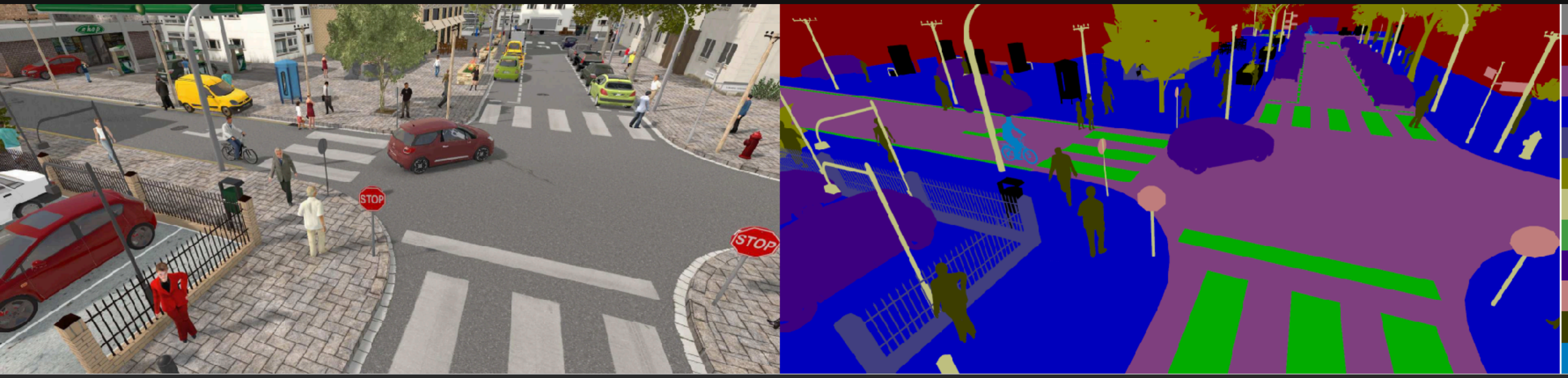
A Walk Through a Digital Savanna

Aerial Wildlife Detection with Synthetic Data

Benjamin Kellenberger, Derek van de Ven, Devis Tuia

EPFL

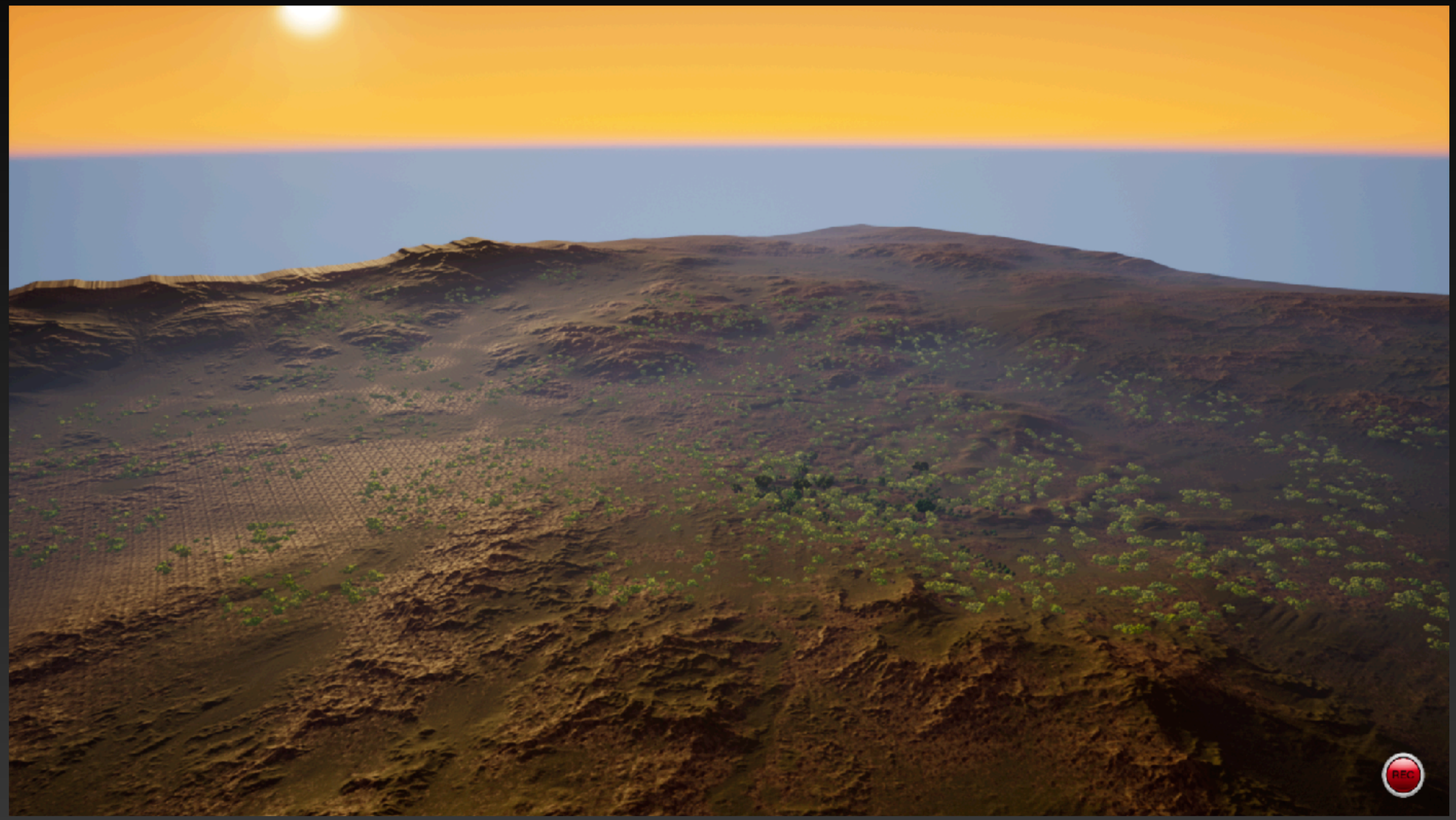
Let's use Synthetic Data!



Ros et al. The SYNTHIA Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes. CVPR, 2016.

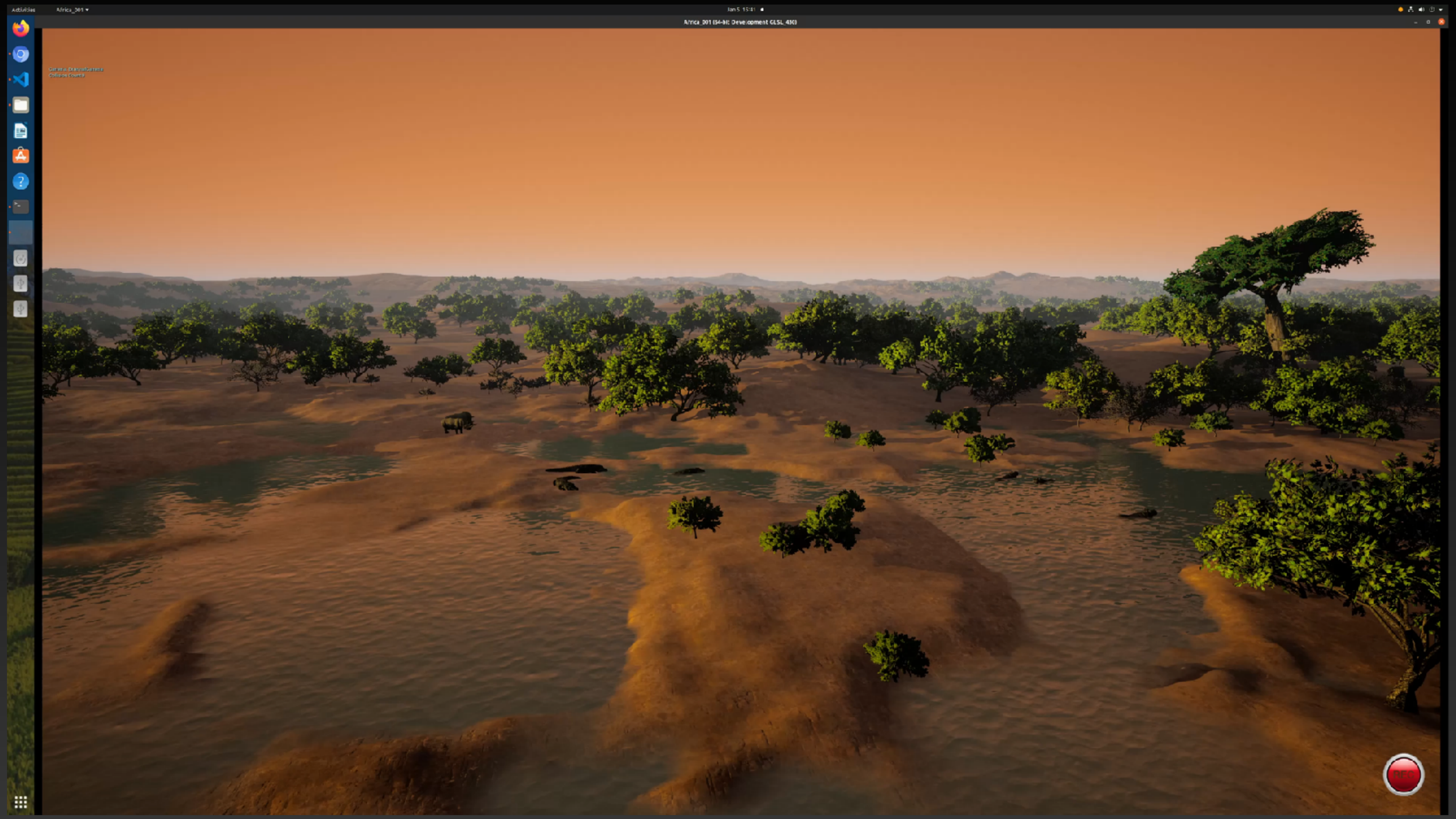
EPFL

AirSim-W



Bondi et al. AirSim-W: A Simulation Environment for Wildlife Conservation with UAVs. ACM SIGCAS, 2018.

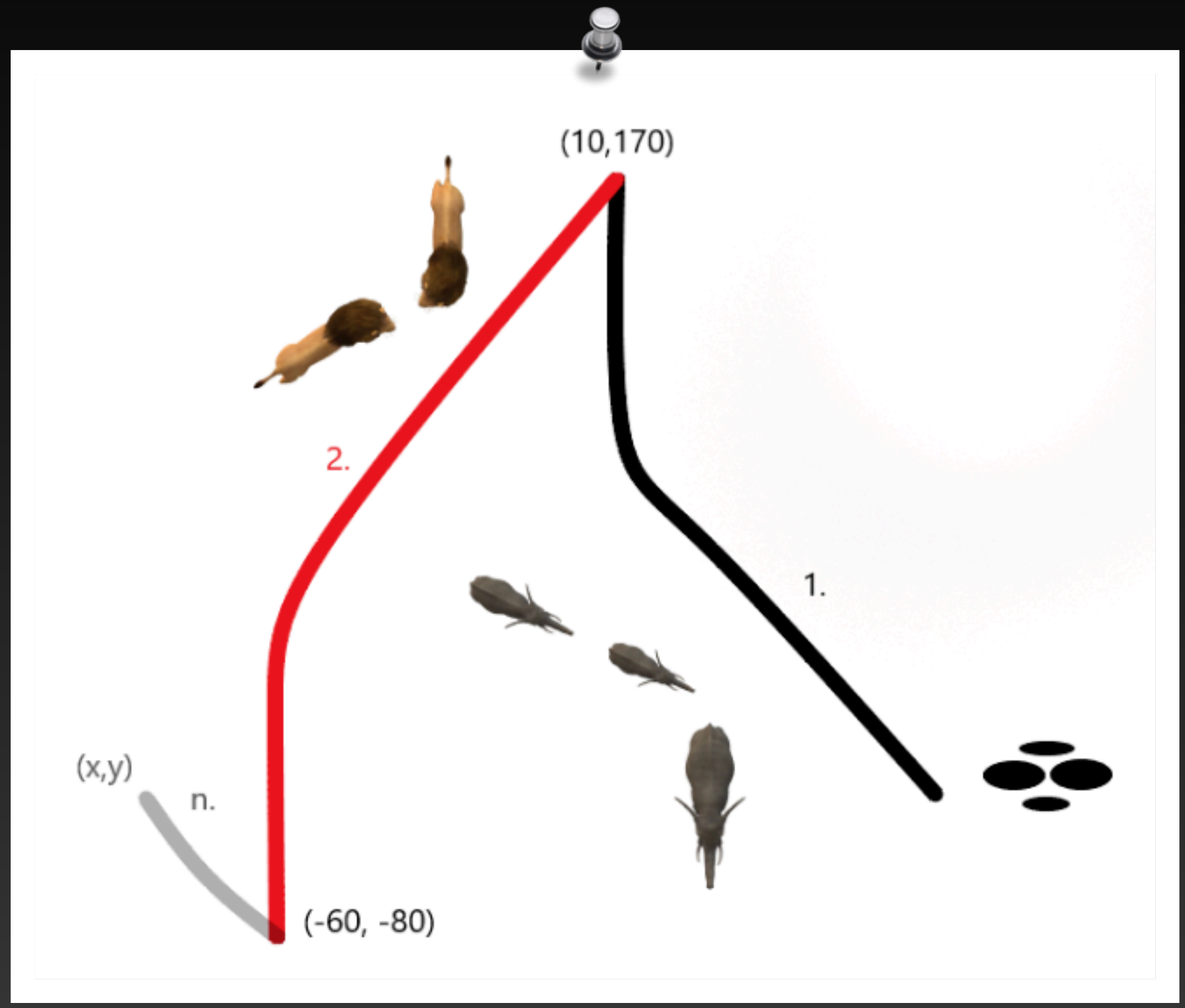
EPFL

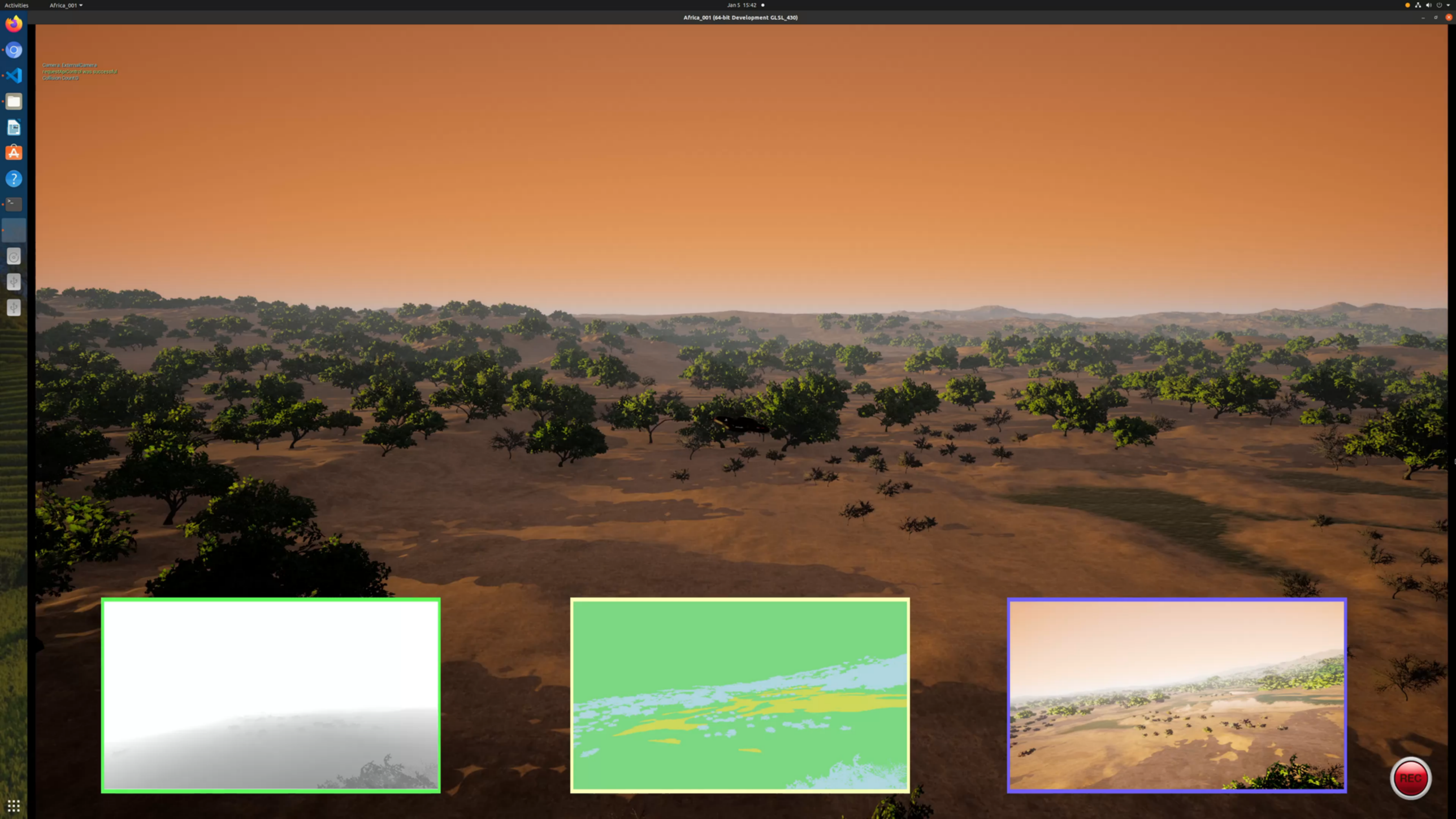


Bondi et al. AirSim-W: A Simulation Environment for Wildlife Conservation with UAVs. ACM SIGCAS, 2018.

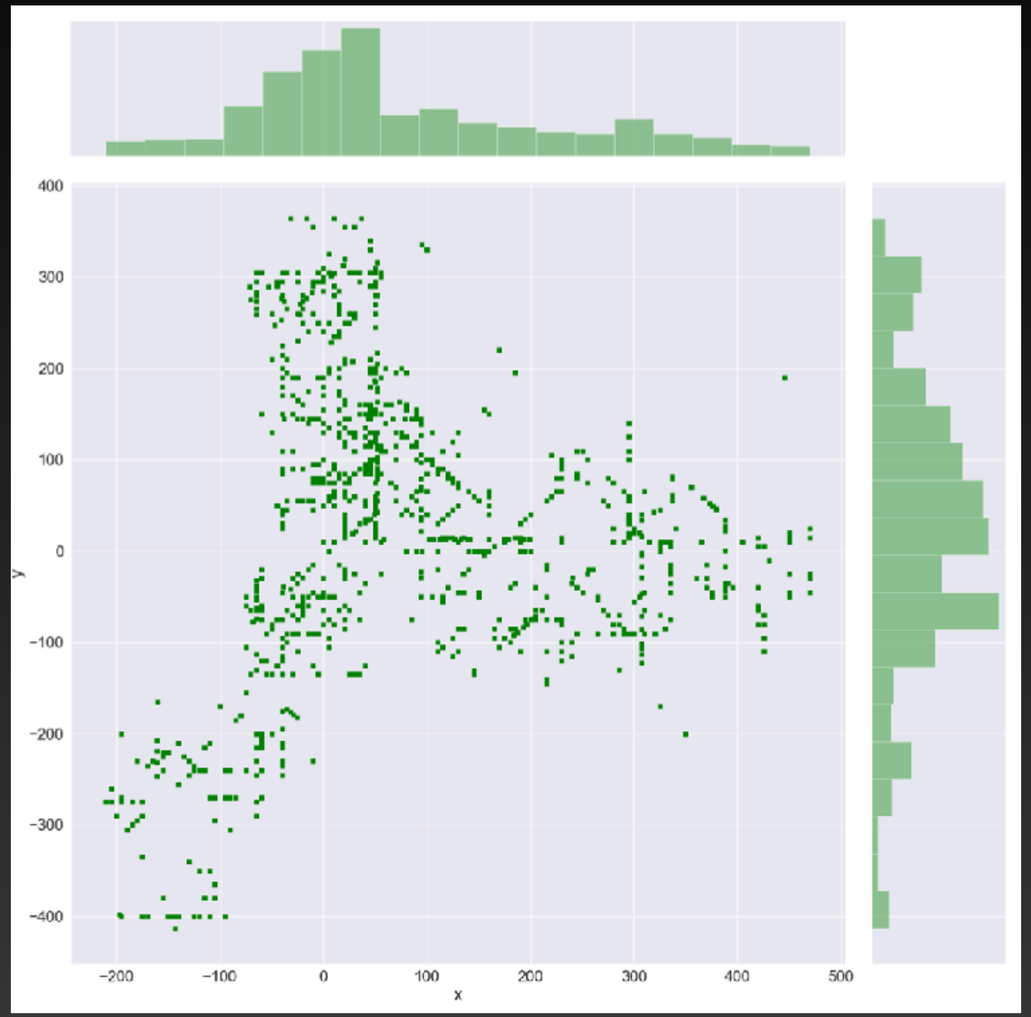
Data Acquisition Schedule

- Nadir images within 1000×1000m around centre
- 20-60m flying altitude
- Flying pattern:
    1. Start from centre
    2. Select random point within bounds
    3. Fly towards it diagonally; take picture every 3-5m

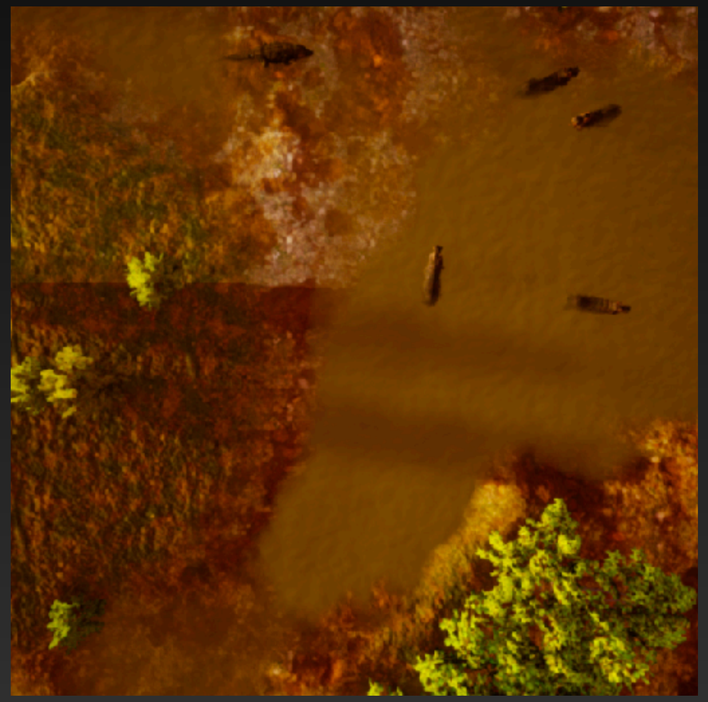



Activities
Africa_001
Jan 5 15:42
Africa_001 (64-bit Development GLSL_430)
Camera ExternalCamera
requestAppControl was successful
Collision Count:0
REC

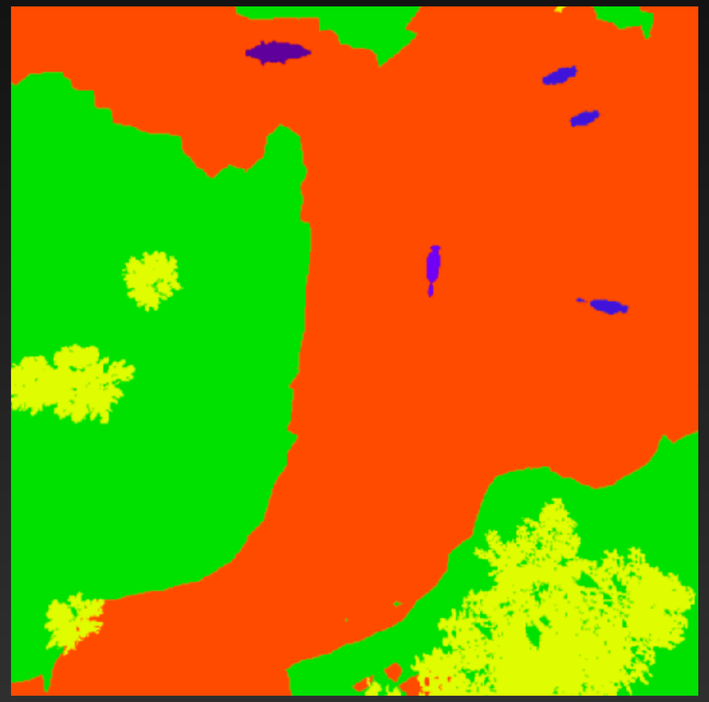
Data Acquisition Schedule



Data Acquisition

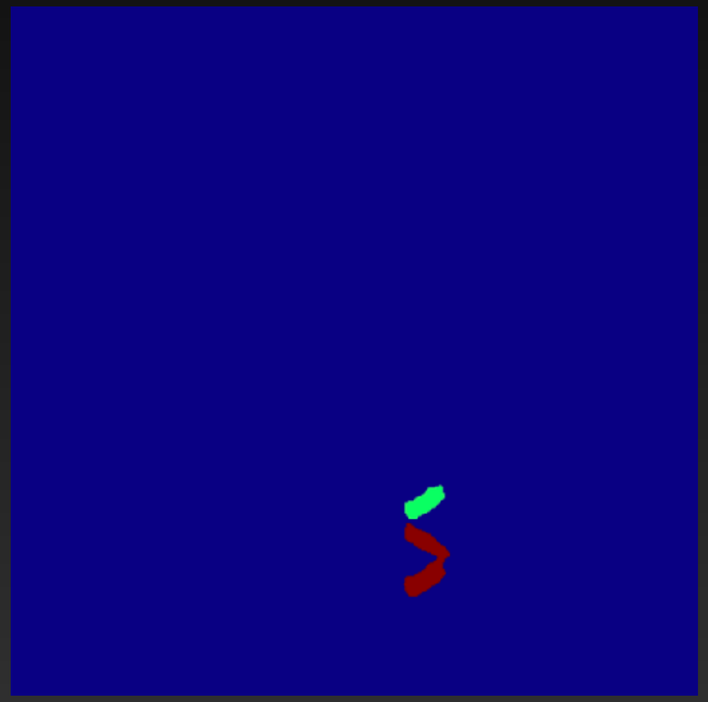


rendered image

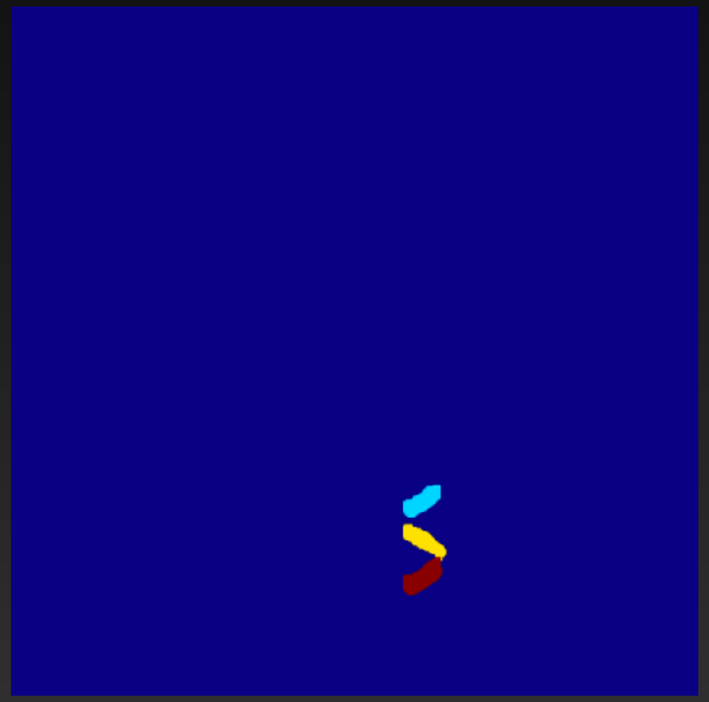


segmentation mask

Data Acquisition

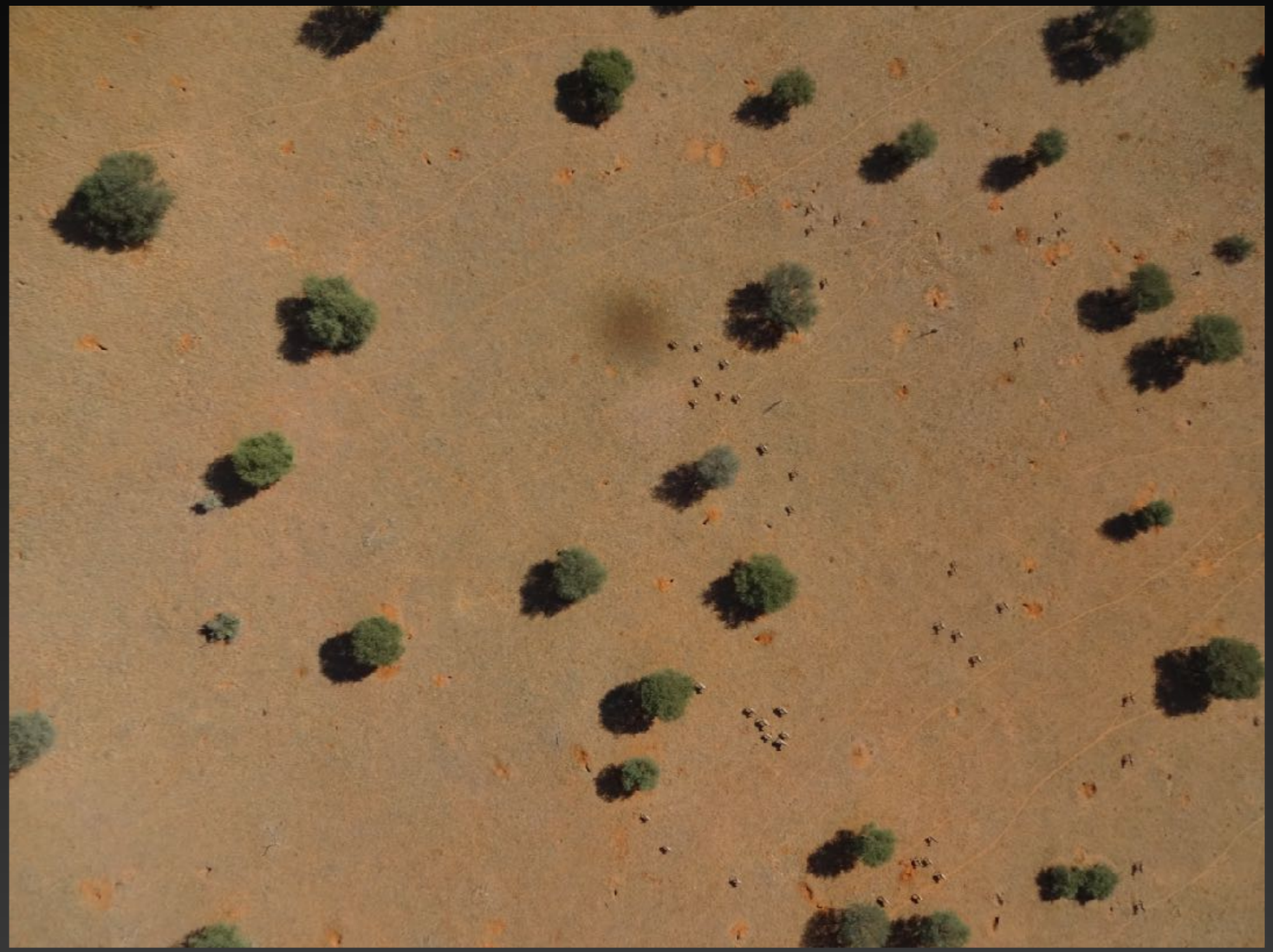


original segmentation mask



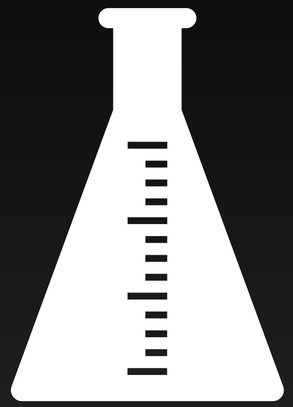
after erosion + dilation

Real Images

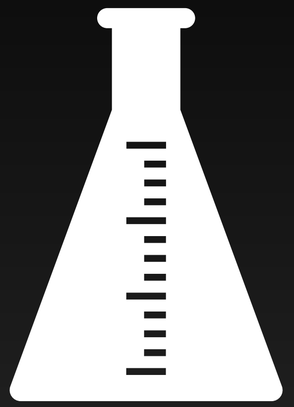


https://kuzikus-namibia.de

Training Data



AirSim-W
(synthetic)



Kuzikus
(real)

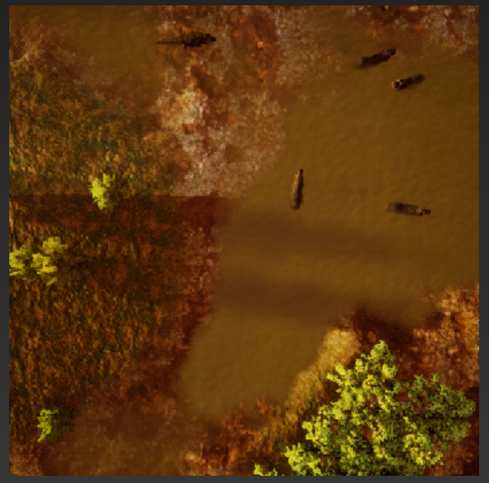
images

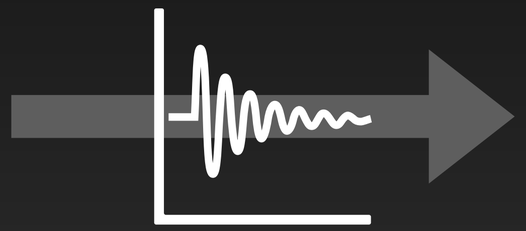
AirSim-W	Kuzikus
5000	0
5000	1000
0	8000

EPFL

Model

RetinaNet with ResNet-18





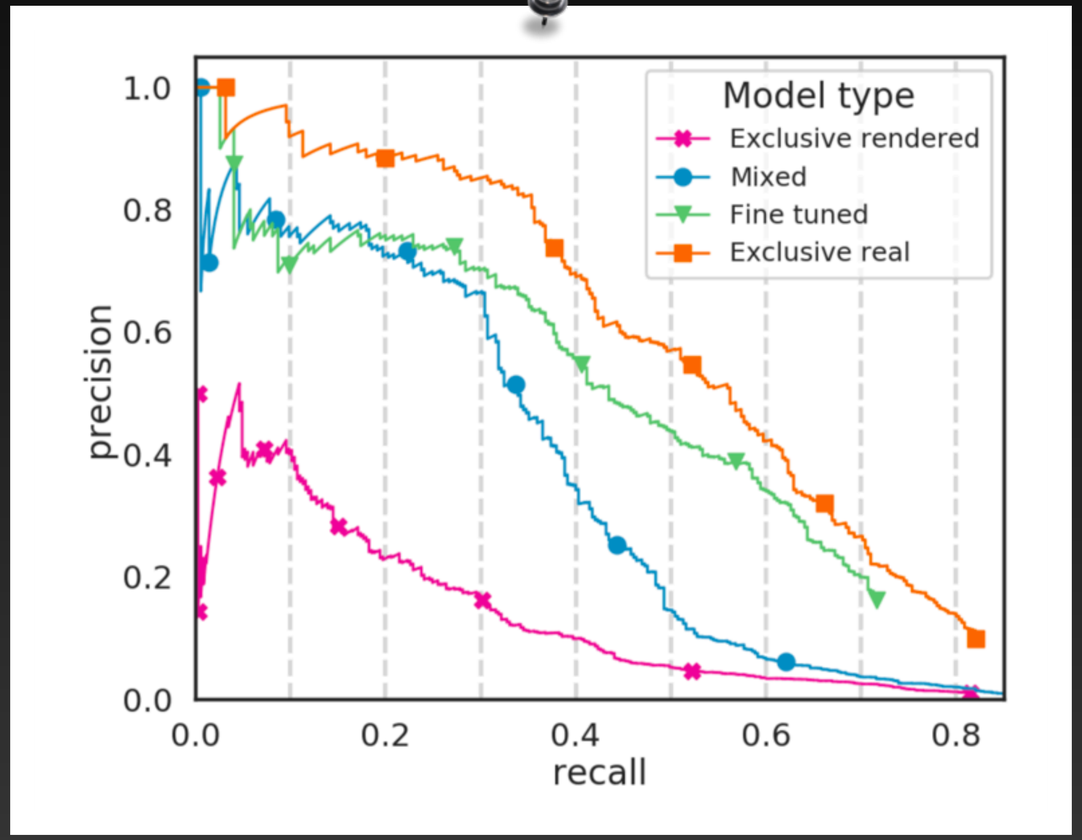
random horizontal flips
Gaussian blurring

non-maximum suppression

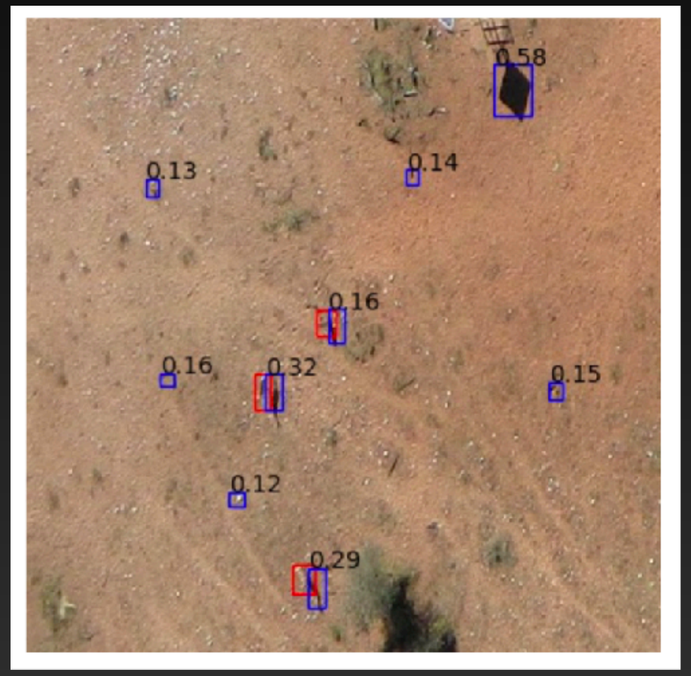
Lin et al. Focal loss for dense object detection. ICCV, 2017.

EPFL

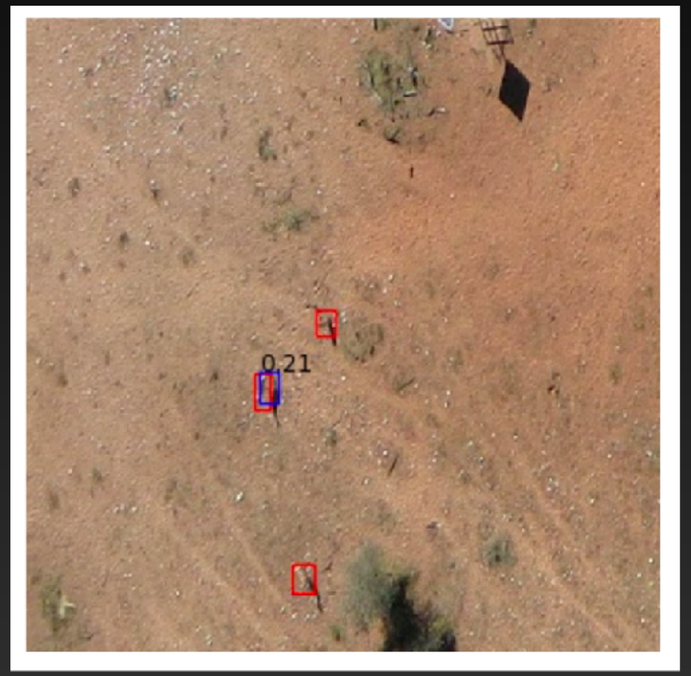
Results


Model type
Exclusive rendered
Mixed
Fine tuned
Exclusive real
precision
recall
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8

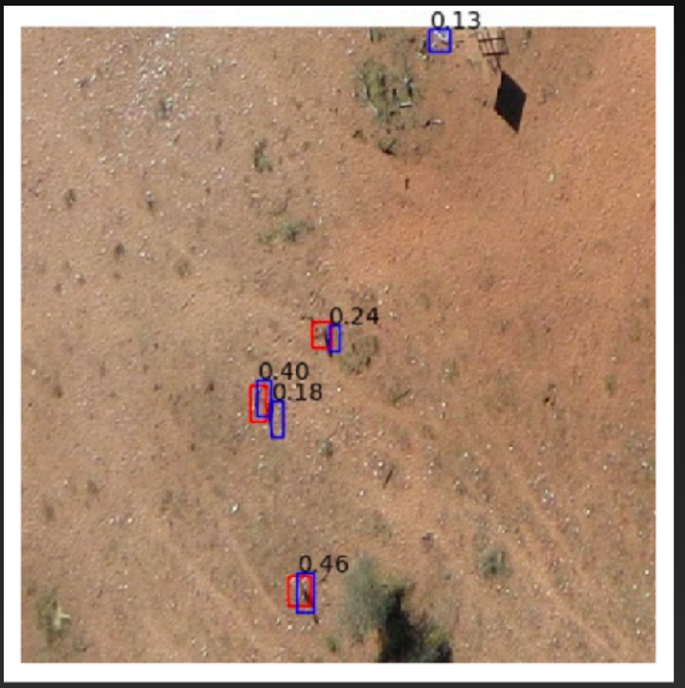
Results



exclusively rendered



fine-tuned



exclusively real

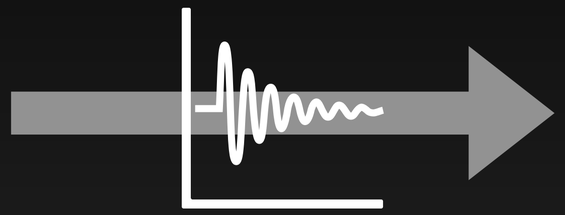
prediction *ground truth*

EPFL

Domain Adaptation? GAN?

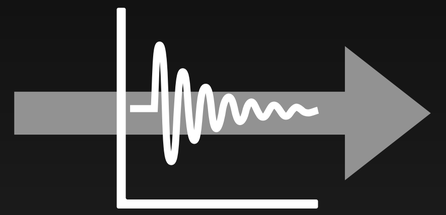
The GAN Story
CycleGAN

real A



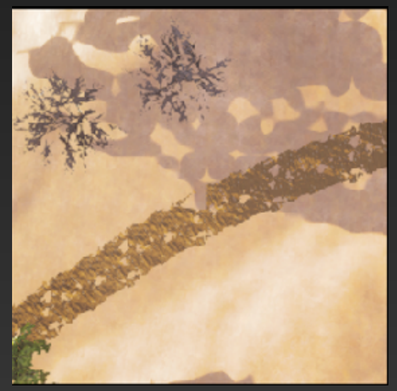
Generator (A → B)

fake B

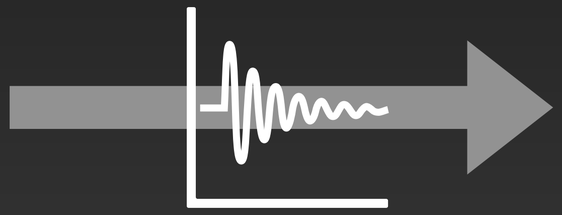


Generator (B → A)

recovered A

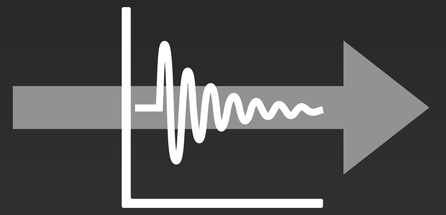


real B

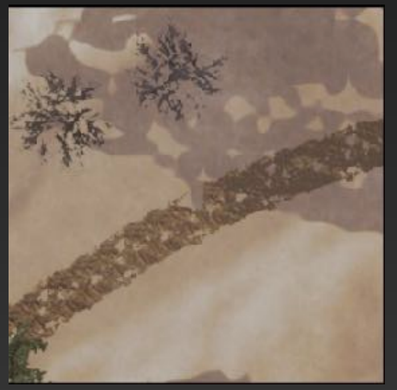


Generator (B → A)

fake A



Generator (A → B)

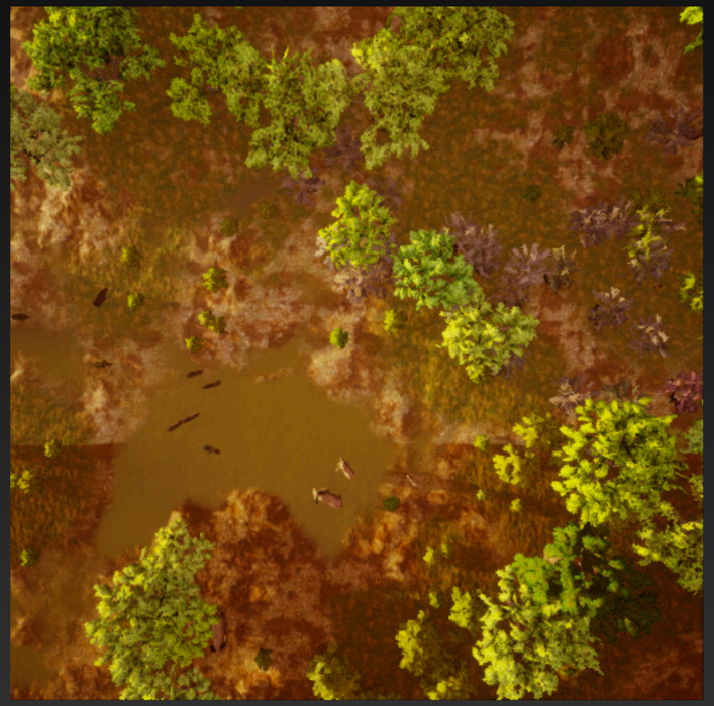


recovered B

Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV, 2017.

EPFL

The GAN Story
CycleGAN



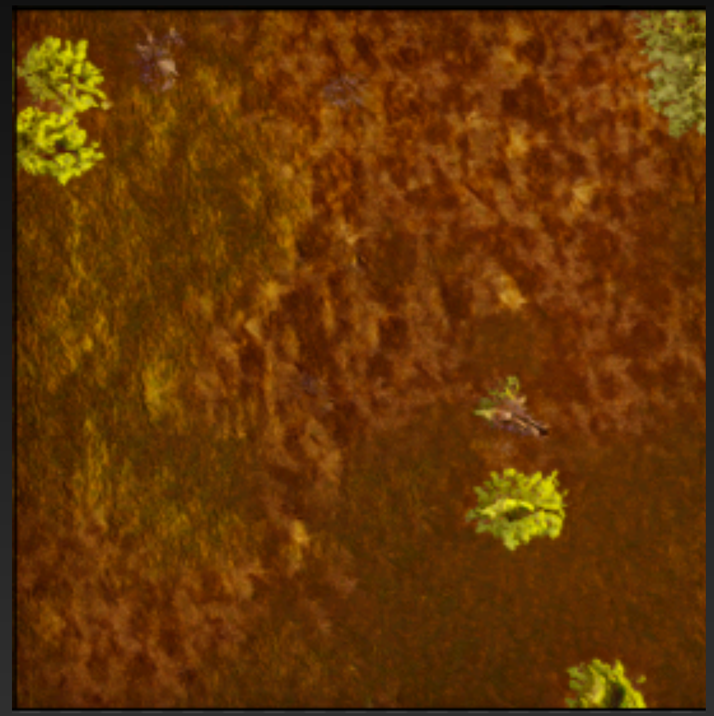
AirSim-W

fake Kuzikus

Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV, 2017.

The GAN Story
CycleGAN

Kuzikus



fake AirSim-W

Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV, 2017.

The GAN Story
CycleGAN

ill-posed

Model doesn’t know what an animal should look like.

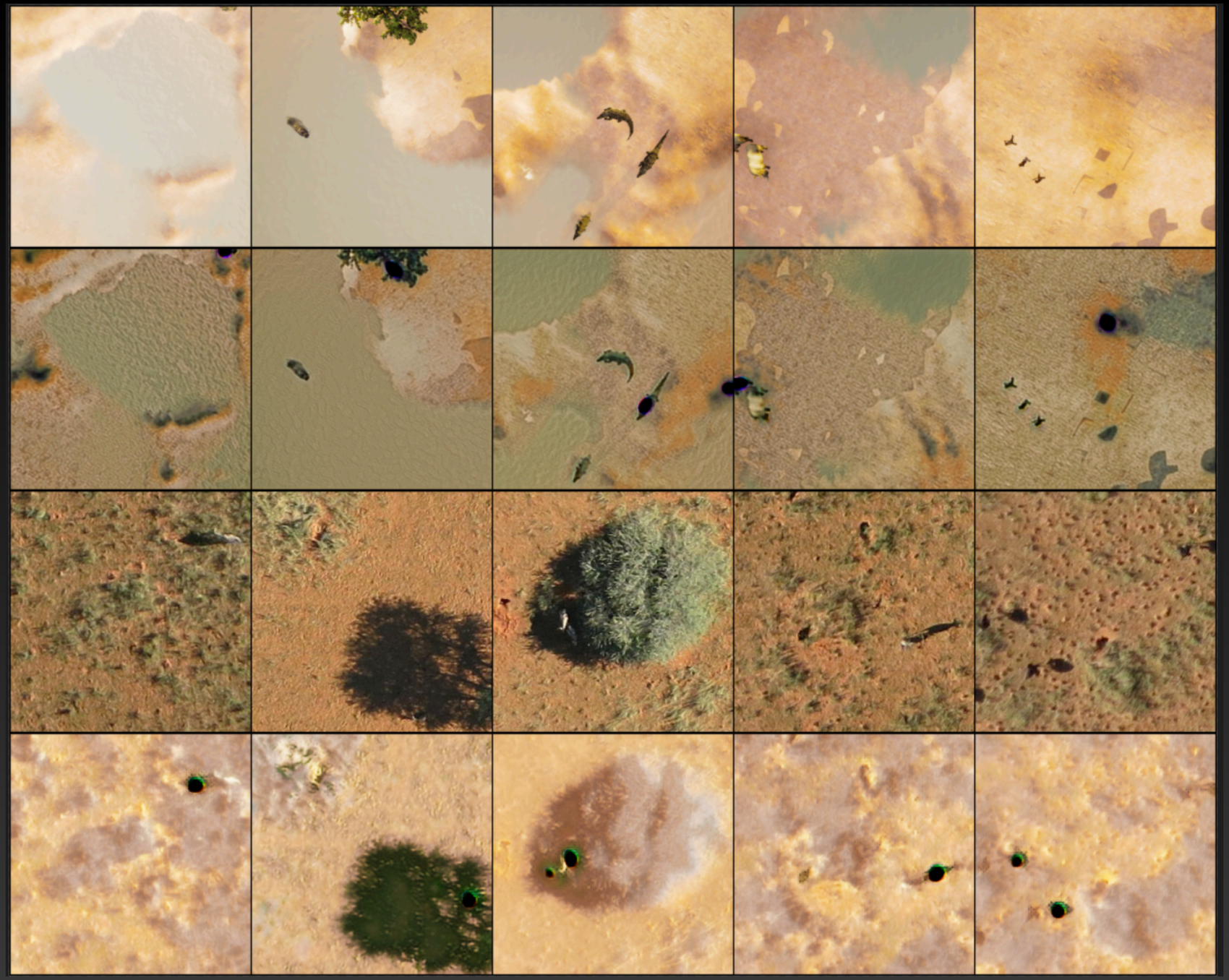
Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV, 2017.

EPFL

The GAN Story
CycleGAN

- Sobel filter L1 norm on images
- Auxiliary segmentation task (U-Net) on AirSim-W ground truth
- Auxiliary detection task (RetinaNet) on fake output
- L1 norm between AirSim-W real & fake animal targets
- Generator architectures (ResNet, U-Net)
- Discriminator architectures (ResNet, U-Net)

Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV, 2017.



Benjamin Kellenberger, Derek van de Ven, Devis Tuia

EPFL