IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Final Restitution: Learning Optical Flow with Convolutional Networks

FlowNet

Maria Cecilia FLORENZA
Dely Catalina ARDILA
Nhan NGUYEN

P.-H. Conze, V. Burdin, R. Fablet, P. Papadakis, L. Bergantin,G. Andrade-Miranda

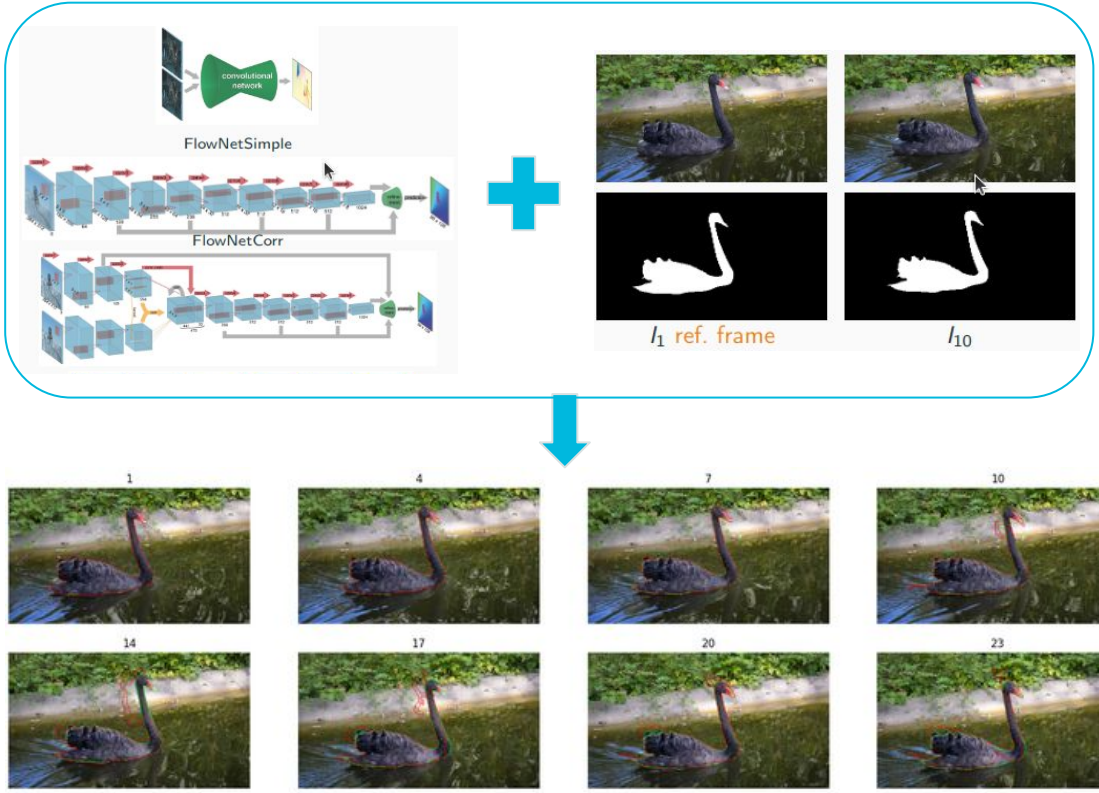
Summary

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

CONTEXT
FlowNet

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

- **Goal:** develop a robust **object tracking method** using **optical flow** computed with **FlowNet**
- **Optical Flow**: Estimates motion between consecutive frames.
- **FlowNet**: A deep learning-based model that learns motion estimation directly from data.

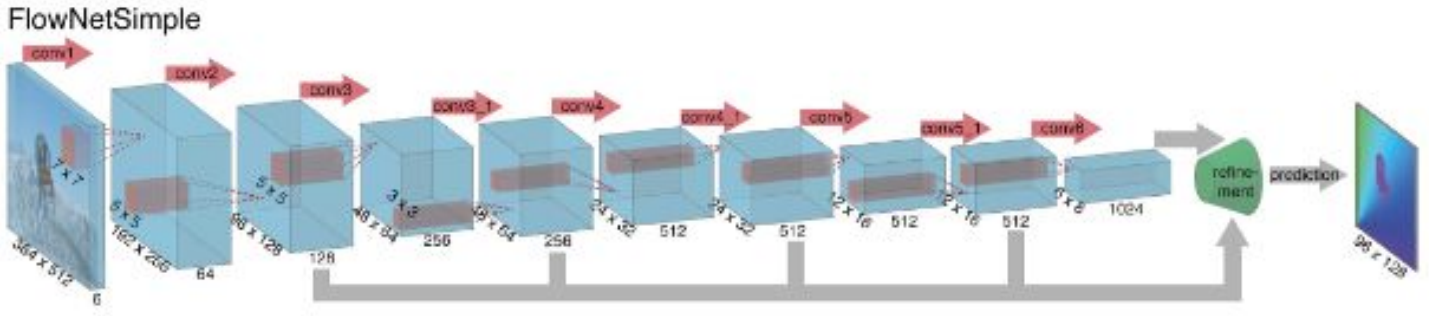


METHODOLOGY
FlowNet

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

- **Pretrained Model:** FlowNetSimple (pretrained) form to compute motion between frames
- Tracking Methods:
  - **Direct Integration**: First frame as the reference throughout the sequence
  - **Sequential Integration**: Previous frame as the reference, updating at each step

GitHub - FlowNetPytorch



Network architecture: FlowNetSimple. The green funnel is a placeholder for the expanding refinement part. The networks including the refinement part are trained end-to-end.

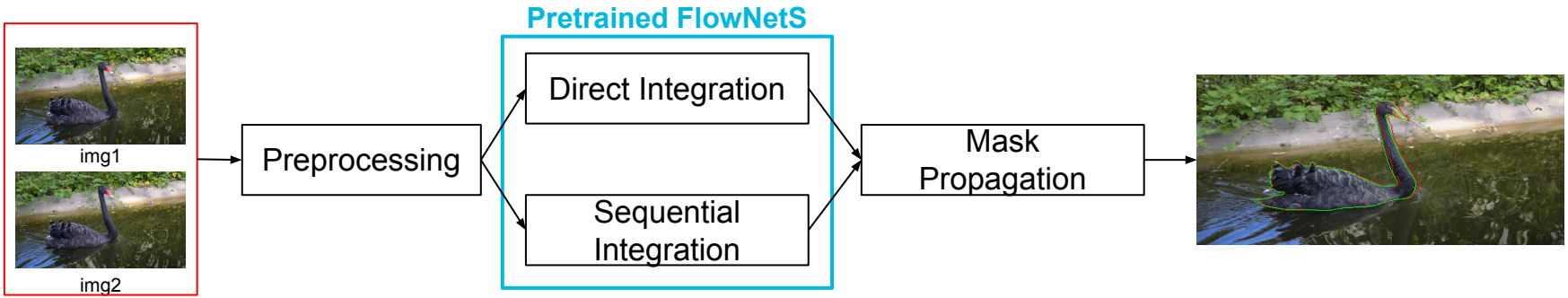
DEVELOPED METHOD

FlowNet

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

3. DEVELOPED METHOD

Self-Supervised Fine-tuning Strategy



Step 1: Optical Flow Estimation

- Run inference with FlowNetS to compute optical flow between frames depending on Integration Method.

Step 2: Mask Generation and Tracking integration

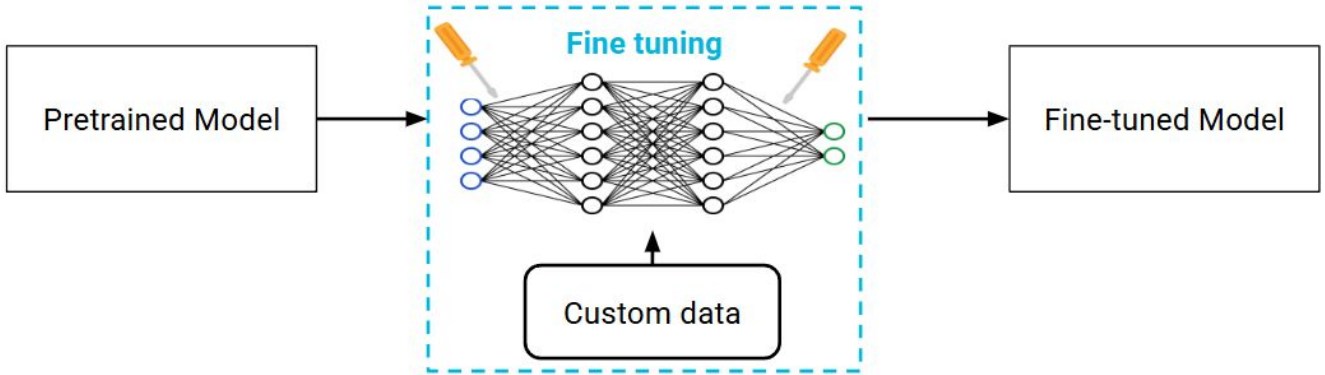
- Use optical flow to estimate object motion regions.
- Generate segmentation masks to highlight moving objects.

Fine-tuning – adapting the model by retraining on our data to improve performance.

✓: Adapt to real, object-centered motion with no ground-truth flow.

!: Training a CNN requires ground truth optical flow, which was unavailable.

- **Pseudo Ground Truth**: Optical flow estimated with Farneback as a reference.
- **Reconstruction Loss**: Warp the image to be processed using the reference image and the predicted optical flow and compare it with the real next frame.



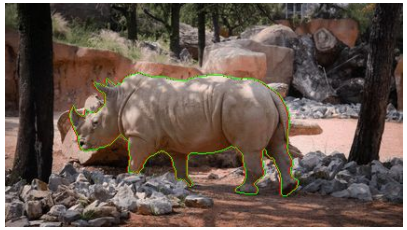
RESULTS

FlowNet

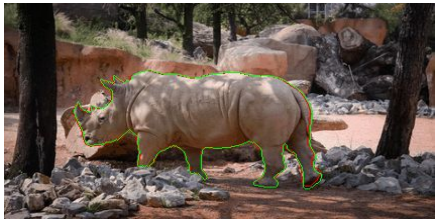
IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

4. RESULTS

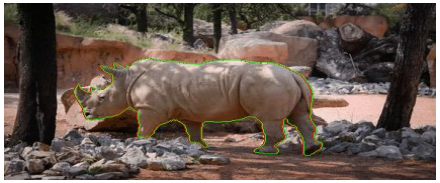
Results on training data



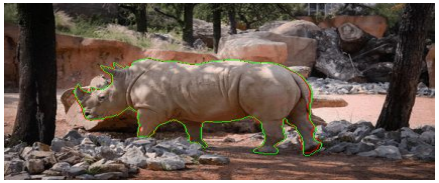
Pretrained model - Direct



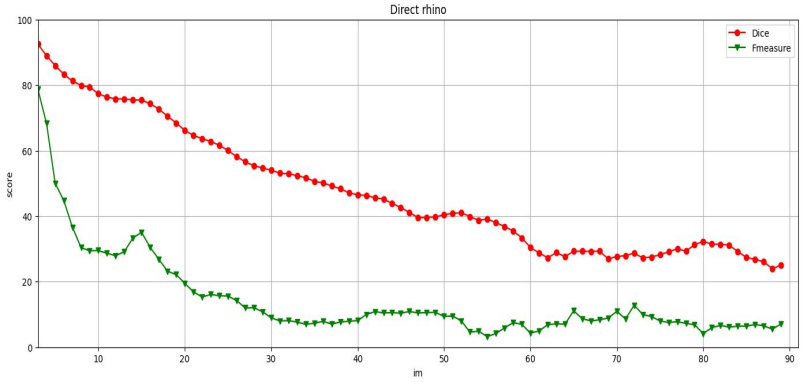
Pretrained model - Sequential

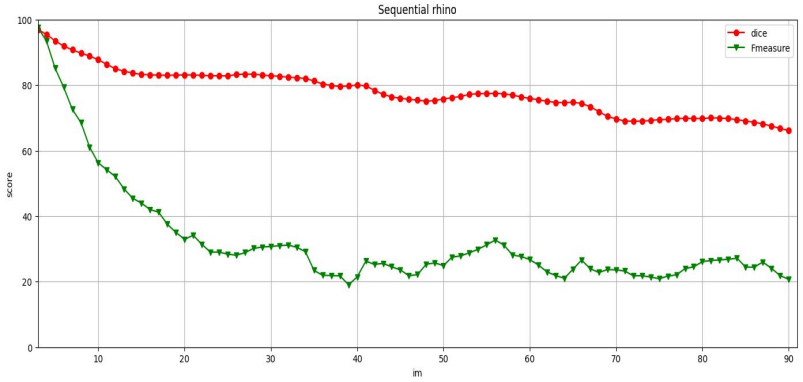


Fine Tuned model - Direct



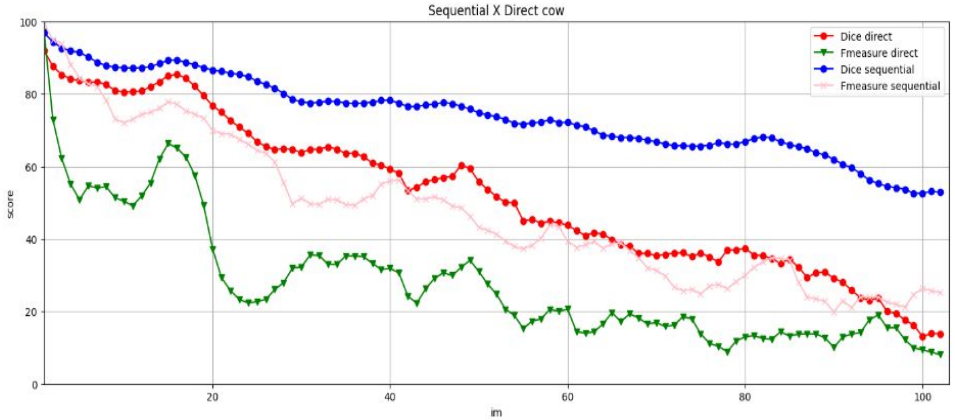
Fine Tuned model - Sequential

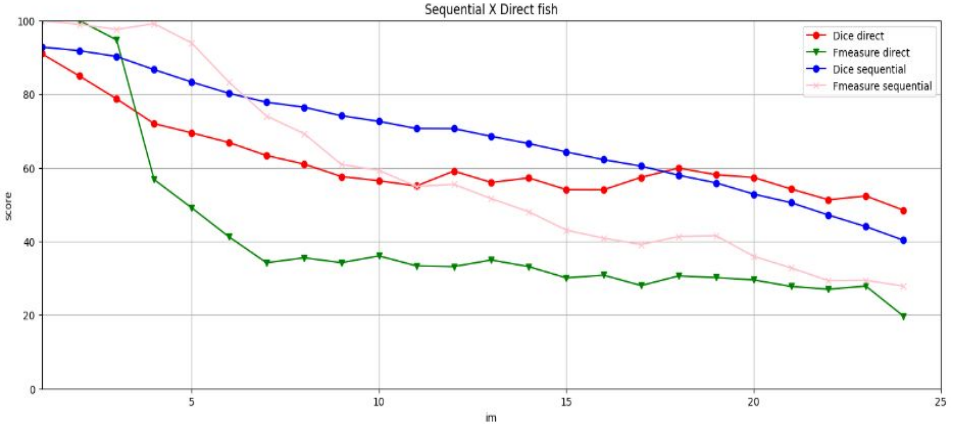


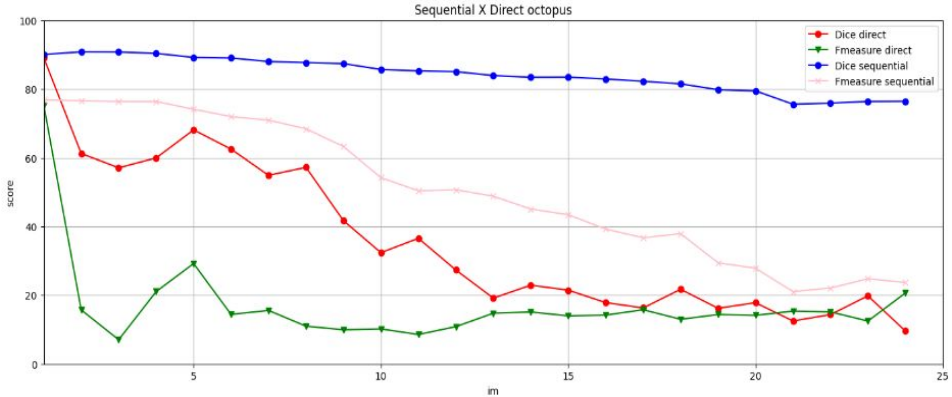


4. RESULTS

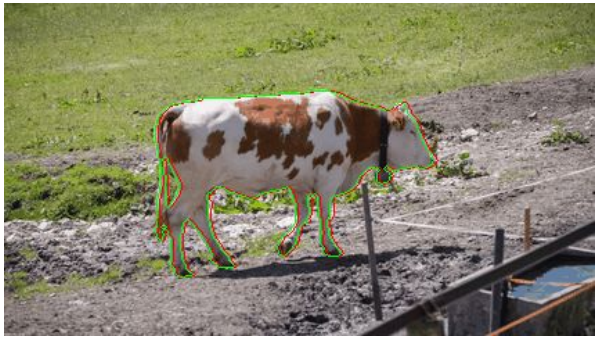
Results on test data



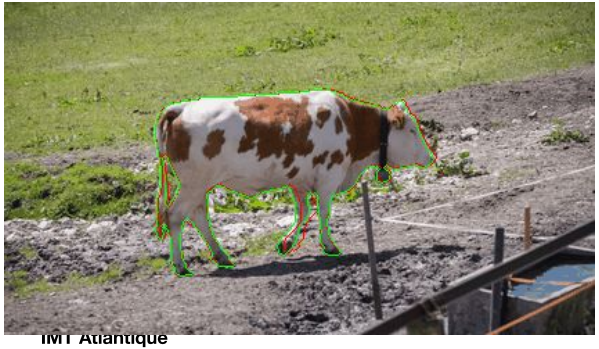




Direct



Sequential



CONCLUSION
FlowNet

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Tracking Performance:

- **Direct** Tracking: Better for smooth motion sequences
- **Sequential** Tracking: More robust but prone to error accumulation

Optical Flow Estimation:

- **FlowNetS** : **accuracy** depends on the **training data**
- **Finetuning** helped **adapt** the model, was limited by the lack of true ground truth

Future Improvements:

- Use a **better pseudo ground truth** for training.
- Fine-tune the model on a **larger dataset** for improved generalization.

Q & A
FlowNet

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

THANK YOU FOR LISTENING !

FlowNet

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom