

Procesamiento de Imágenes Digitales

PROYECTO

SUPERRESOLUCIÓN EN IMÁGENES CAPTADAS POR DRON PARA LA DETECCIÓN DE FRAILEJONES

PAULA UZCATEGUI LEÓN
JOAN SEBASTIAN PEDRAZA



Delegación de Responsabilidades



Paula Uzcategui León

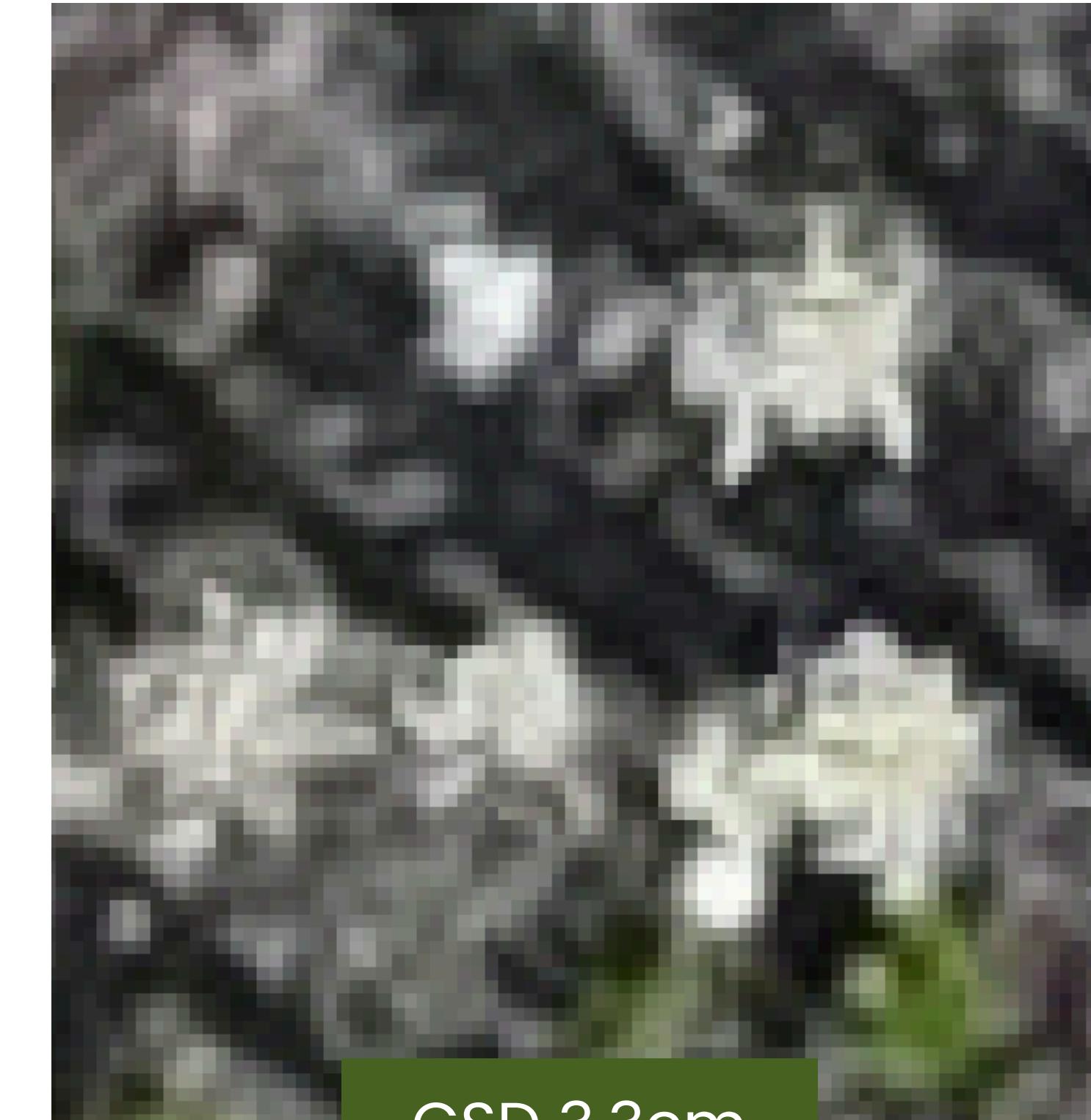
Generación del Dataset
Inferencia de modelo de
Superresolución y detección
Redacción del Informe
Creación de Diapositivas



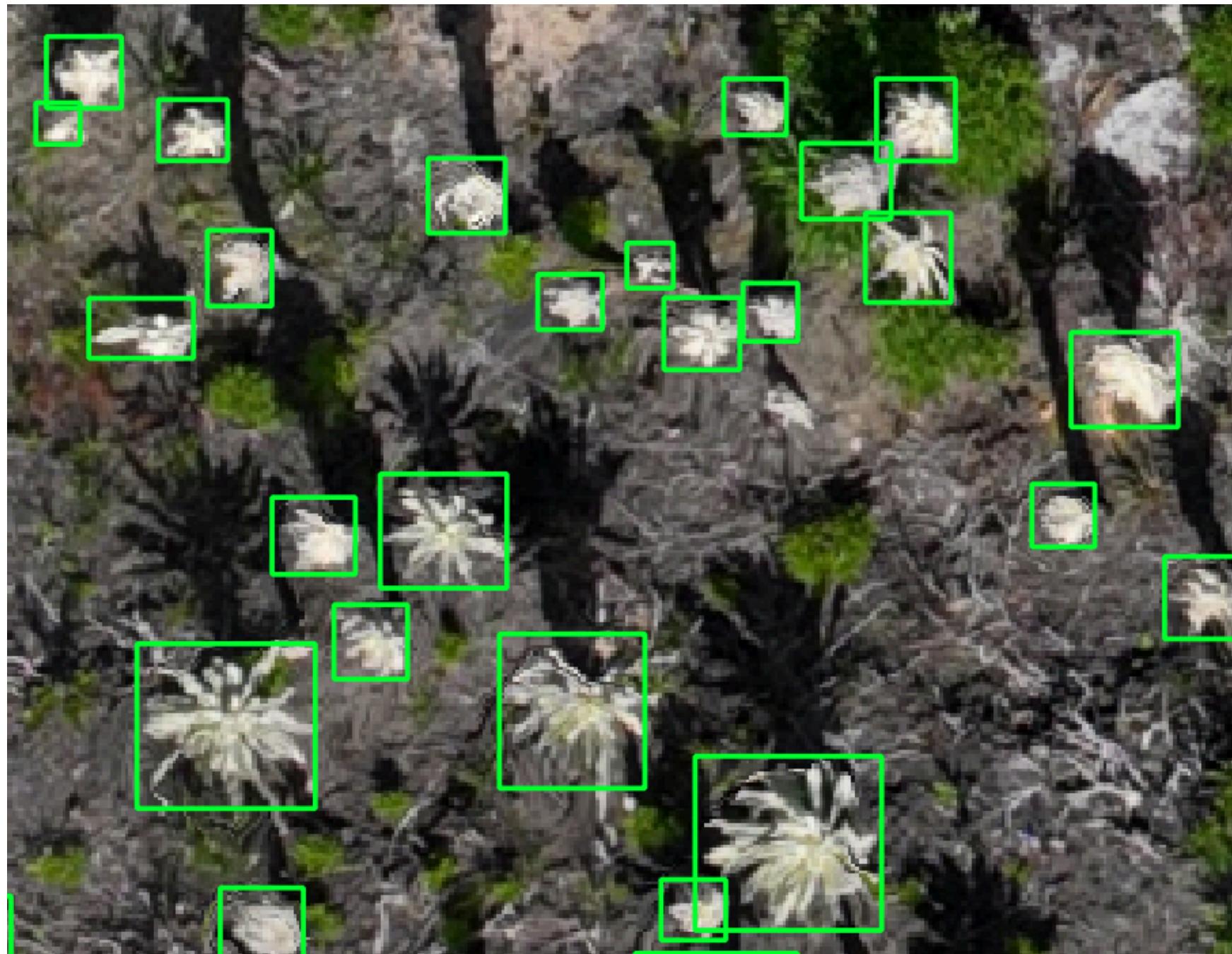
Joan Sebastian Pedraza

Depuración del Dataset
Implementación y entrenamiento
de Real-ESRGAN
Redacción del Informe
Creación de diapositivas

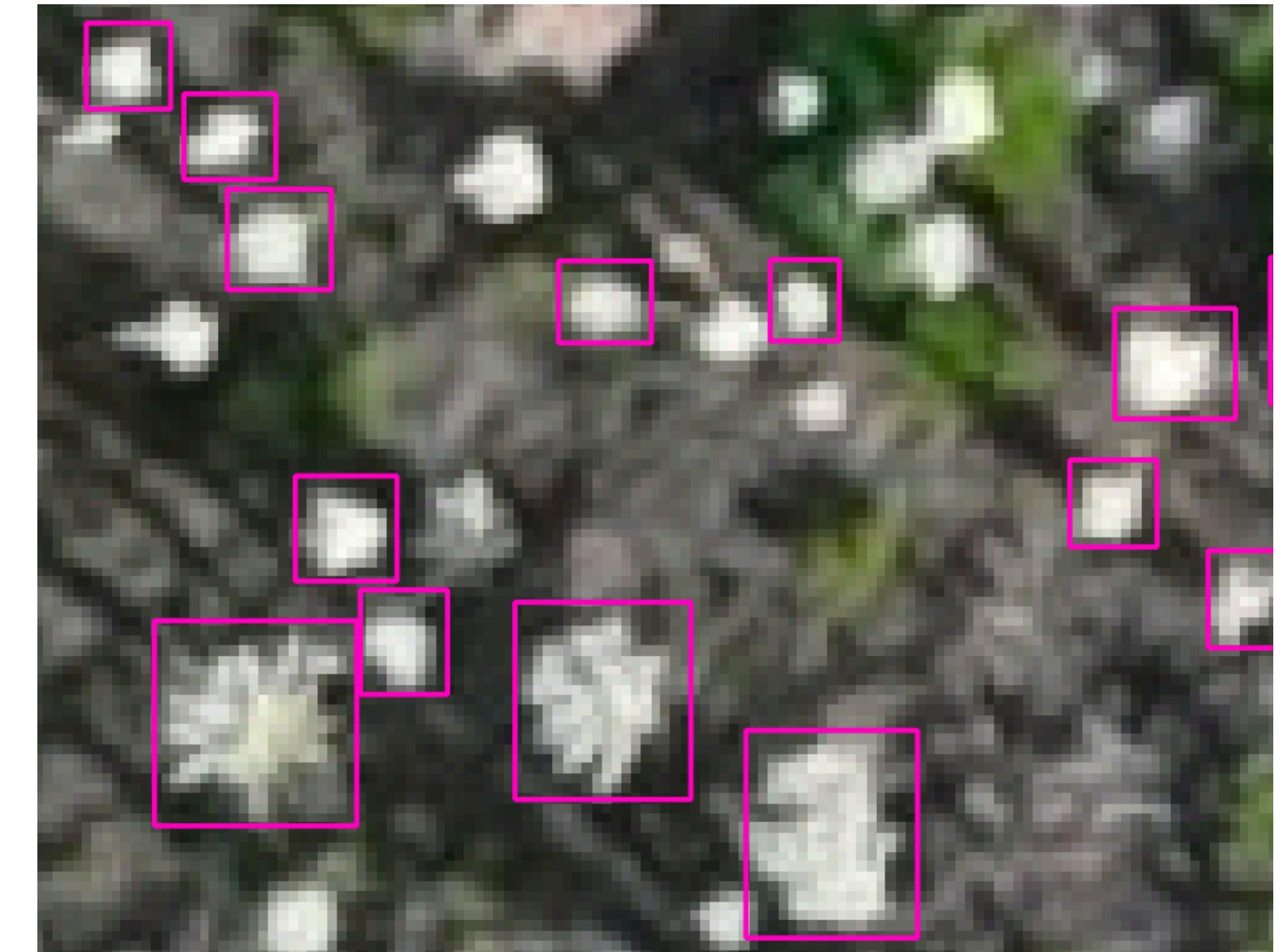
Objetivo: mejorar la resolución de imágenes de frailejones captadas por dron



Objetivo: mejorar la resolución de imágenes de frailejones captadas por dron

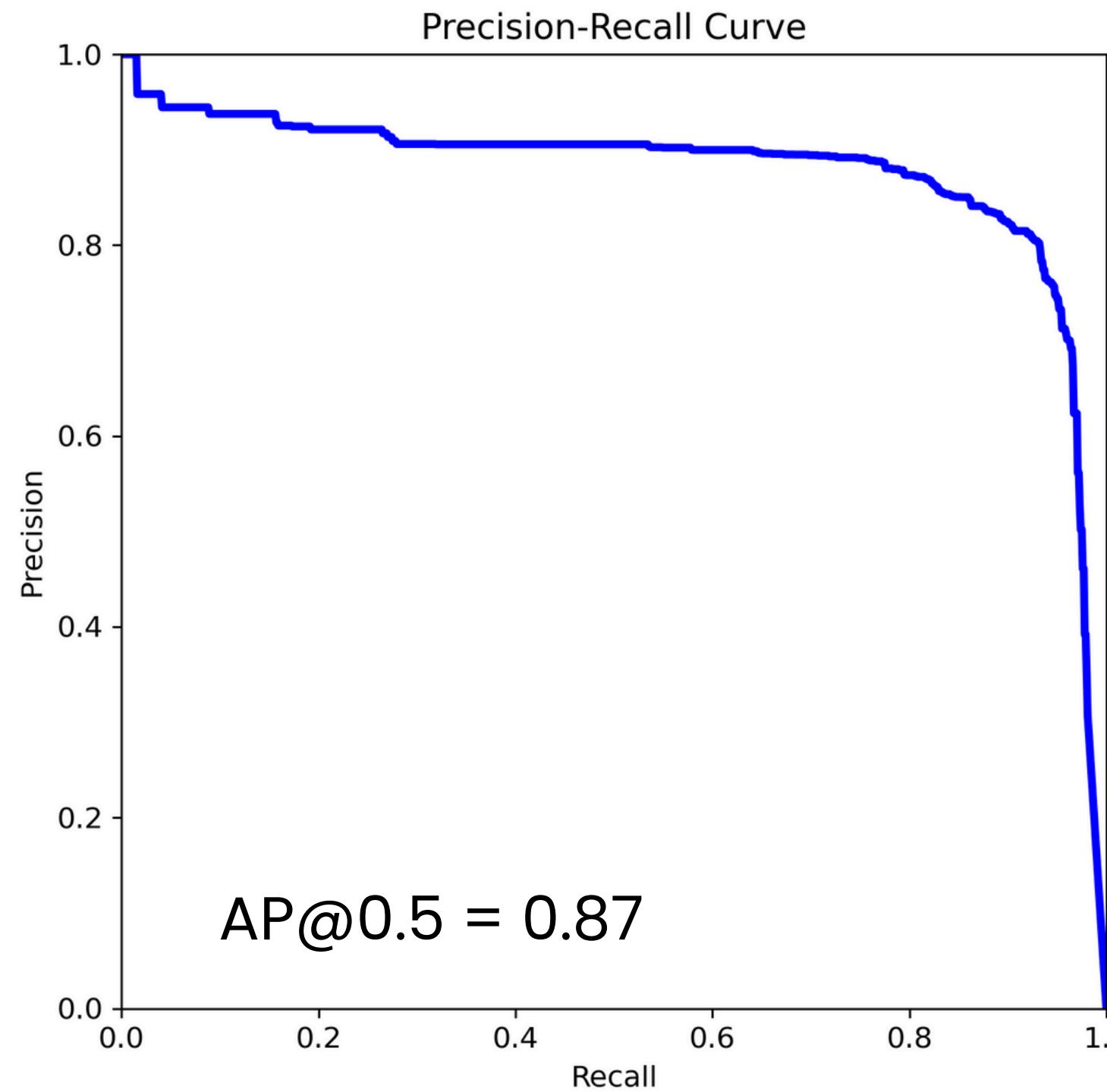


GSD 1.3cm

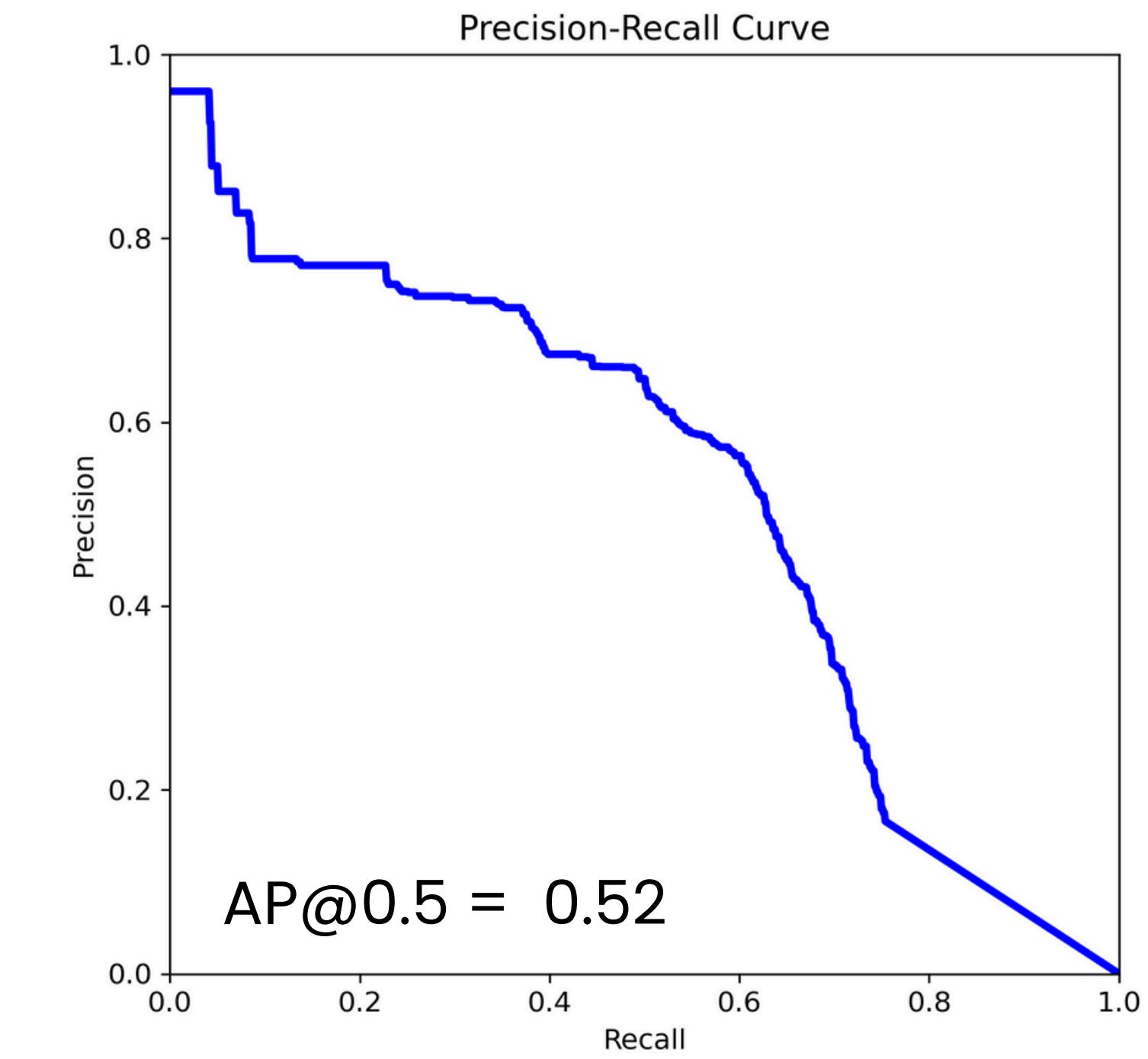


GSD 3.3cm

Objetivo: mejorar la resolución de imágenes de frailejones captadas por dron



GSD 1.3cm



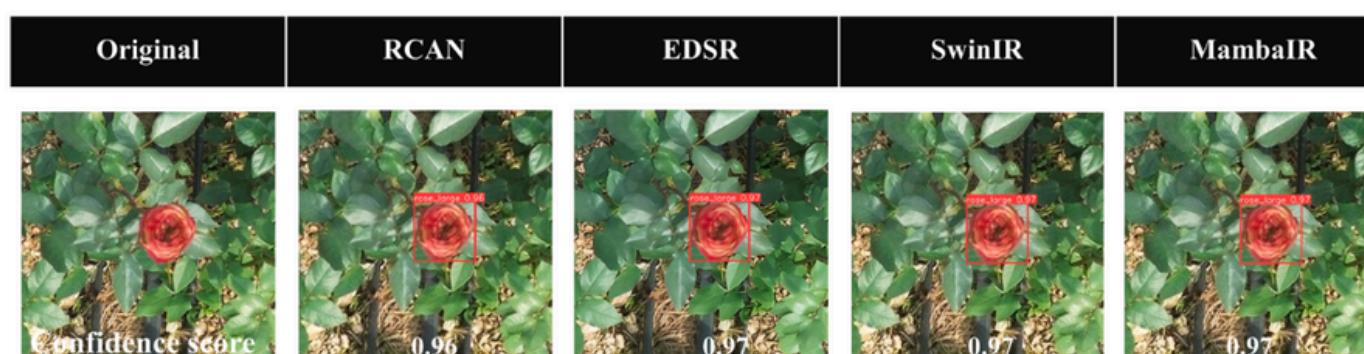
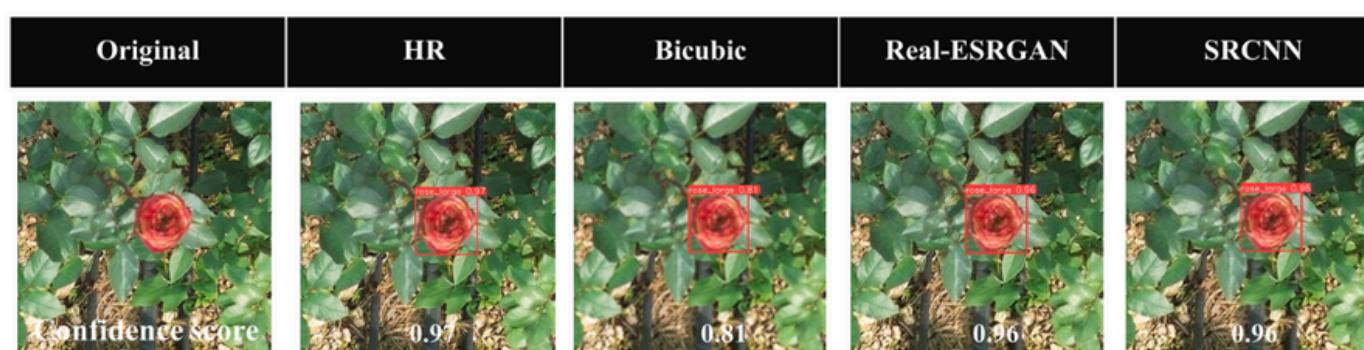
GSD 3.3cm

GSD 3.3 cm

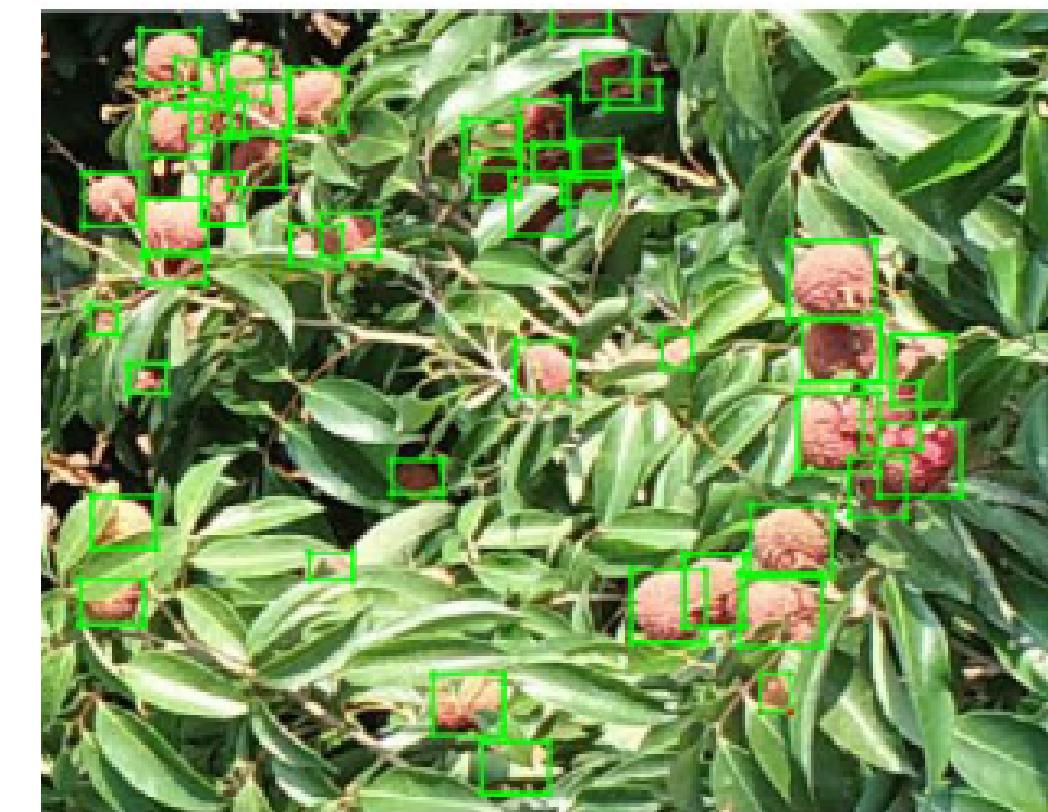
GSD 1.3 cm



Zhao et al., "Smart UAV-assisted rose growth monitoring with improved YOLOv10 and Mamba restoration techniques", 2024.



Liang et al., "Enhanced visual detection of litchi fruit in complex natural environments based on UAV remote sensing", 2023.



(b3) Labeled number: 391



(c3) Detection number: 46
(Before super-resolution)

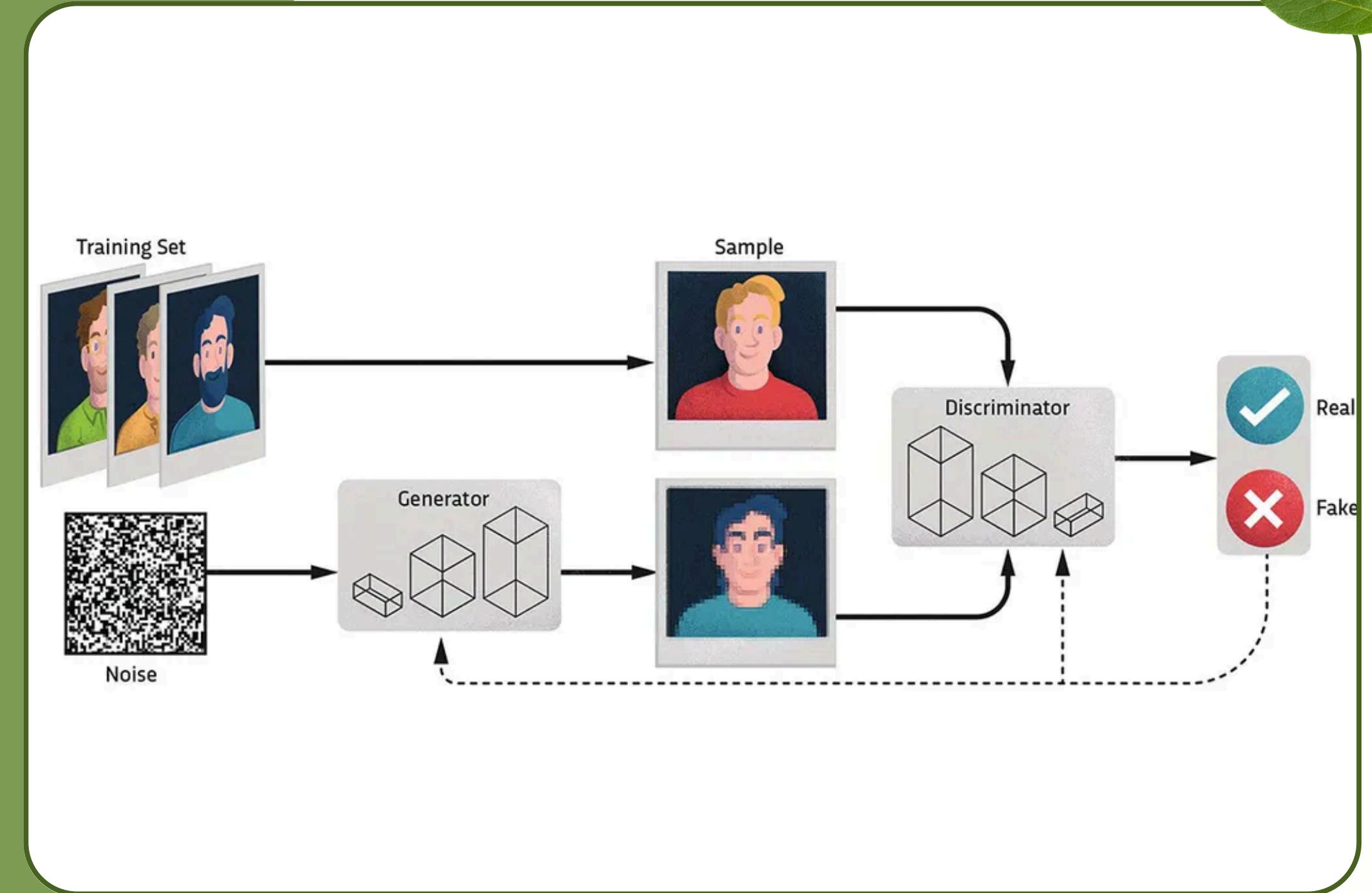
(d3) Detection number: 337
(After super-resolution)

GANs

GAN (Red generativa antagonica)

Entrena dos redes neuronales de modo que compitan entre sí para generar nuevos datos más auténticos a partir de un conjunto de datos de entrenamiento determinado.

Consta de dos redes neuronales profundas: la red generadora y la red discriminadora



SRGANs

Super-Resolution Generative Adversarial Networks

Redes generativas adversarias que están enfocadas a la super-resolución

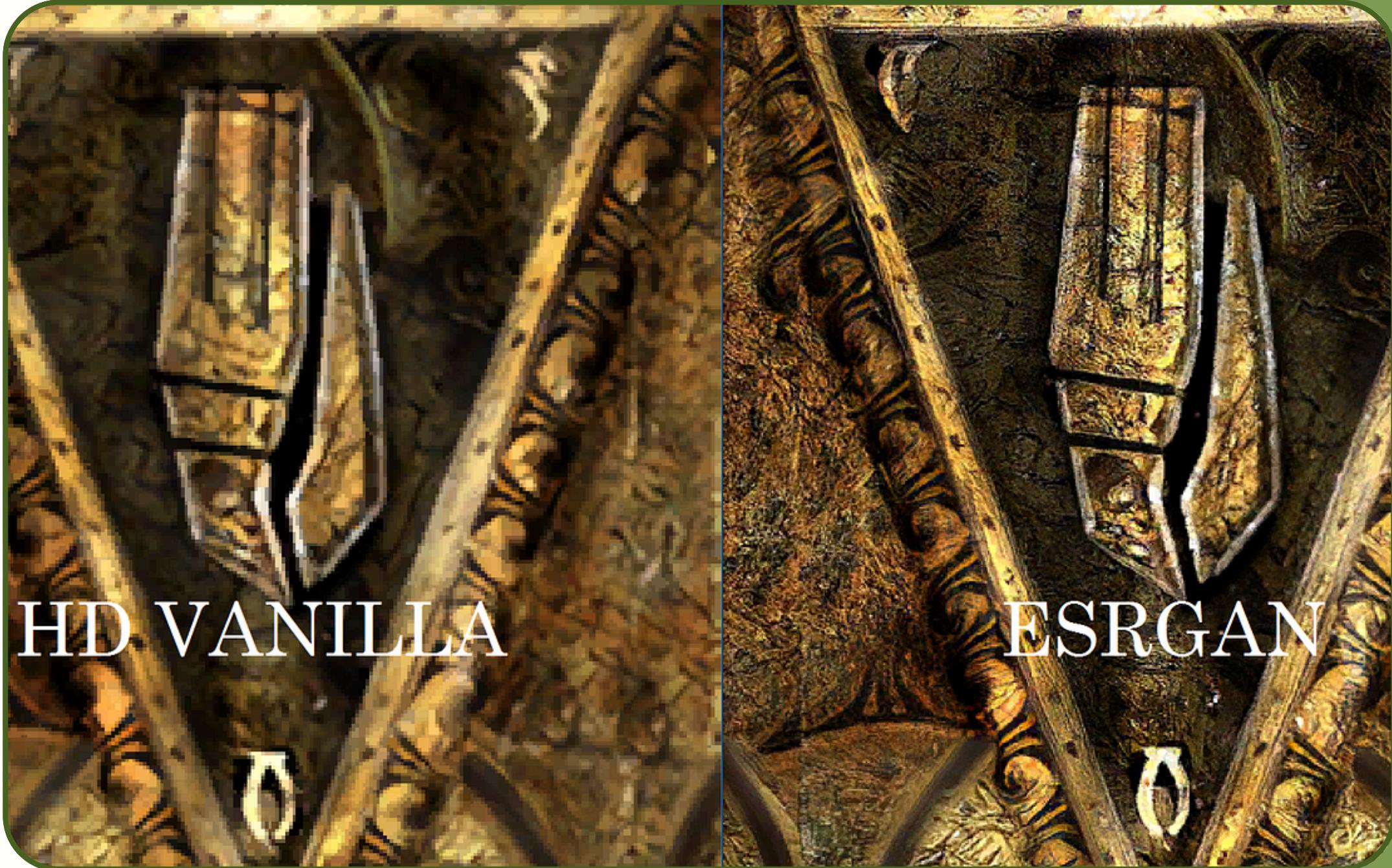
ESRGANs

Enhanced Super-Resolution Generative Adversarial Networks

Arquitectura similar a SRResGAN, pero con unas modificaciones para obtener mejores resultados

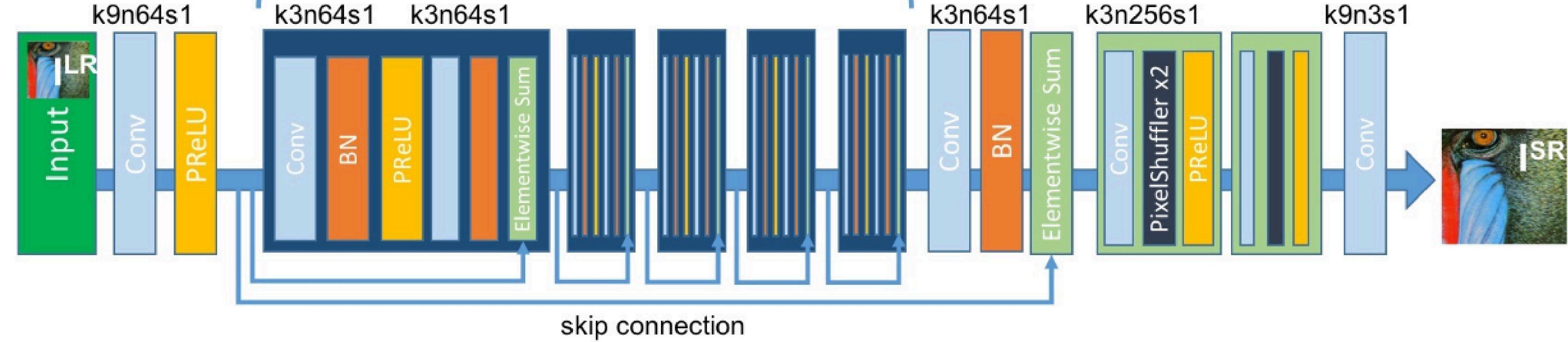
Real-ERSGAN

Genera degradaciones artificialmente muy similares a las que ocurren en la realidad para los datos de entrenamiento

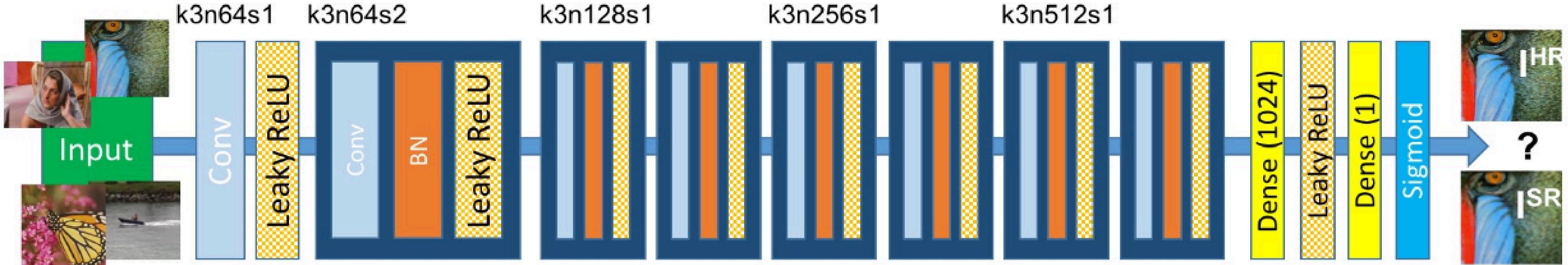


Generator Network

B residual blocks



Discriminator Network





METODOLOGÍA



Creación del dataset

Reconstrucción con Superresolución

Detección de Frailejones

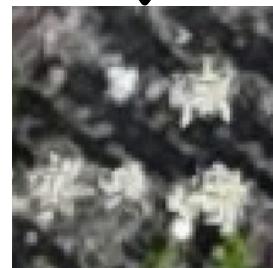
Toma de datos con dron



Datos de Baja resolución (GSD 3.3cm)

Datos de Alta resolución (GSD 1.13cm)

Generación de teselas de 8m x 8m



256 x 256



512 x 512

Disminución de resolución a diferentes escalas



Entrenamiento

Real-ERSGAN-Frailejones

256 x 256

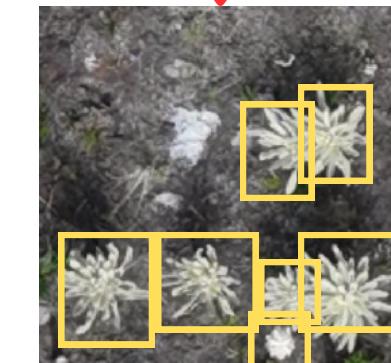


512 x 512

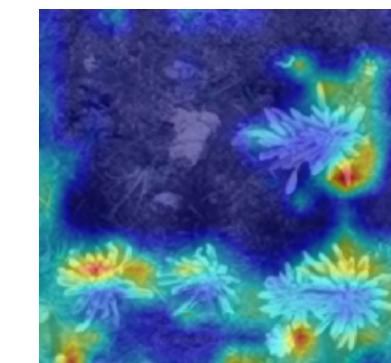
Entrenamiento

Yolov11

512 x 512
GSD 1.13cm



Gradcam





Depuración del Dataset



Las imágenes fueron degradadas a tres escalas diferentes para entrenar



0.75

0.5

0.33



RESULTADOS

#ecology

PSNR

Peak Signal-to-Noise Ratio

Es la proporción entre la energía máxima de una imagen y la energía de ruido que la afecta.

Para calcularla es necesario comparar la imagen ruidosa con una imagen sin ruido.

$$PSNR = 10 \log_{10}(R^2 / MSE)$$

Calculamos un PSNR promedio para el modelo base y el modelo entrenado, obteniendo los siguientes resultados:

Base: 21.59 dB

Entrenado: 21.05 dB

Original
(Ref.)



Base
PSNR: 19.48 dB



Trained
PSNR: 20.35 dB



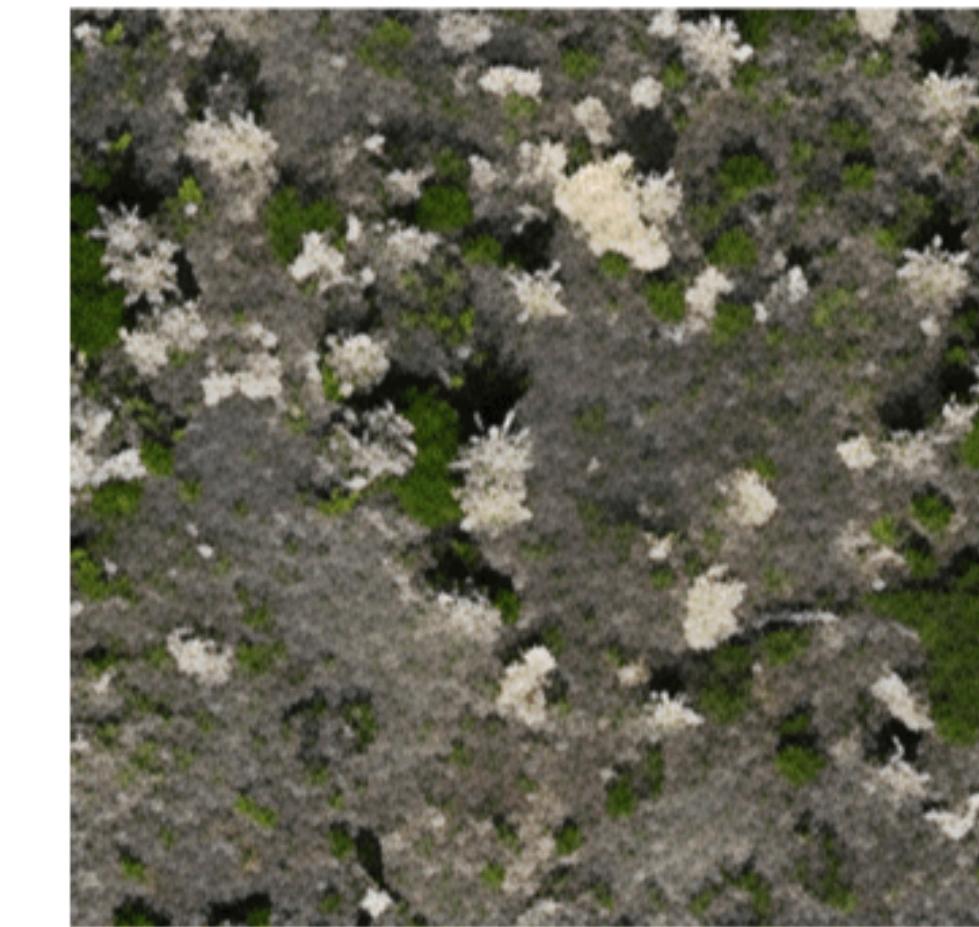
Original
(Ref.)



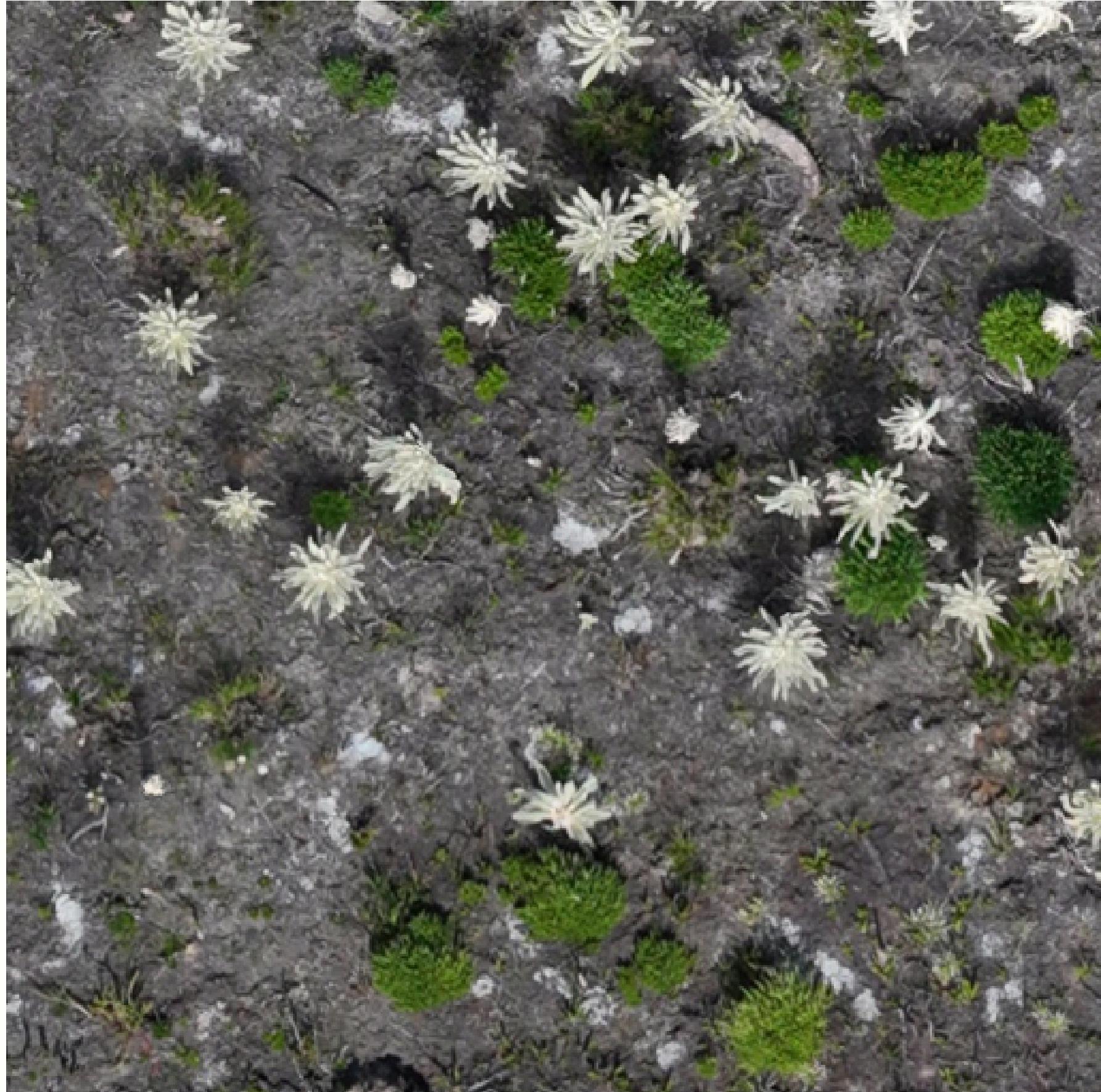
Base
PSNR: 19.90 dB



Trained
PSNR: 19.51 dB



GSD 1.3cm



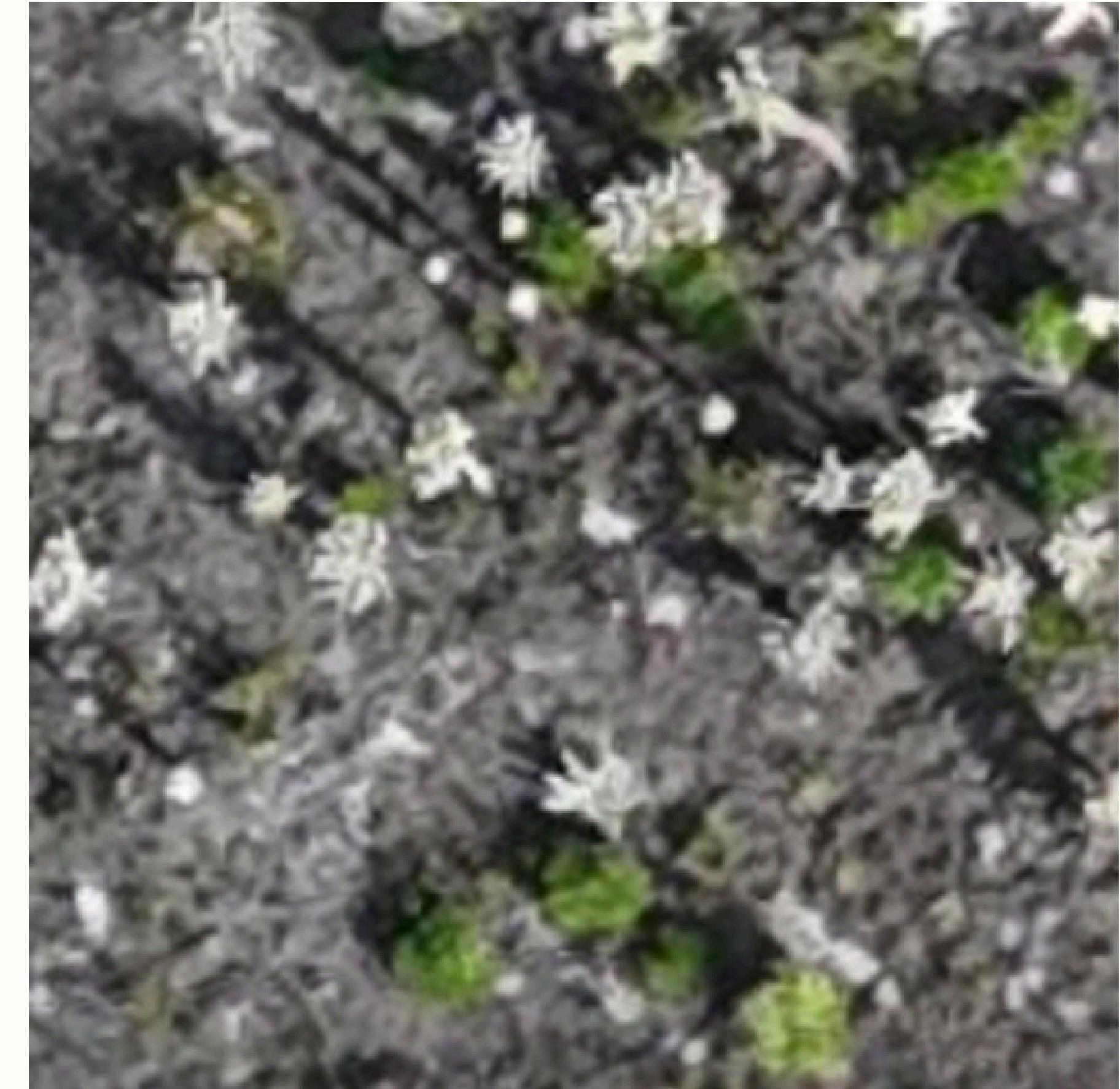
GSD 3.3cm



GSD 3.3cm



GSD 3.3cm Bicubic



GSD 1.3cm



GSD 3.3cm Real-ERSGAN



GSD 1.3cm



GSD 3.3cm Real-ERSGAN entrenado

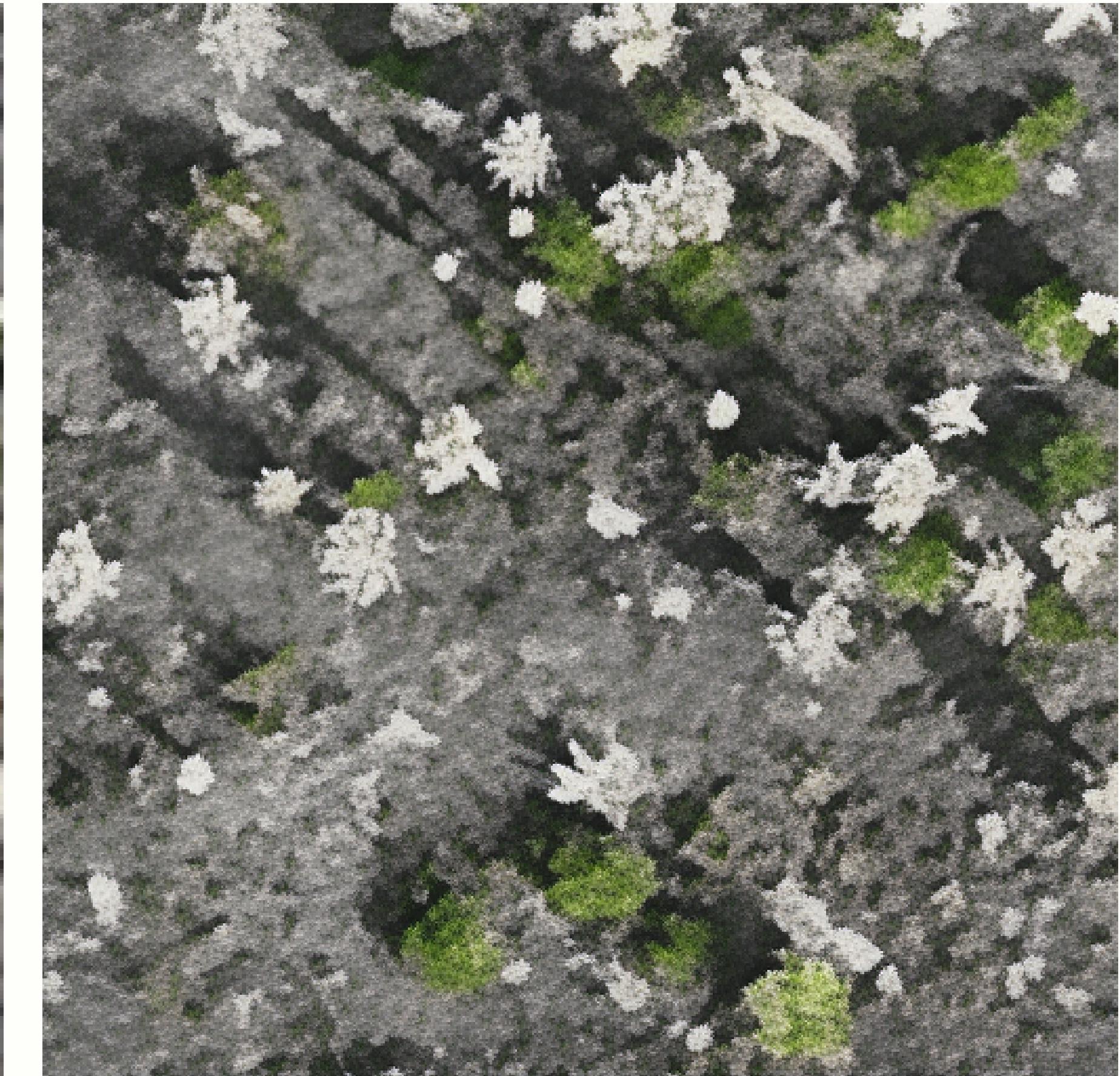
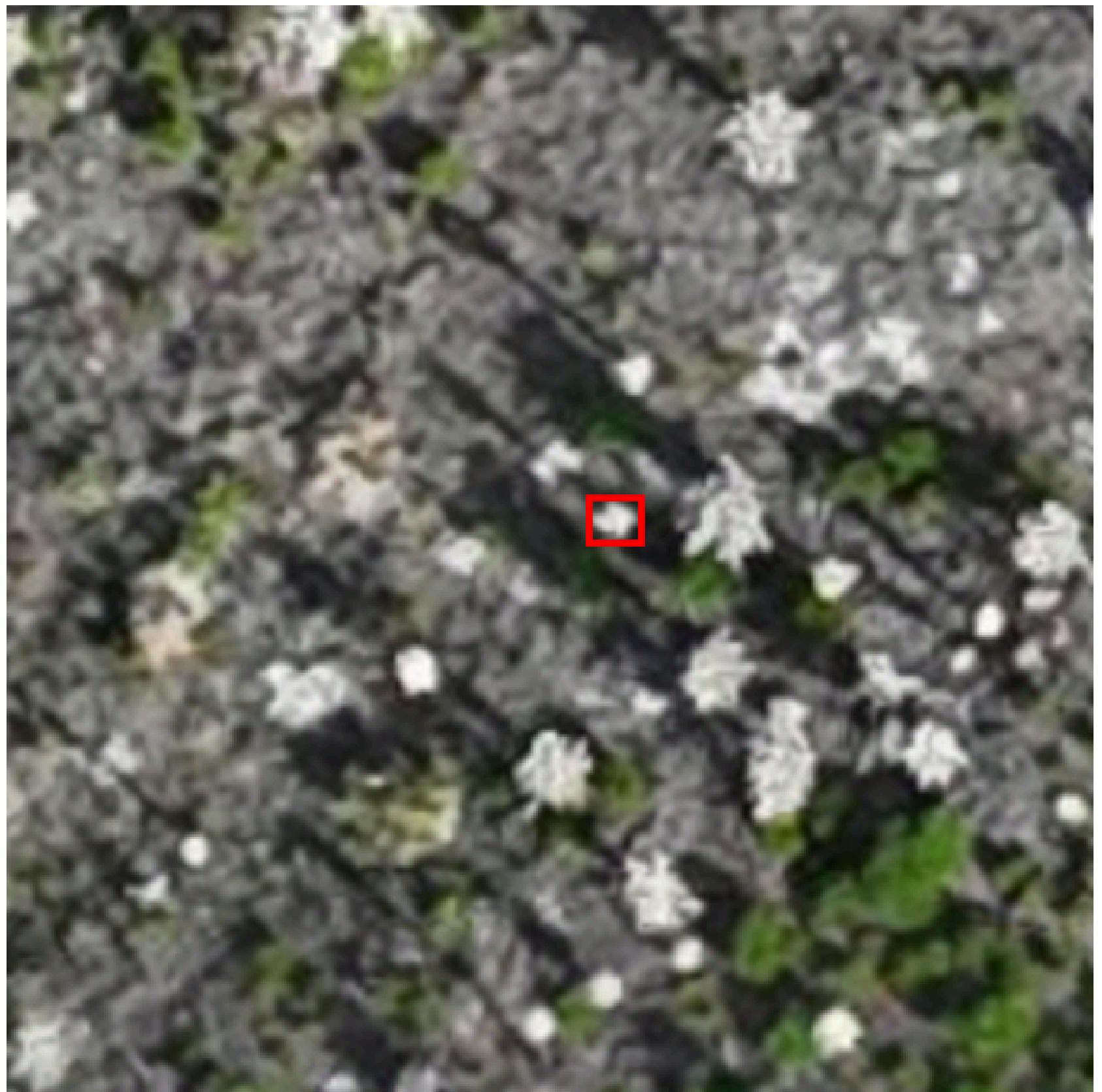
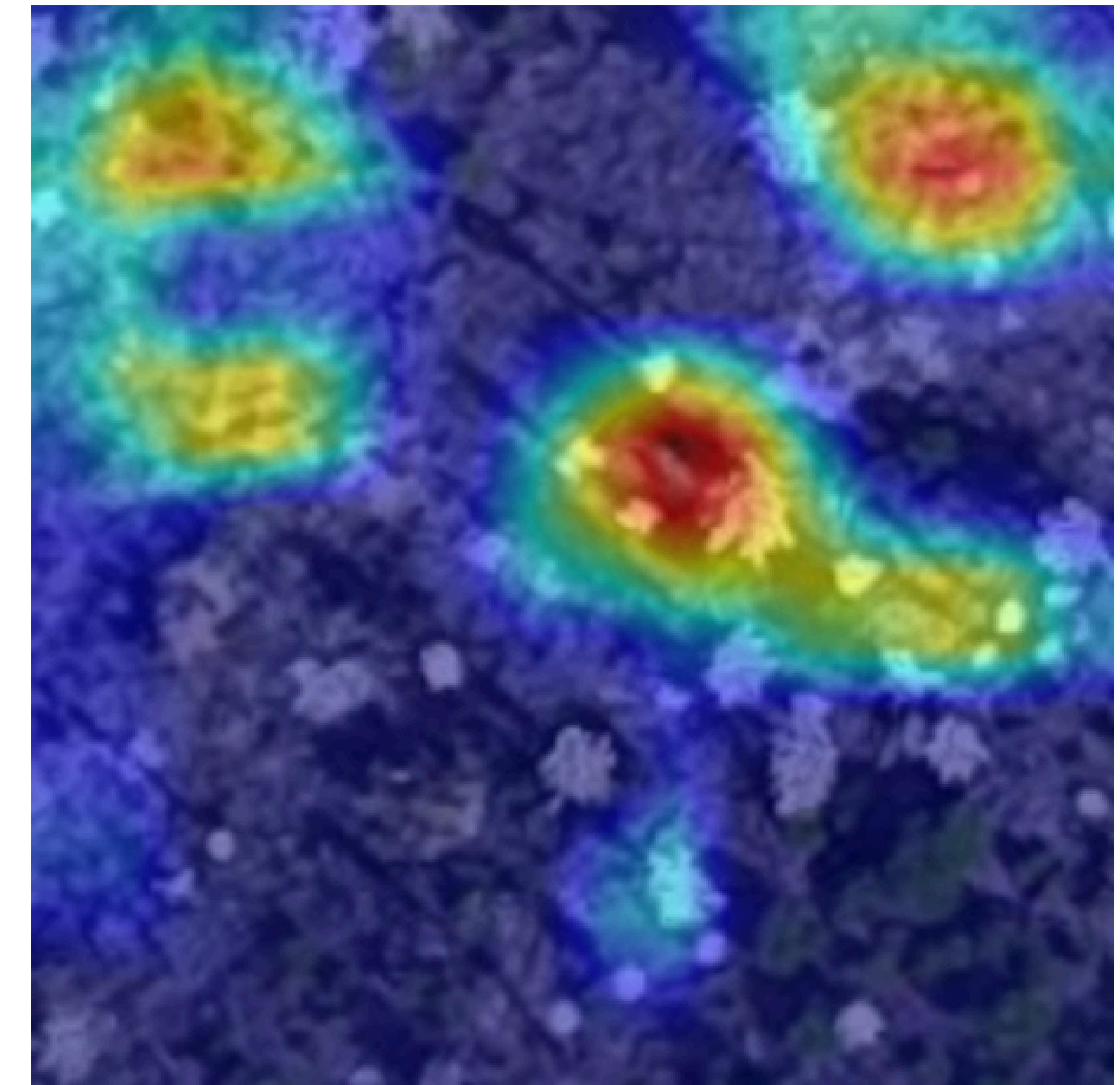


Imagen	<i>AP@0.5</i>
GSD 1.3 cm	0.877
GSD 3.3 cm	0.519
GSD 3.3 cm – Interpolación Bicúbica	0.526
GSD 3.3 cm – Real-ERSGAN	0.541
GSD 3.3 cm – Real-ERSGAN reentrenado	0.142

Interpolación Bicúbica



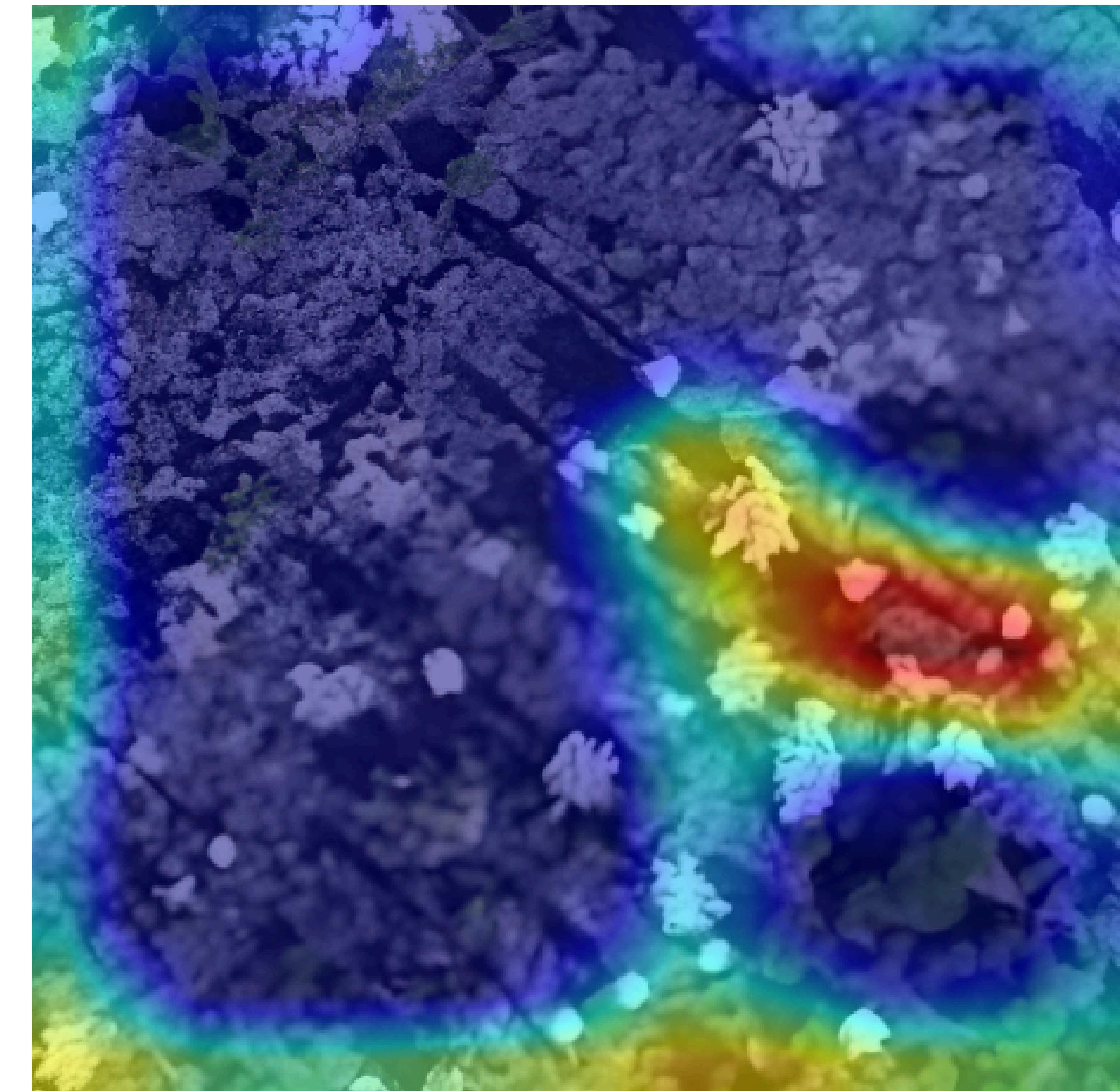
Grad-CAM



RealERSGAN



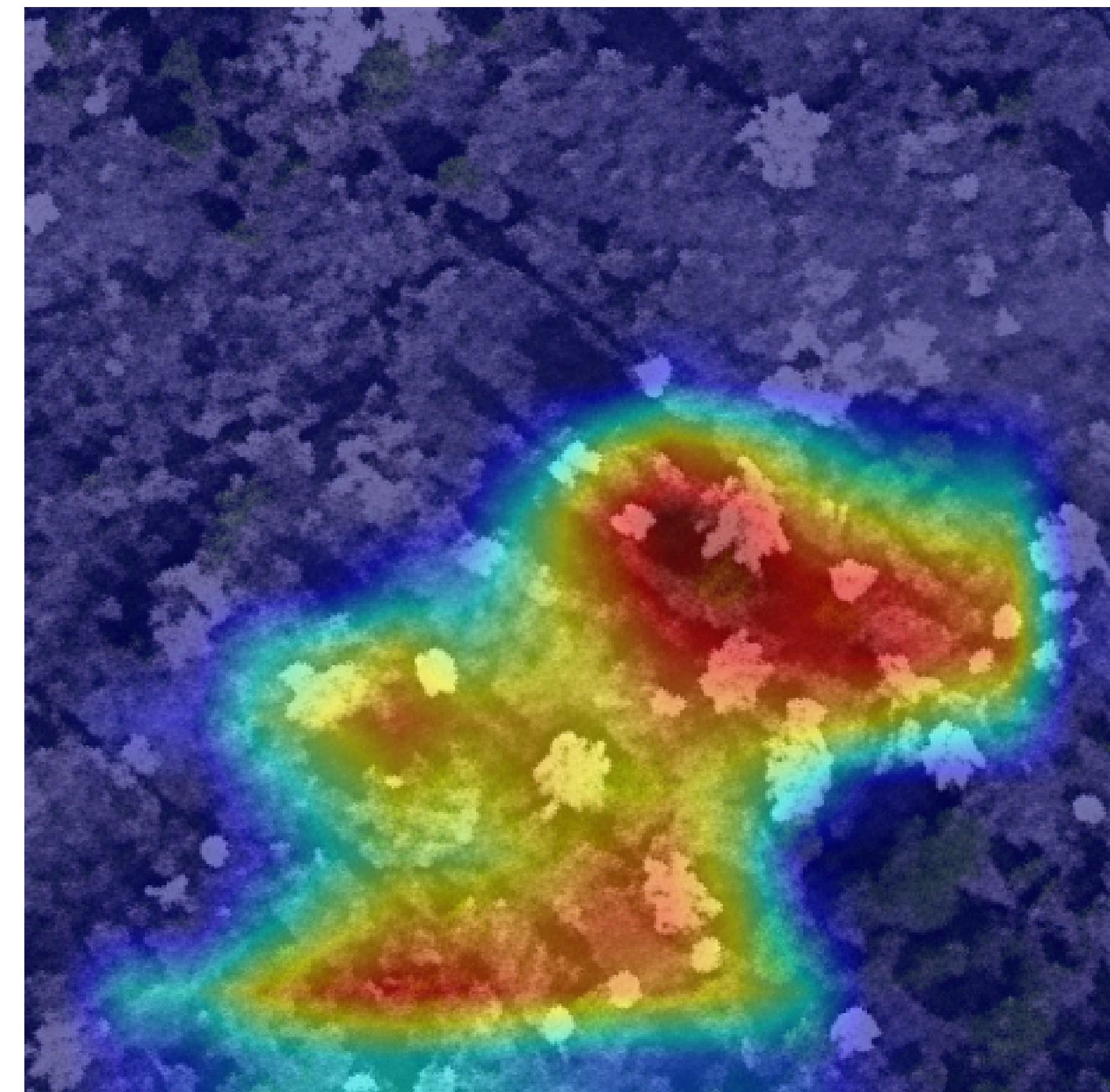
Grad-CAM



RealERSGAN Entrenada



Grad-CAM



Conclusiones



La superresolución con Real-ERSGAN mejora ligeramente la detección, pero el aumento es pequeño frente a métodos más simples.

Esto puede deberse a:

1. Diferencias entre el dominio de las imágenes y el de entrenamiento de Real-ERSGAN.
2. Pérdida de información irreparable en imágenes de 3.3 cm/píxel.

En conclusión, se deben explorar otras estrategias para mejorar la resolución.



MUCHAS
GRACIAS



REFERENCIAS

- [1] DJI. MAVIC 3 MULTISPECTRAL. DJI AGRICULTURE, 2025. RETRIEVED FEBRUARY 22, 2025. 1
- [2] CHRISTIAN LEDIG, LUCAS THEIS, FERENC HUSZ'AR, JOSE CABALLERO, ANDREW CUNNINGHAM, ALEJANDRO ACOSTA, ANDREW AITKEN, ALYKHAN TEJANI, JOHANNES TOTZ, ZEHAN WANG, ET AL. PHOTO-REALISTIC SINGLE IMAGE SUPER-RESOLUTION USING A GENERATIVE ADVERSARIAL NETWORK. IN PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, PAGES 4681-4690, 2017. 2, 3
- [3] CHANGJIANG LIANG, JUNTAO LIANG, WEIGUANG YANG, WEIYI GE, JING ZHAO, ZHAORONG LI, SHUDAI BAI, JIAWEN FAN, YUBIN LAN, AND YONGBING LONG. ENHANCED VISUAL DETECTION OF LITCHI FRUIT IN COMPLEX NATURAL ENVIRONMENTS BASED ON UNMANNED AERIAL VEHICLE (UAV) REMOTE SENSING. PRECISION AGRICULTURE, 26(1):23, 2025. 1
- [4] XINTAO WANG, LIANGBIN XIE, CHAO DONG, AND YING SHAN. REAL-ESRGAN: TRAINING REAL-WORLD BLIND SUPER-RESOLUTION WITH PURE SYNTHETIC DATA. IN INTERNATIONAL CONFERENCE ON COMPUTER VISION WORKSHOPS (ICCVW). 1
- [5] XINTAO WANG, KE YU, SHIXIANG WU, JINJIN GU, YIHAO LIU, CHAO DONG, YU QIAO, AND CHEN CHANGE LOY. ESRGAN: ENHANCED SUPER-RESOLUTION GENERATIVE ADVERSARIAL NETWORKS. IN PROCEEDINGS OF THE EUROPEAN CONFERENCE ON COMPUTER VISION (ECCV) WORKSHOPS, PAGES 0-0, 2018. 2
- [6] FAN ZHAO, ZHIYAN REN, JIAQI WANG, QINGYANG WU, DIANHAN XI, XINLEI SHAO, YONGYING LIU, YIJIA CHEN, AND KATSUNORI MIZUNO. SMART UAV-ASSISTED ROSE GROWTH MONITORING WITH IMPROVED YOLOV10 AND MAMBA RESTORATION TECHNIQUES. SMART AGRICULTURAL TECHNOLOGY, 10:100730, 2025.