AI4AS | *Monday June 29th*

SYNFISH

Testing synthetic data integration for automated fish detection on beam trawl vessels

Hanne De Rijcke, *Lukas Snoeck, Siebren Cosijn, Sander Delacauw*

ILVO, Belgium

Co-funded by the European Union

Europees Fonds voor Maritieme Zaken en Visserij

ILVO vlaanderen

Why synthetic data for fisheries monitoring?

- Increasing importance **of automated catch monitoring** for **scalable and reliable fisheries data collection**
- Need for **diverse, well-labelled data**
- **Real data imagery is challenging** in trawl fisheries:
  - Mixed catches
  - Overlapping individuals
  - Variable lighting
  - Rare or protected species
  - High annotation effort

Synthetic data could help **fill these gaps** - if it **transfers reliably to real imagery**

Medegefinancierd door de Europese Unie

From synthetic data promise to workflow sensitivity

Main question:

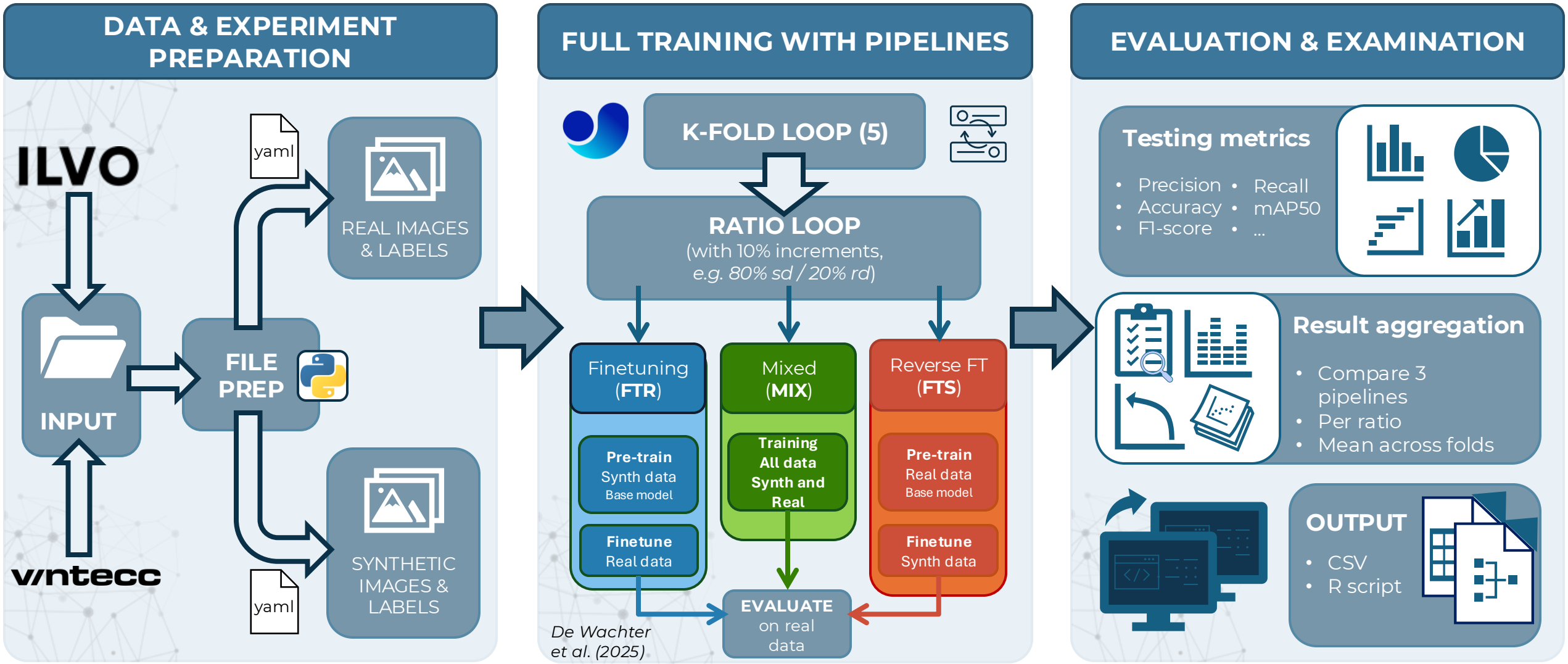
Can **synthetic fish** imagery **improve detection models** for real fisheries monitoring, and which **workflow choices** determine whether this **improvement is reliable**?

We evaluate sensitivity to:

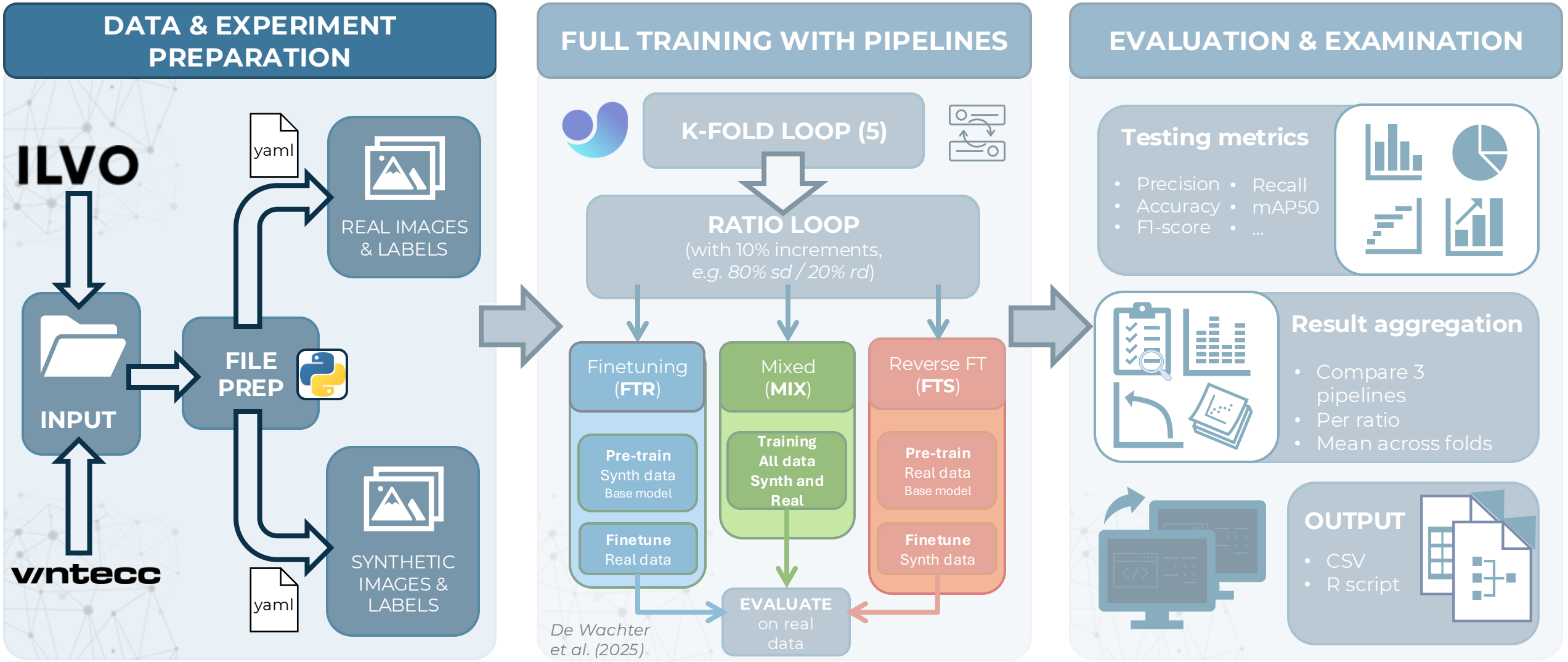
1. Synthetic–real training strategy
2. Synthetic–real data ratio
3. Annotation and preprocessing workflow
4. Biological and visual diversity

Medegefinancierd door de Europese Unie

Project workflow
detection & segmentation model training



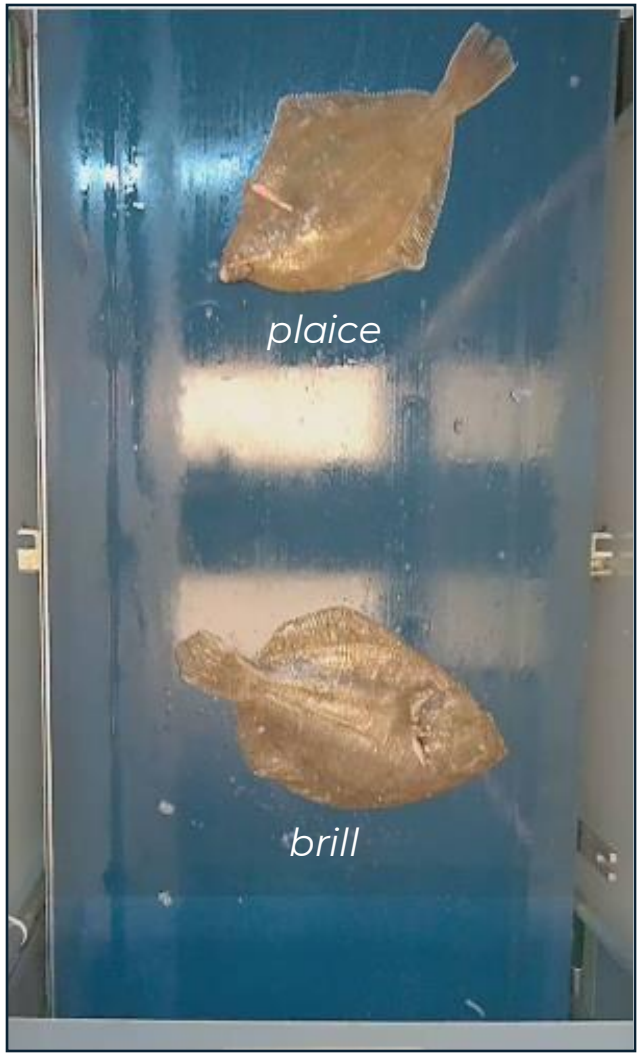
Project workflow
detection & segmentation model training



Project workflow
Real data

Controlled lab scenario's

- Two **species**: brill or plaice



Medegefinancierd door de Europese Unie

Project workflow
Real data

Controlled lab scenario's

- Two **species**: brill or plaice
- **Orientation**:
  - belly, back, or mixed

back *mixed* *belly*

Medegefinancierd door de Europese Unie

Project workflow
Real data

Controlled lab scenario's

- Two **species**: brill or plaice
- **Orientation**:
  - belly, back, or mixed
- **Overlap**:
  - none, moderate, or heavy

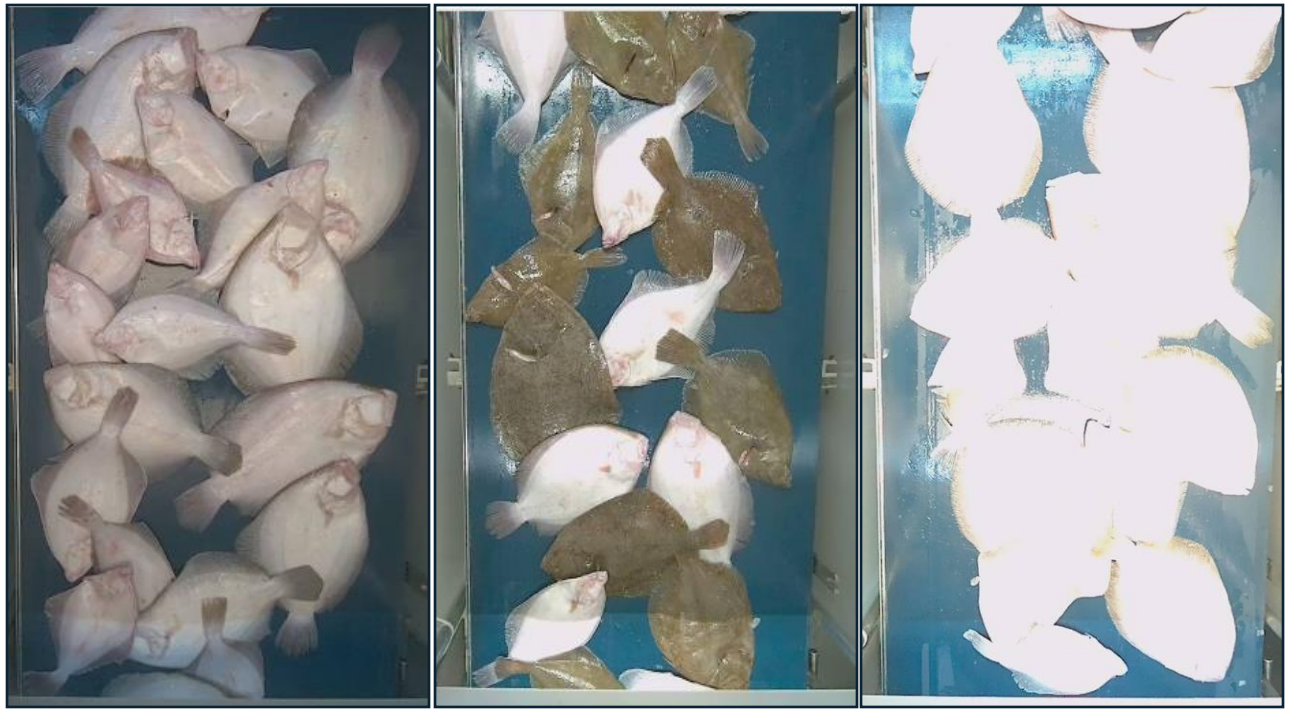
none *moderate* *heavy*

Medegefinancierd door de Europese Unie

Project workflow
Real data

Controlled lab scenario's

- Two **species**: brill or plaice
- **Orientation**:
  - belly, back, or mixed
- **Overlap**:
  - none, moderate, or heavy
- **Illumination**:
  - moderate, normal, or strong



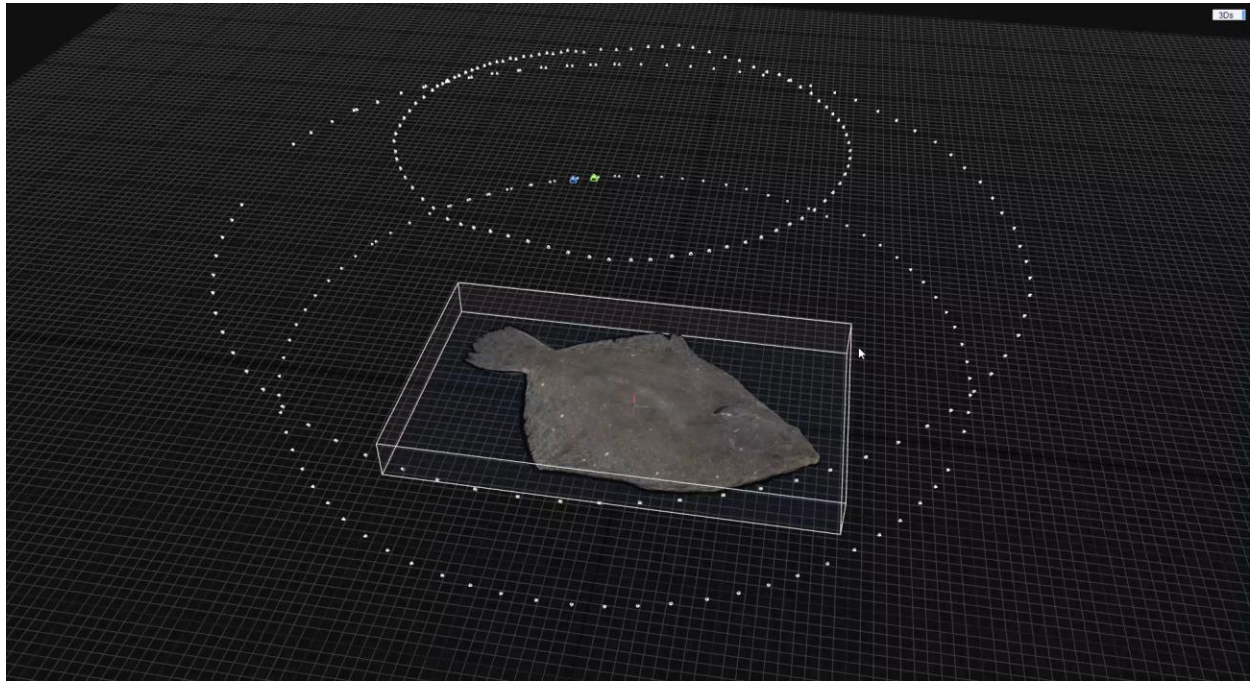
moderate *normal* *strong*

Medegefinancierd door de Europese Unie

Project workflow
Synthetic data

Created by Vintecc

- Photogrammetry-based 3D fish models
- Rendered under varying pose, lighting, and backgrounds
- Automatically labelled during generation



Medegefinancierd door de Europese Unie

Project workflow
Synthetic data

Created by Vintecc

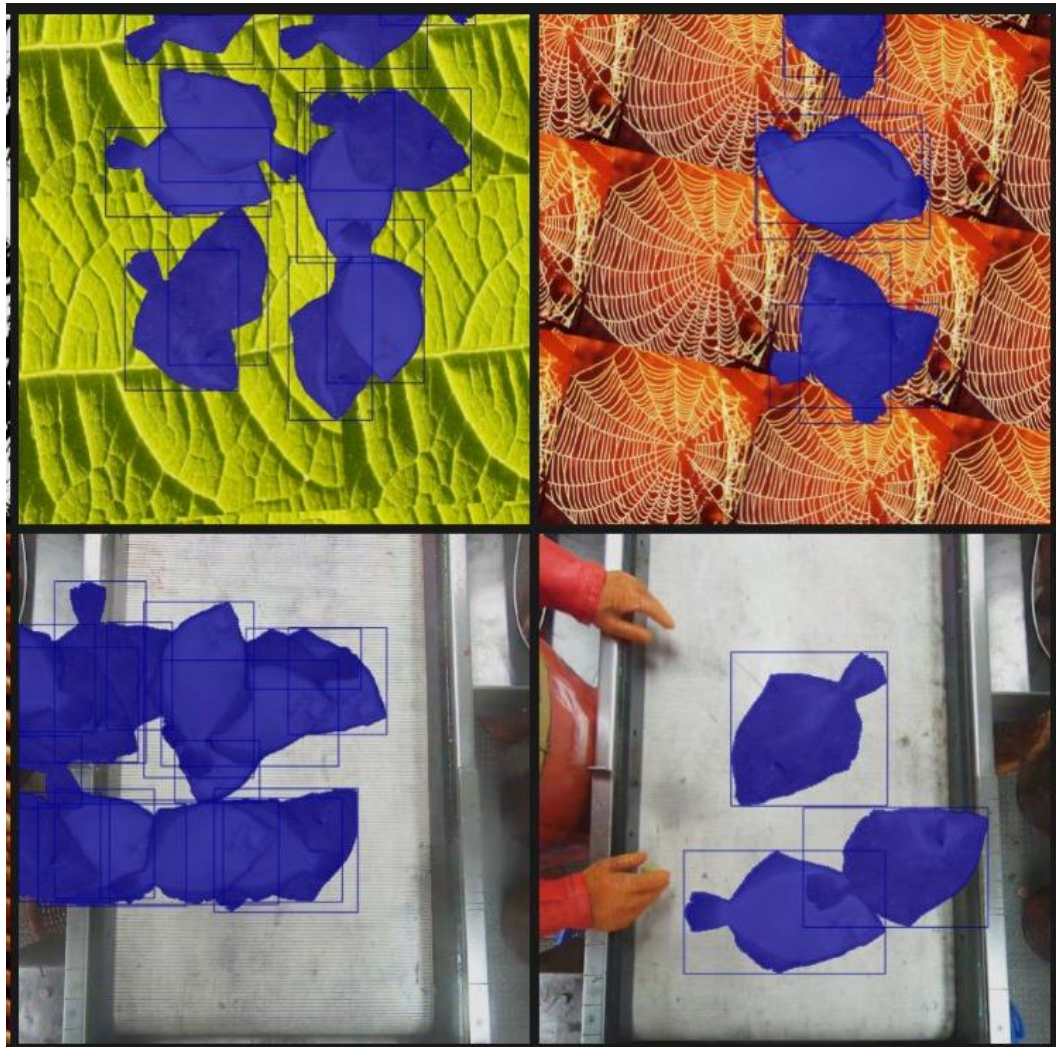
- Photogrammetry-based 3D fish models
- Rendered under varying pose, lighting, and backgrounds
- Automatically labelled during generation

Medegefinancierd door de Europese Unie

Project workflow
Synthetic data

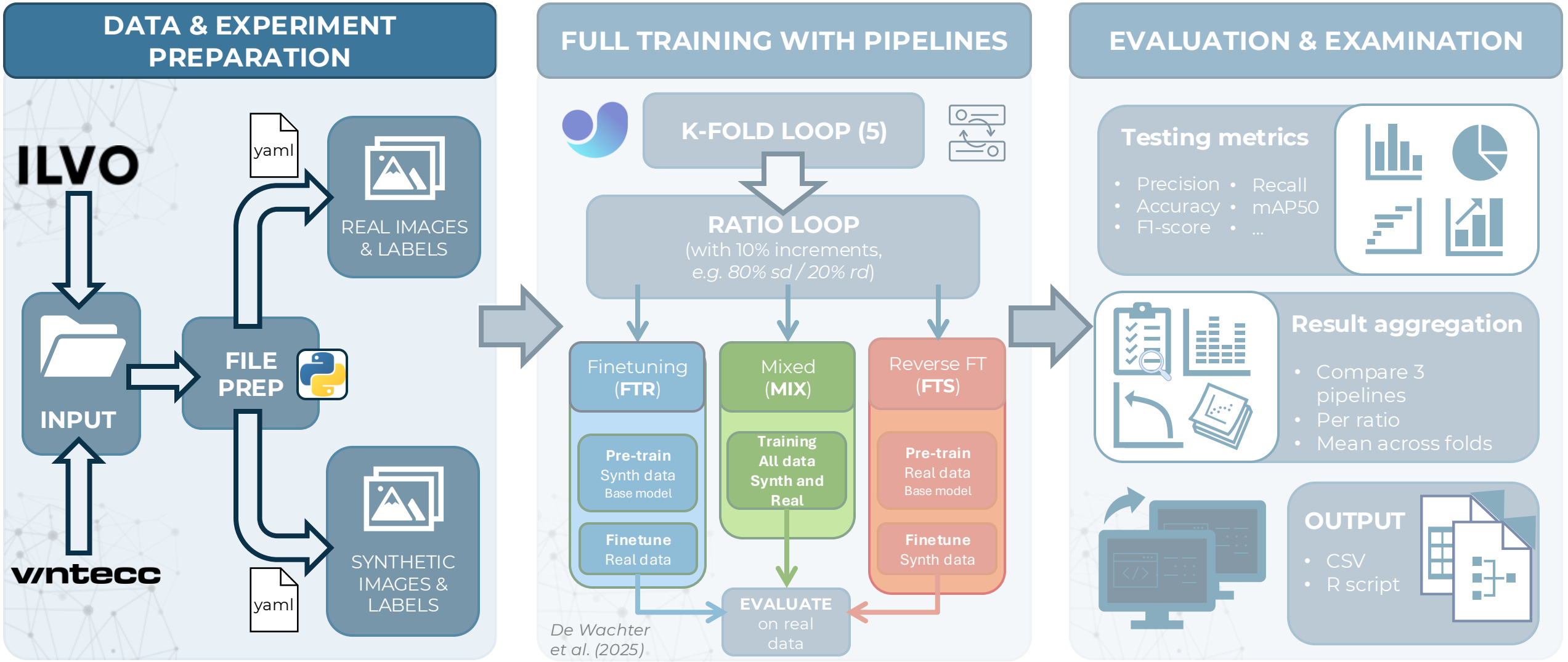
Created by Vintecc

- Photogrammetry-based 3D fish models
- Rendered under varying pose, lighting, and backgrounds
- Automatically labelled during generation

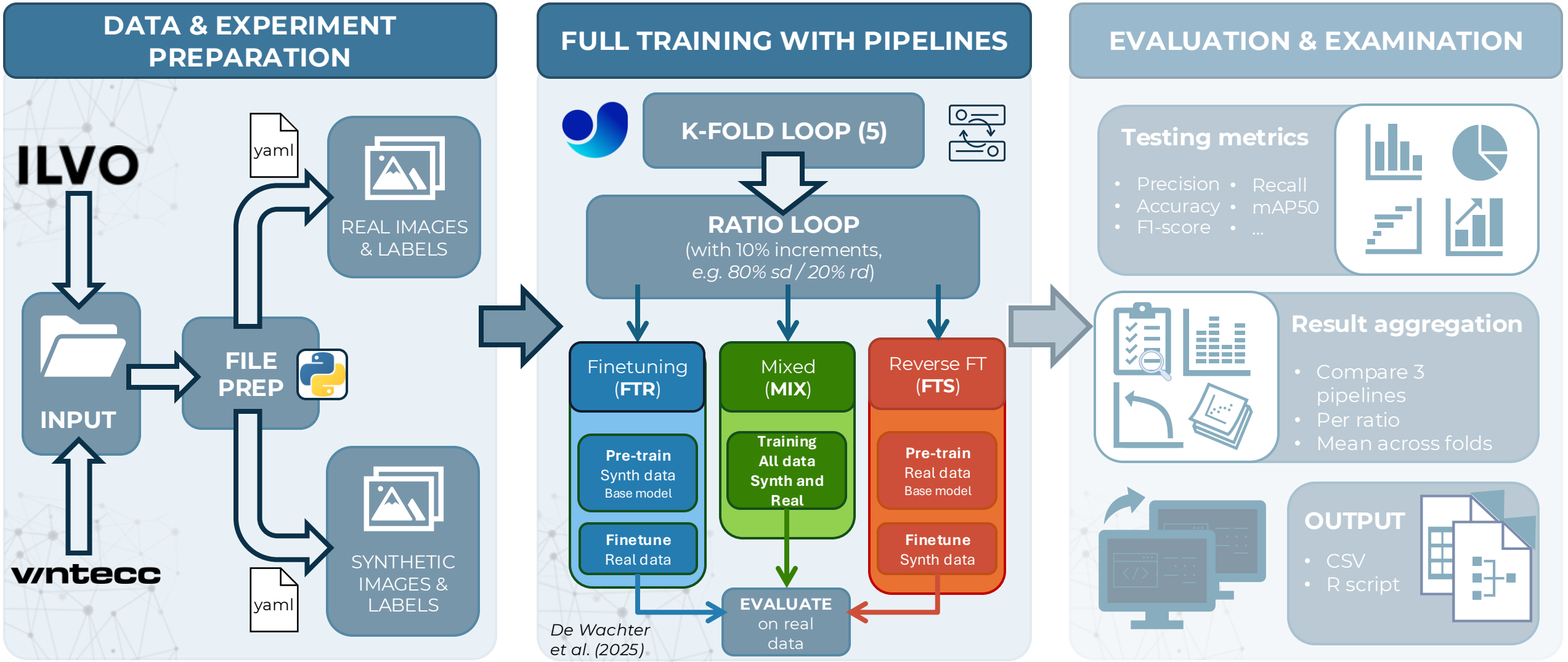


Medegefinancierd door de Europese Unie

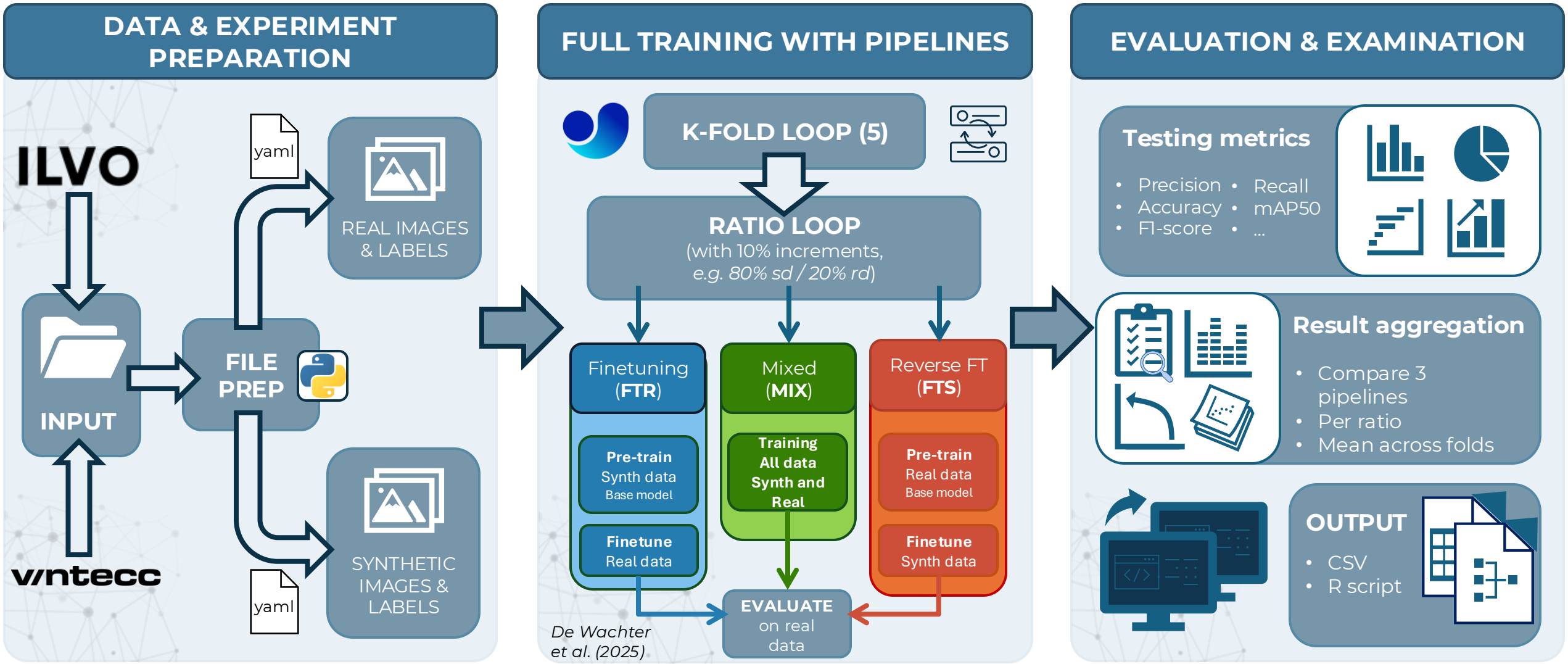
Project workflow
detection & segmentation model training



Project workflow
detection & segmentation model training



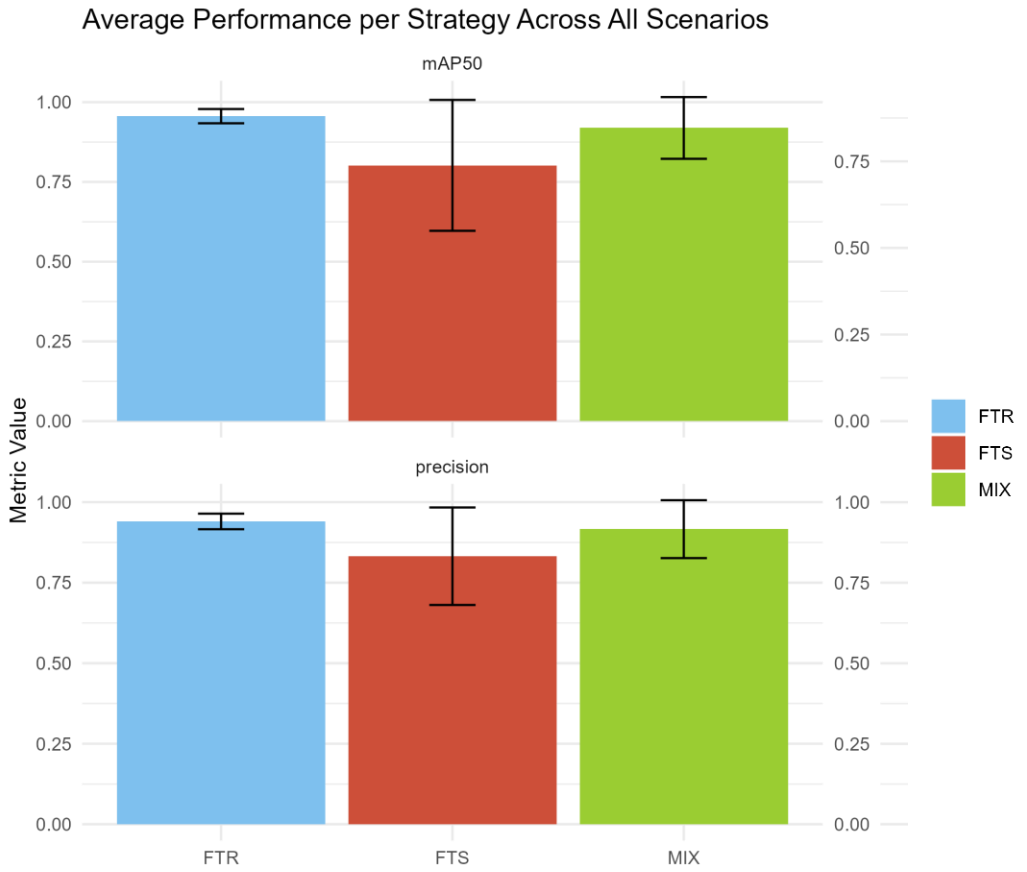
Project workflow
detection & segmentation model training



Initial results
YOLO-based

On the **1st synthetic dataset** (*only brill*) results showed:

- Differences between training strategies
- FTR appeared to perform best
- Performance remained high at low–moderate synthetic ratios
- First decline observed around higher synthetic proportions

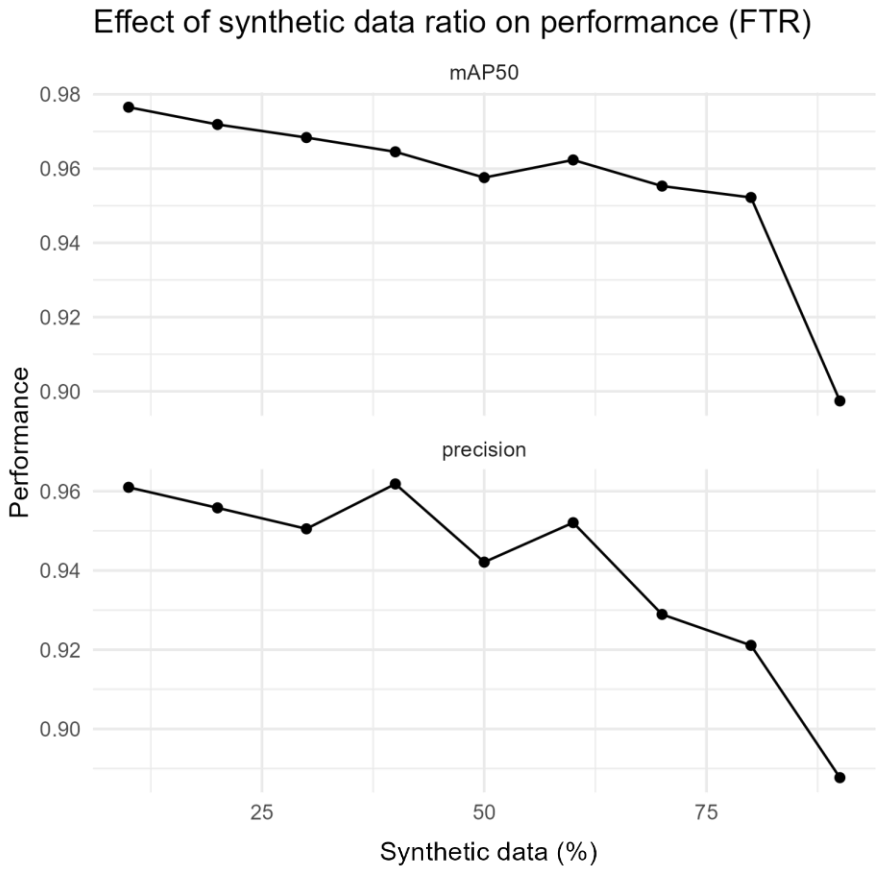


Medegefinancierd door de Europese Unie

Initial results
YOLO-based

On the **1st synthetic dataset** (*only brill*) results showed:

- Differences between training strategies
- FTR appeared to perform best
- Performance remained high at low–moderate synthetic ratios
- First decline observed around higher synthetic proportions

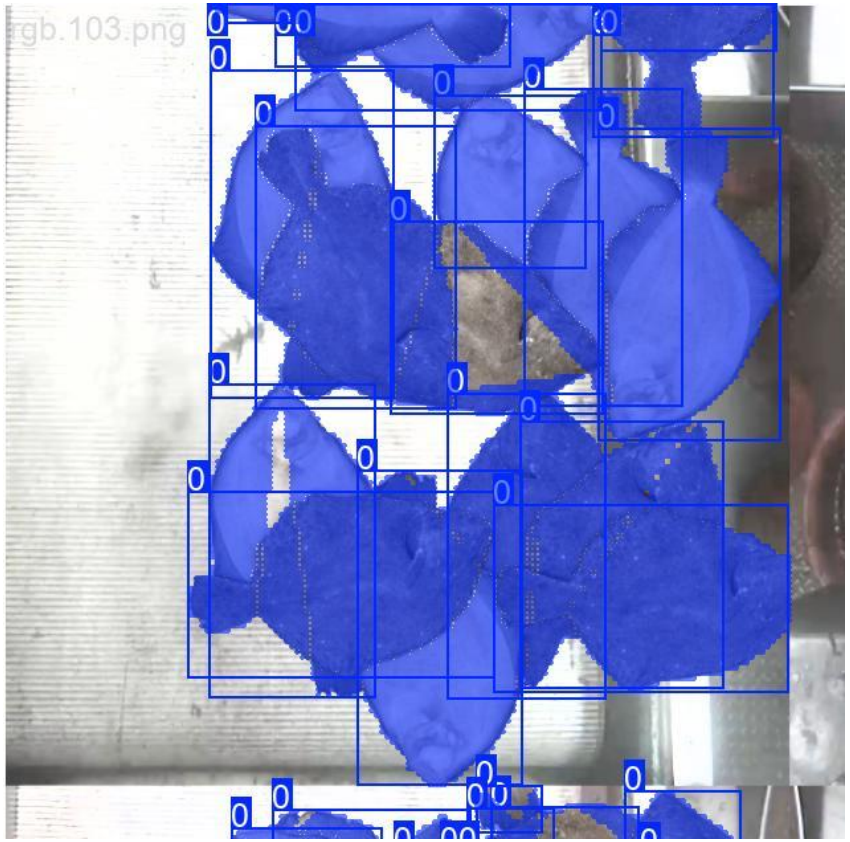


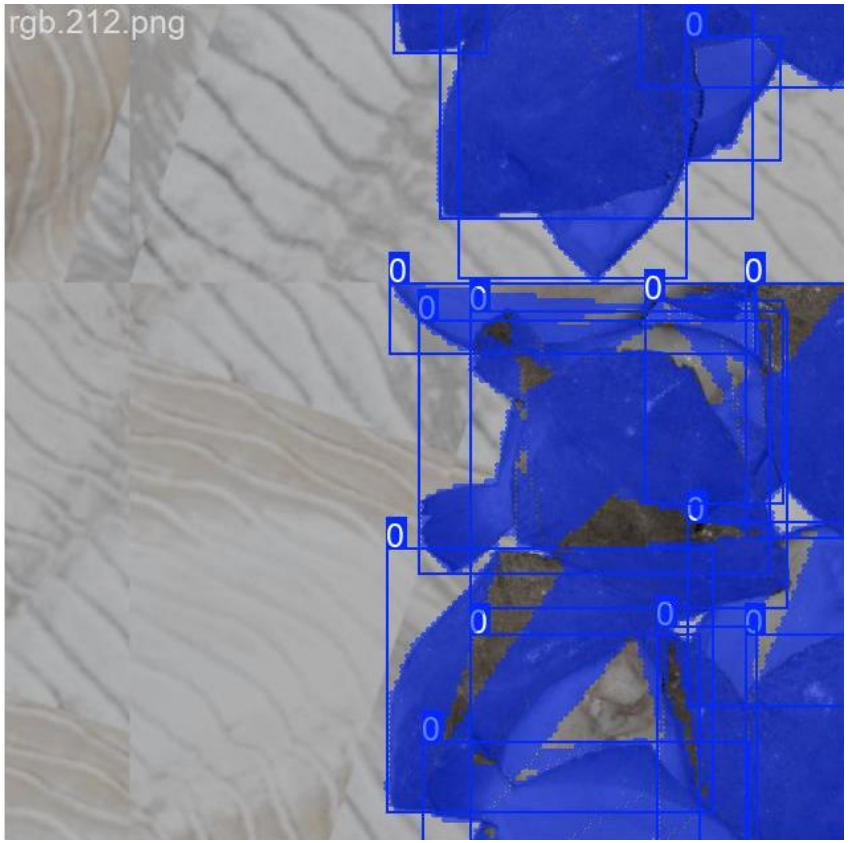
We still performed a **critical workflow review** before accepting interpretation.

Medegefinancierd door de Europese Unie

Where workflow sensitivity appeared

Annotation conversion





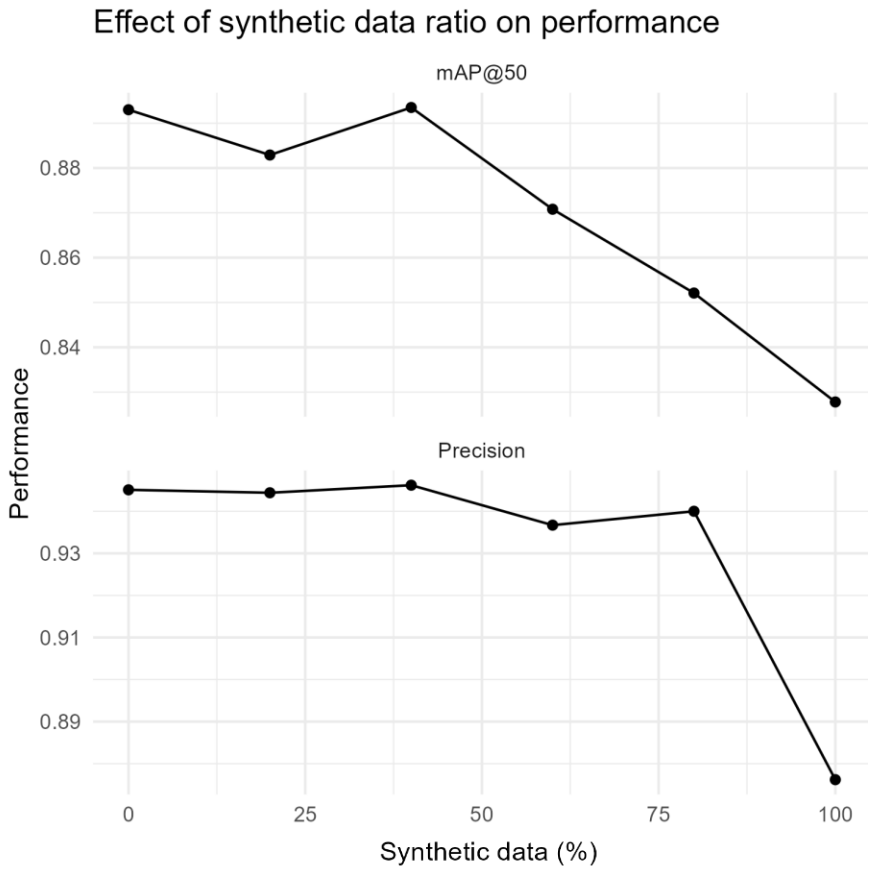
Medegefinancierd door de Europese Unie

Workflow adjustment
RF-DETR

Current status

- Preliminary analysis
- Mixed (MIX) training only
- Ratios tested at 20% increments

Preliminary trend: performance remains higher up to ~40-60% synthetic data, followed by a decline at higher synthetic proportions.



Next Steps

- **Re-run all training** branches using the corrected workflow
- Repeat ratio analysis in **10% increments**
- **Evaluate robustness** across controlled and operational real-world scenarios
- **Scale** synthetic data **diversity**:
  - More individuals
  - Plaice as additional species
  - More complex mixed-species scenes
- Define **where synthetic data adds value**, and where it does not

Take-home Message

- ✓ Synthetic data is **promising**, but **not plug-and-play**
- ✓ Its **value** depends on the **full workflow**: generation, annotation, conversion, training, and validation

SynFish aims to **identify when synthetic data becomes reliable and useful** for AI-based fisheries monitoring, as well as which **workflow choices** determine that **reliability**.

AI4AS | Monday June 29th

Thank you for listening!

Any questions?

Hanne De Rijcke
hanne.derijcke@ilvo.vlaanderen.be
ILVO Marien, Jacobsenstraat 1, 8400 Ostend

Co-funded by the European Union
Europees Fonds voor Maritieme Zaken en Visserij
ILVO
vlaanderen