aéreas del PNOA

Dataset

Dos fuentes de información proporcionadas por el CESGA*:

- Imágenes RGB, procedentes del PNOA (años 2005 y 2010)
- Imágenes SIOSE (años 2006 y 2011)

+ 150 imágenes de 54000x38000 píxeles (+ 2000 millones de píxeles cada una)

* Procesado realizado CESGA: (reescalado, recortado, filtrado)



Clasificación Automática de Ortofotografías Aéreas del PNOA

Dataset. Dificultades

- *Imágenes del SIOSE con fotointerpretación incompleta*

Sucede cuando no toda la superficie de una imagen está etiquetada.

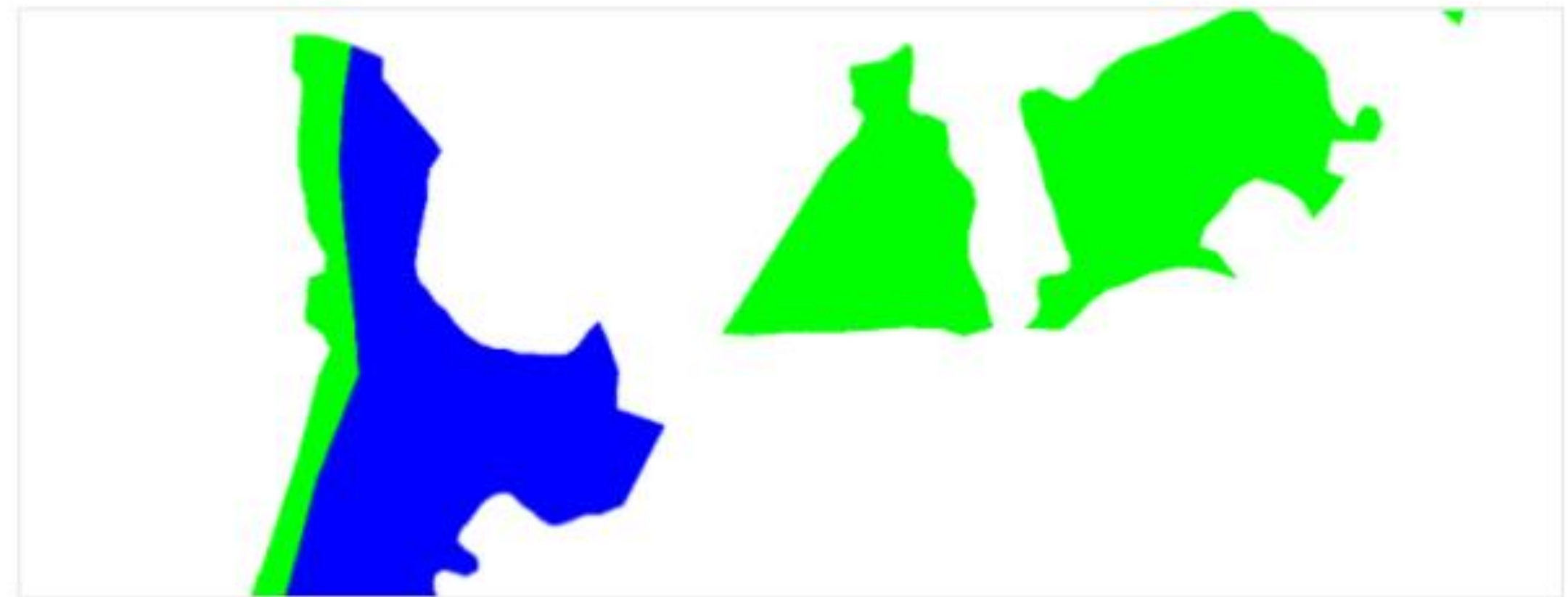


Ilustración 4. Arriba: Detalle de imagen del PNOA. Abajo: Correspondiente capa de SIOSE con cobertura forestal (azul oscura), otra (verde) y no interpretada (blanco)



Clasificación Automática de Ortofotografías Aéreas del PNOA

Dataset. Dificultades

- ***Imágenes del SIOSE con fotointerpretación incompleta***

Sucede cuando no toda la superficie de una imagen está etiquetada.

- ***Fotointerpretación inexacta.***

La interpretación no se corresponde al 100% con la realidad de la imagen PNOA a la que hace referencia.

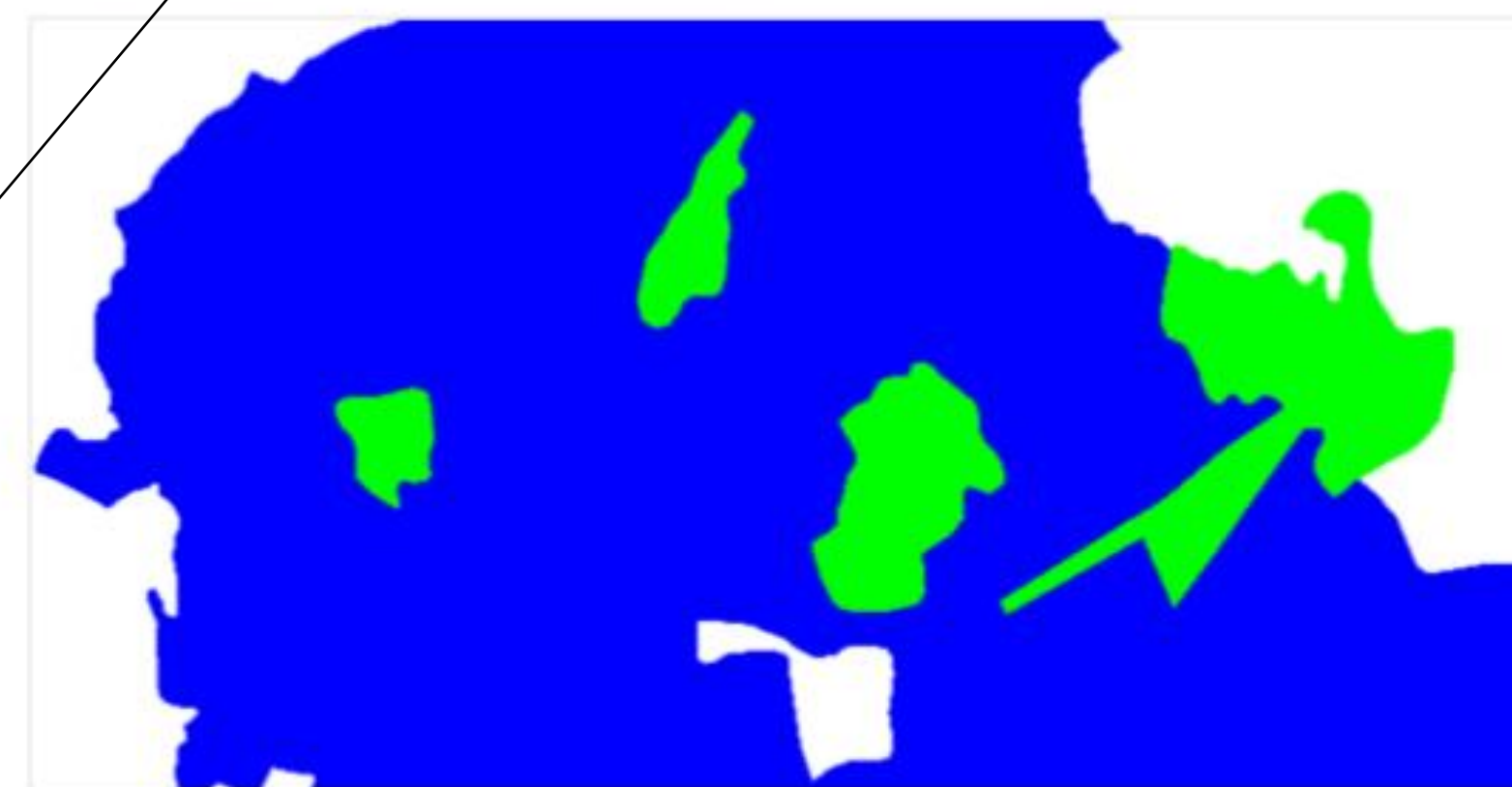


Ilustración 5. Arriba: Detalle de imagen del PNOA. Abajo: Correspondiente capa de SIOSE con cobertura forestal (azul oscura), otra (verde) y no interpretada (blanco). Los verdes son respectivamente Pastizal, Pastizal, Asentamiento agrícola, Matorral



Clasificación Automática de Ortofotografías Aéreas del PNOA

Dataset. Dificultades

- ***Imágenes del SIOSE con fotointerpretación incompleta***

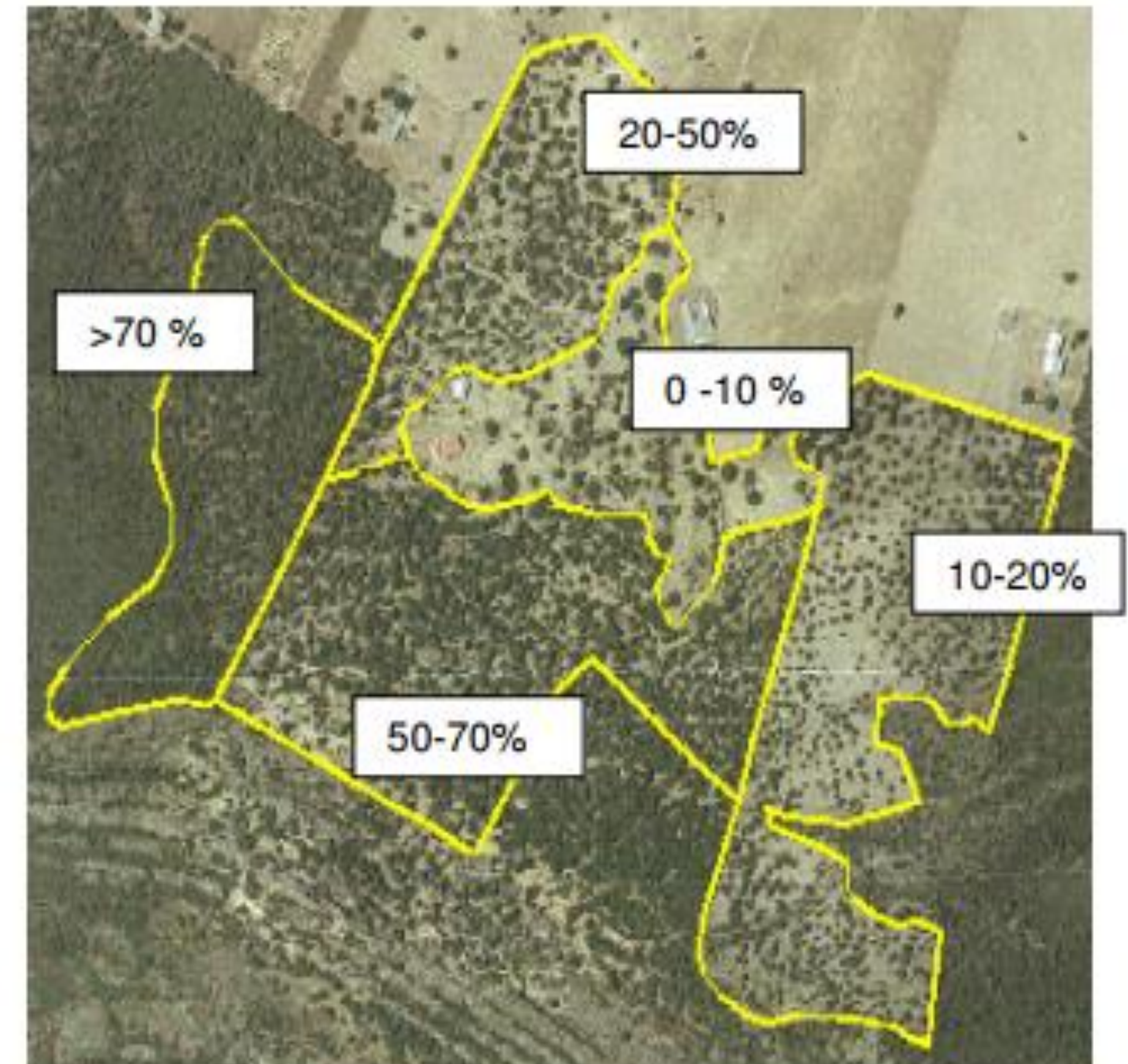
Sucede cuando no toda la superficie de una imagen está etiquetada.

- ***Fotointerpretación inexacta.***

La interpretación no se corresponde al 100% con la realidad de la imagen PNOA a la que hace referencia.

- ***Existencia de coberturas compuestas***

Cuando una cierta área aparece identificada con una cobertura compuesta, es decir, se le ha asignado más de un tipo de cobertura y cada una de ellas con un cierto porcentaje.

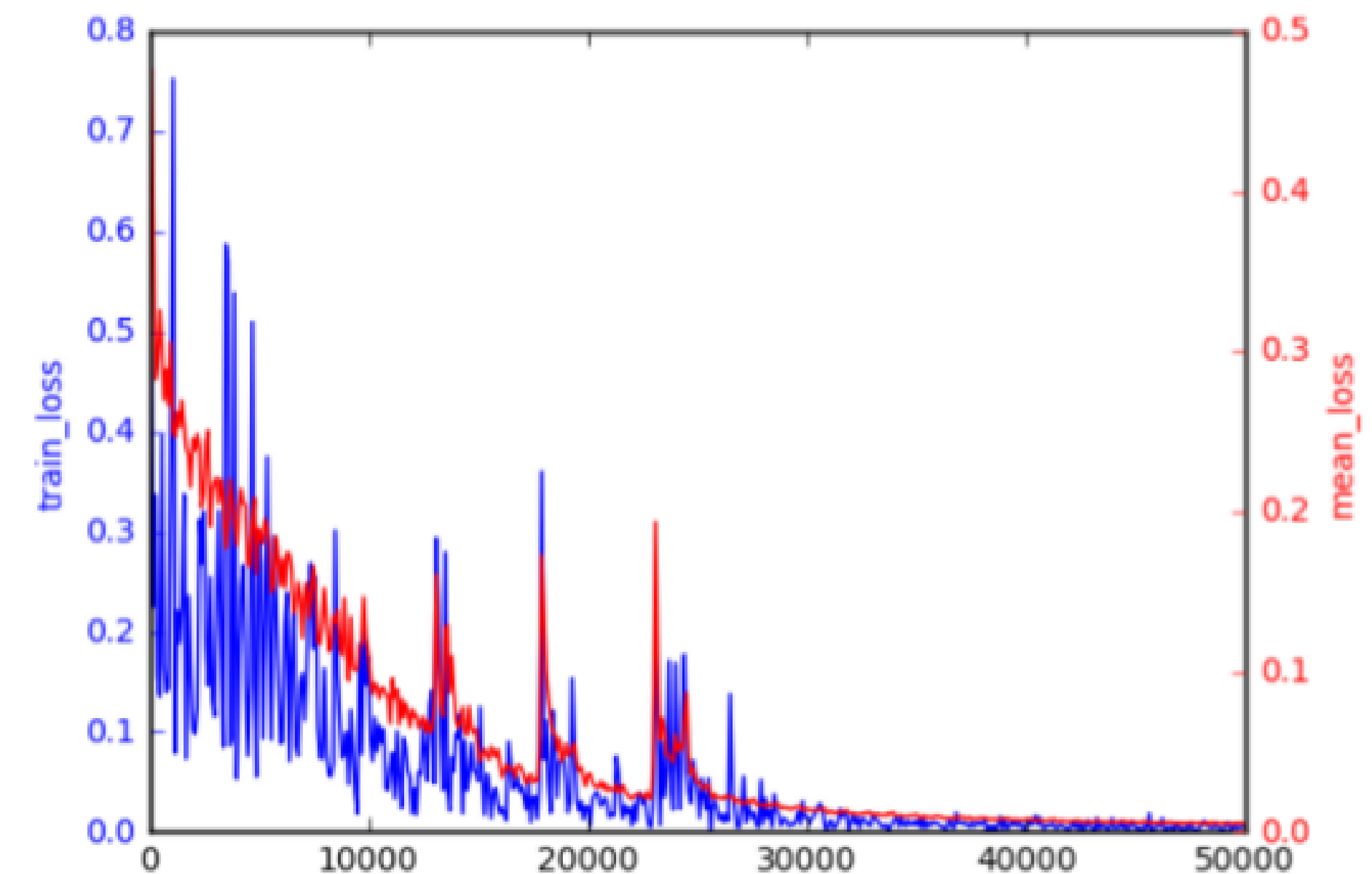
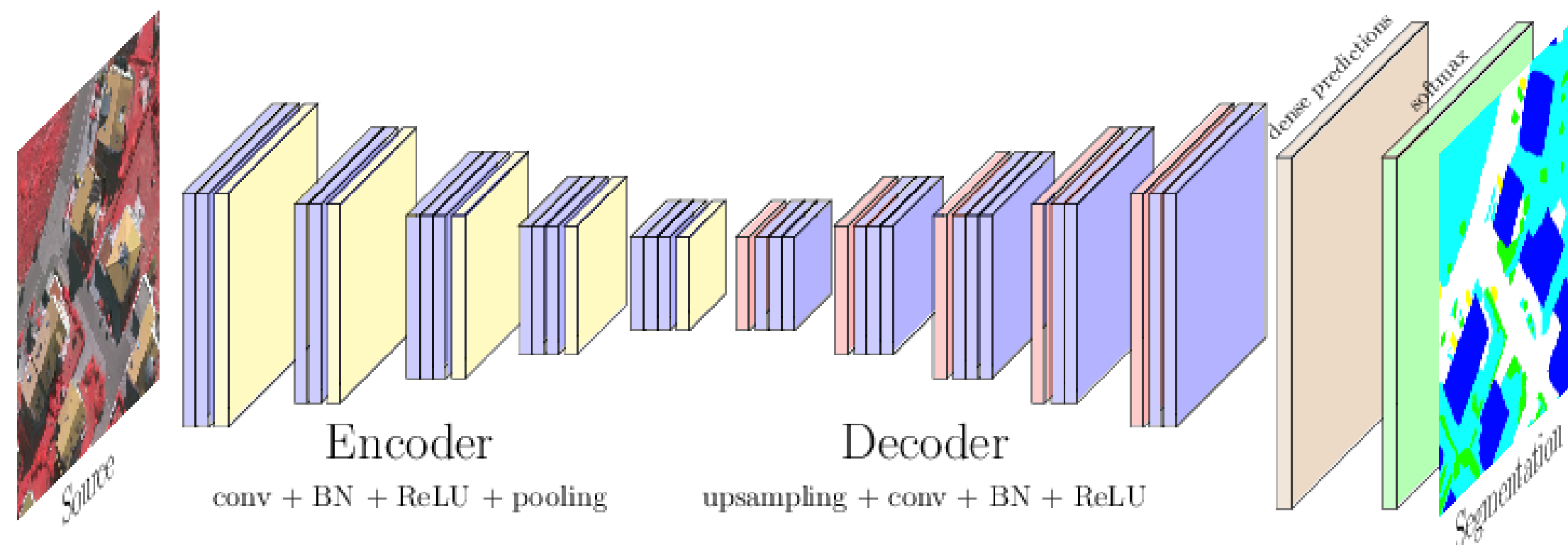


Clasificación Automática de Ortofotografías Aéreas del PNOA

Resultados. Entrenamiento del modelo.

Entrenamiento del modelo

- SegNet
- 50000 iteraciones



Clasificación Automática de Ortofotografías Aéreas del PNOA

Resultados: caso A



Imagen original PNOA



Fotointerpretación SIOSE: ■ No forestal; ■ Forestal; □ Sin fotointerpretación

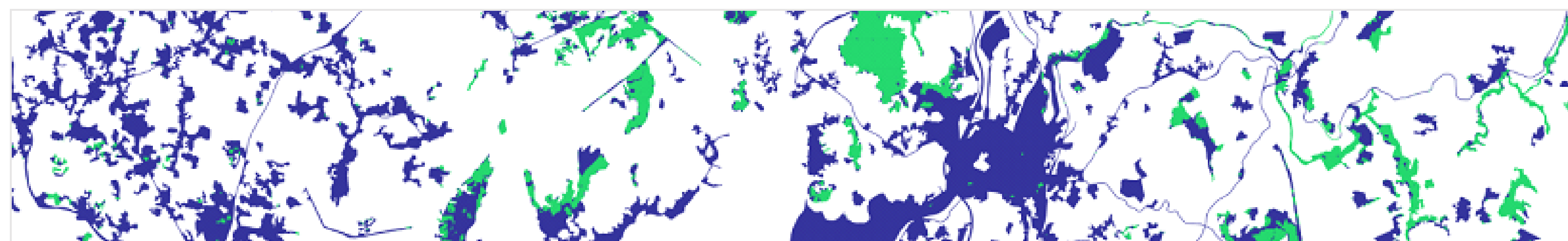


Color: Píxel de *groun-truth* con color PNOA original; Negro: Píxel sin fotointerpretación SIOSE

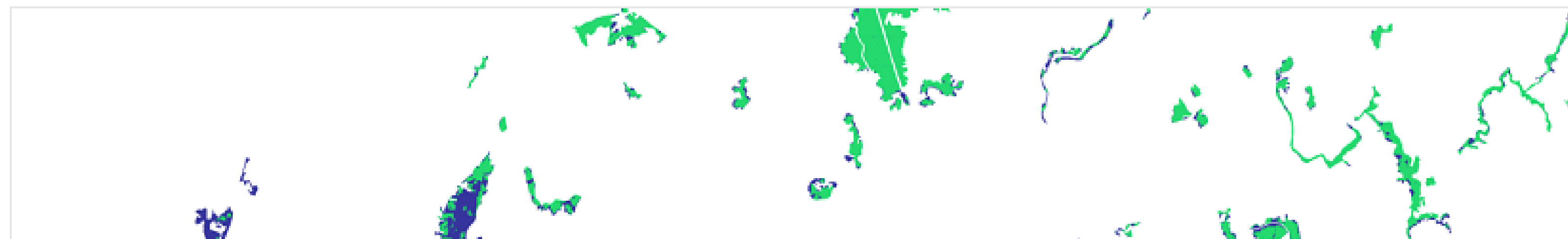


Clasificación Automática de Ortofotografías Aéreas del PNOA

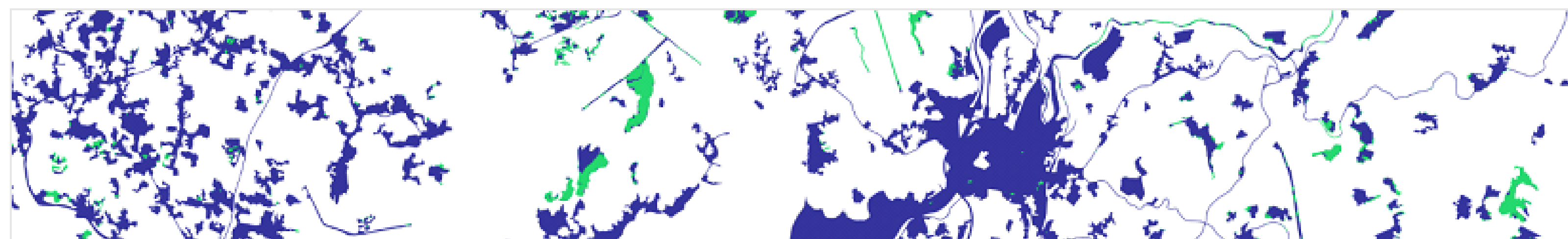
Resultados: caso A



Predicción: ■ No forestal; ■ Forestal; ■ Sin fotointerpretación



Predicción en parcelas forestales: ■ Error (FN); ■ Correcto (TP); ■ Sin fotointerpretación



Predicción en parcelas no forestales: ■ Correcto (TN); ■ Error (FP); ■ Sin fotointerpretación



Clasificación Automática de Ortofotografías Aéreas del PNOA

Resultados: caso A



Imagen original PNOA



Fotointerpretación SIOSE: ■ No forestal; ■ Forestal; ■ Sin fotointerpretación



Color: Píxel de *groun-truth* con color PNOA original; Negro: Píxel sin fotointerpretación SIOSE



Clasificación Automática de Ortofotografías Aéreas del PNOA

Resultados: caso A



Clasificación Automática de Ortofotografías Aéreas del PNOA

Resultados: matrices de confusión

- Matriz de Confusión

		Clase Real	
		Arbolado Forestal	Resto
Clase Predicha	Arbolado Forestal	847682090	200993073
	Resto	297048081	4053969030

- Matriz de Confusión Normalizada

		Clase Real	
		Arbolado Forestal	Resto
Clase Predicha	Arbolado Forestal	0,81	0,19
	Resto	0,07	0,93

Clasificación Automática de Ortofotografías Aéreas del PNOA

Análisis de Recursos Computacionales

Tiempo de computación.

- Deep Learning computacionalmente muy intensivo → Máquinas de altas prestaciones

Capacidad de almacenamiento y recursos de memoria.

- Tamaño del dataset necesario para Deep Learning
 - 42 imágenes del PNOA por año (3 años) de un tamaño aproximado de más de 6,4 GB cada una
 - 42 imágenes del SIOSE por año (2 años) de más de 4,3 GB cada una.

+ 1 TB de información en imágenes!



Clasificación Automática de Ortofotografías Aéreas del PNOA

Análisis de Recursos Computacionales

Infraestructura CESGA:

- Almacenamiento en disco:
 - LUSTRE 6TB y 1000000 de ficheros (capacidad de almacenamiento en local no resultaba suficiente)
- Computación:
 - Finis Terrae II: FAT node para la fase de creación de los ficheros LMDB correspondientes a los dataset de entrenamiento y test (350 GB de memoria)
 - Finis Terrae II: Thin nodes, utilizados para el resto de procesos.



Clasificación Automática de Ortofotografías Aéreas del PNOA

Análisis de Recursos Computacionales

Tiempos de Computación.

- Extracción de imágenes

	Imágenes (PNOA)	Etiquetas (SIOSE)
Entrenamiento	176 minutos	16.4 minutos
Validación	85'18 minutos	11'9 minutos

- Creación de ficheros LMDB

	Creación LMDB (PNOA + SIOSE)
Entrenamiento	14'28 minutos
Validación	9'83 minutos

- Entrenamiento del modelo

	Entrenamiento del modelo
Entrenamiento	16'17 horas

Clasificación Automática de Ortofotografías Aéreas del PNOA

Conclusiones

Viabilidad del Sistema

- Utilización de redes neuronales profundas es una herramienta adecuada para la clasificación semántica de coberturas del terreno
- Accuracy \approx 90%
- Resultados extrapolables a otras categorías de alto nivel
- Posibilidad de clasificaciones en categorías de más bajo nivel

Ventajas en la utilización de la infraestructura del CESGA.

- Reducción de los tiempos de computación (resultado esperado)
- Capacidades CESGA (ahora mismo imposibilidad de replicar los resultados en otro entorno)

Inconvenientes.

- Dificultades adecuación de algoritmos y distribución Caffe al entorno CESGA
- Pérdida de flexibilidad debido al sistema de colas



Clasificación Automática de Ortofotografías Aéreas del PNOA

Conclusiones. Trabajo Futuro

Tiempos de Cómputo

- En la fase de preparación de datos: código secuencial altamente paralelizable
- Fase entrenamiento: análisis de hiperparámetros

Exactitud en la predicción

- Arquitectura.
- Disponibilidad de datos de entrenamiento con un nivel de fotointerpretación más elevado
- Técnicas de Data Augmentation

Otros problemas

- Clasificación simultánea de más de dos categorías, lo que requeriría modificaciones en la arquitectura.
- Explorar otros escenarios y/o conjuntos de imágenes en otras bandas espectrales.



Semantic Segmentation

Semantic Segmentation

Áreas de interés:

- Análisis de escenas de carretera
 - Aprendizaje de carretera, coche, humano,... (real-time)



Semantic Segmentation

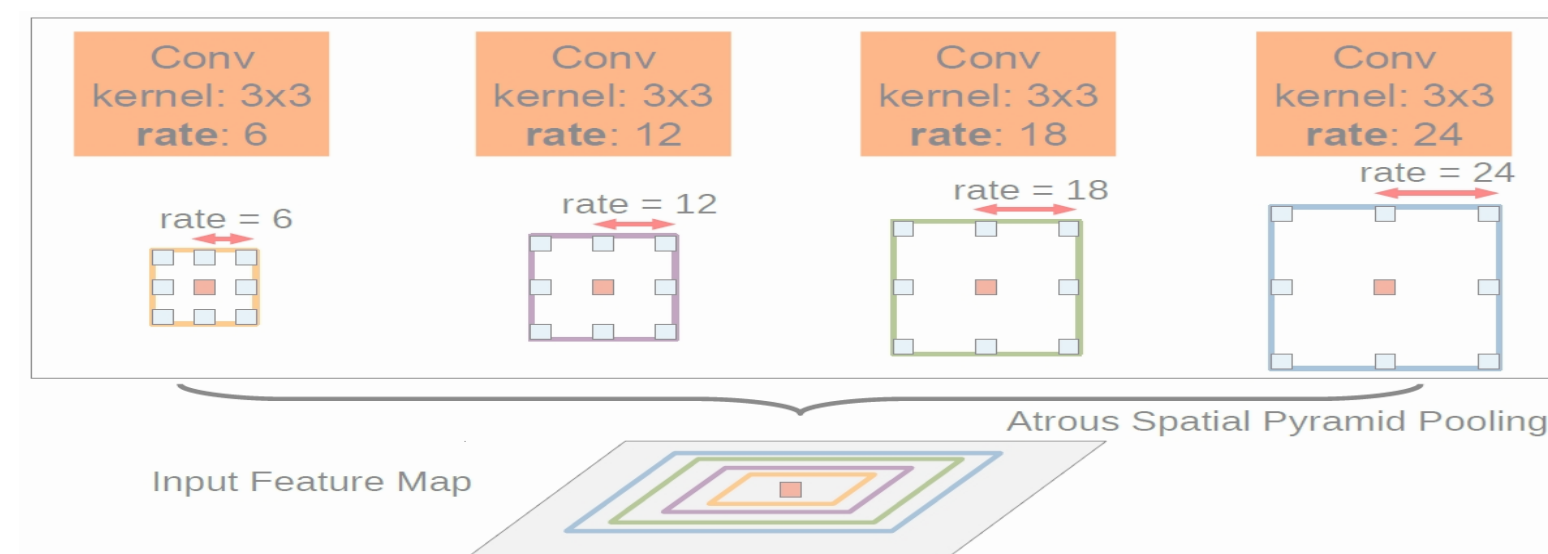
Otras arquitecturas:

Fully convolutional Network (FNC)

- Lower level activations added for upsampling detail reconstruction

DeepLab (FNC)

- Atrous spatial pyramid pooling for multi-scale recognition
- Conditional RandomFields (CRF) for Refinement

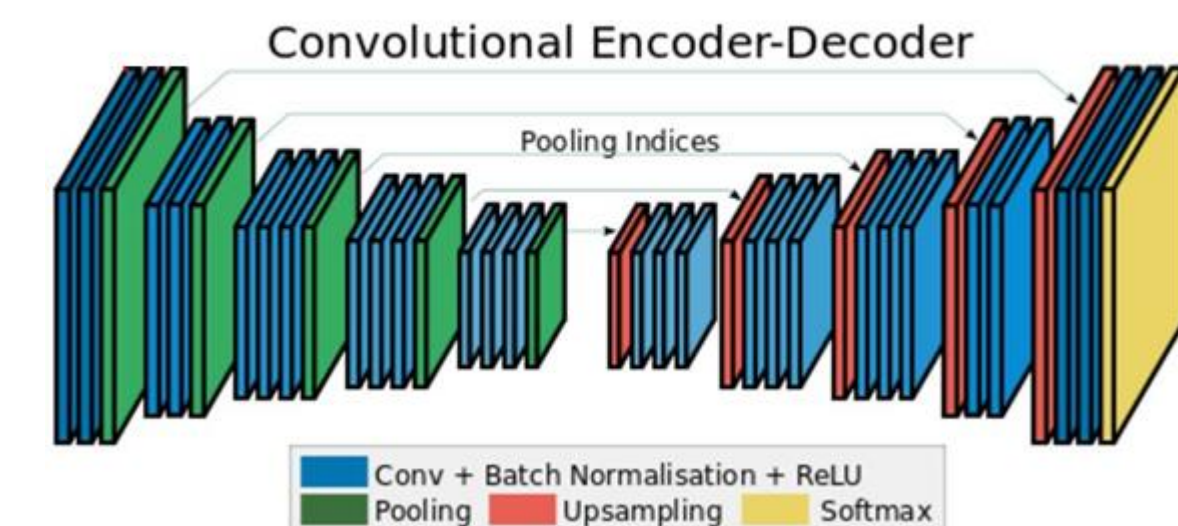


Enet

- Few data needed for training (no pretraining)
- Very Small + fast

SegNet

- Store pooling indices for reconstruction
- Symmetric





Orientados hacia las necesidades de la industria

(+34) 986 120 430 | gradiant@gradiant.org | www.gradiant.org