



IMT Atlantique

Bretagne-Pays de la Loire
École Mines-Télécom

FLOWNET2 IMPLEMENTATION LILIAN, HOANG & PIERRE

SOMMAIRE

1. FLOWNET2 VS FLOWNET
2. IMPLEMENTATION
 - 2.1 Code used
 - 2.2 Adaptation to the challenge
3. METHODS & RESULTS ON TRAINING SEQUENCE
 - 3.1 Flownet2 output
 - 3.2 Results
4. RESULTS ON TEST SEQUENCE
5. CONCLUSION
6. APPENDIX



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

SECTION 1

FLOWNET2 VS FLOWNET



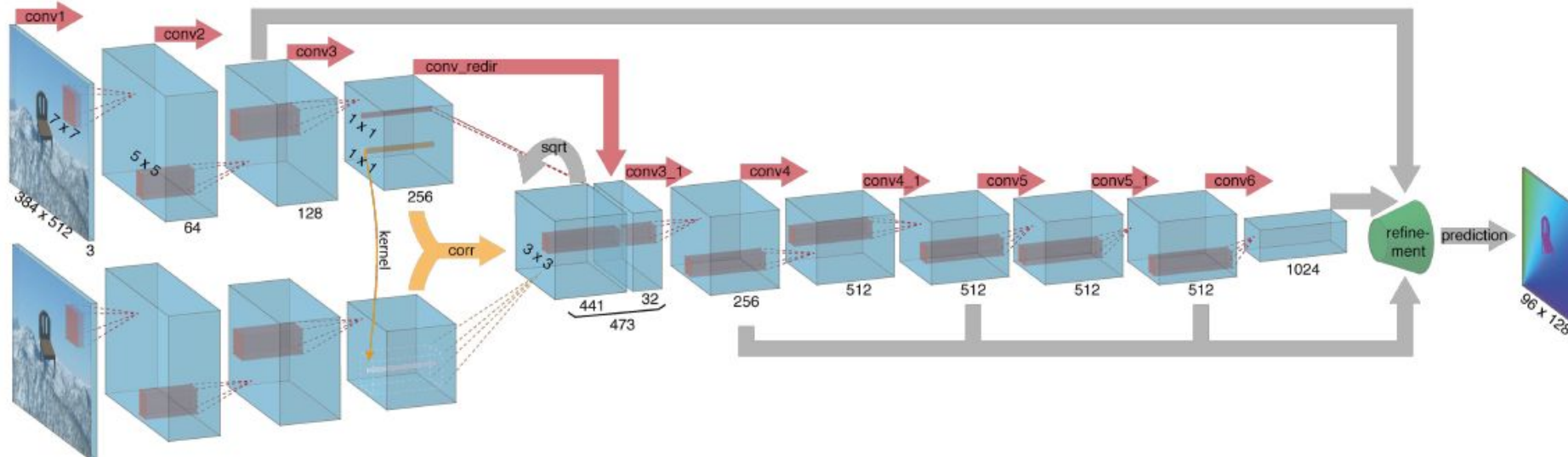
IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

SECTION 1 : FLOWNET2 VS FLOWNET

Remember :
Flownet

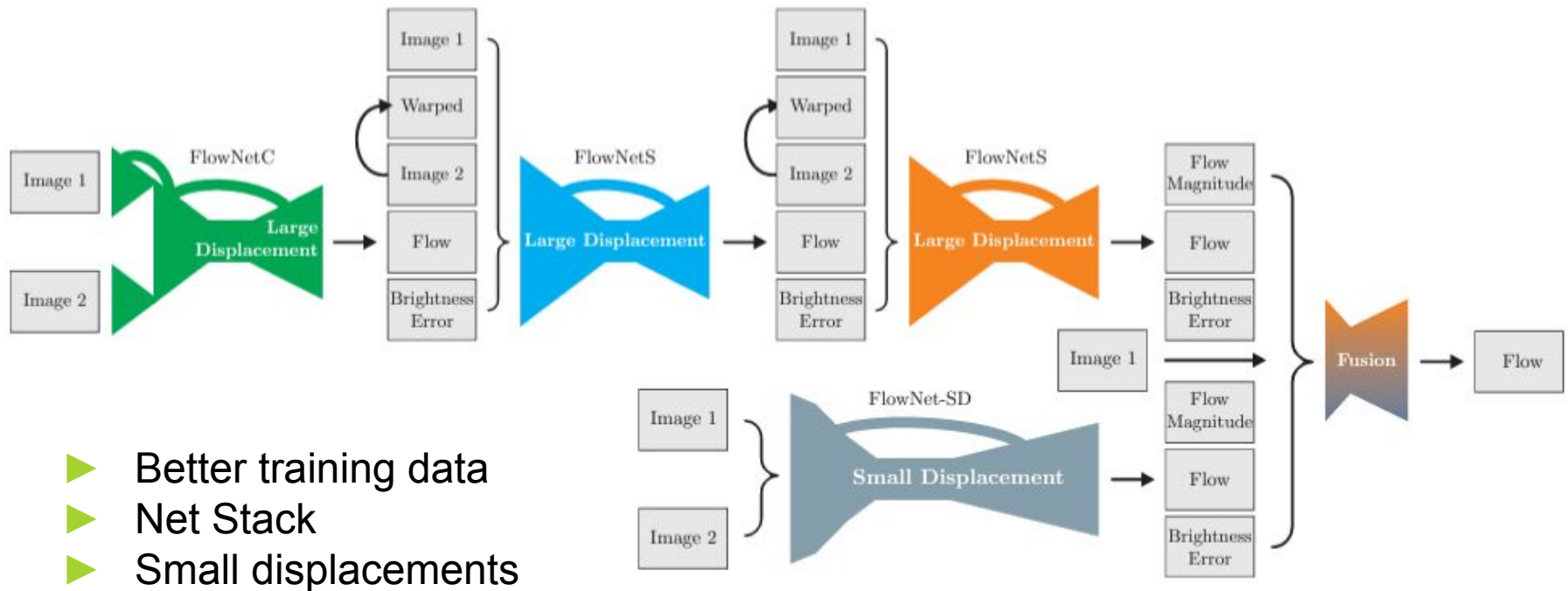
- ▶ CNN
- ▶ Correlation & interpretability

FlowNetCorr



SECTION 1 : FLOWNET2 VS FLOWNET

FlowNet2 :



SECTION 2

IMPLEMENTATION



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

SECTION 2 : IMPLEMENTATION

2.1 code used

Which code we used? :

NVIDIA implementation in pytorch

- ▶ Most used git of flownet implementation
- ▶ Weights converted from caffe
- ▶ Use of colab

BUT

- ▶ Still old (python 3.8) -> incompatibilities
- ▶ home made libraries
- ▶ Hard to understand & to adapt

SECTION 2 : IMPLEMENTATION

2.3 Adaptation to the challenge

Create folders with 2 images in each:

For sequential integration:

img 1 - img 2 | img 2 - img 3 ...

For direct integration:

img1 - img2 | img 1 - img 3 ...

Run inference on each folder -> get an optical flow

Process the optical flow (right format) & either apply to mask or concatenate and apply to mask.

SECTION 3

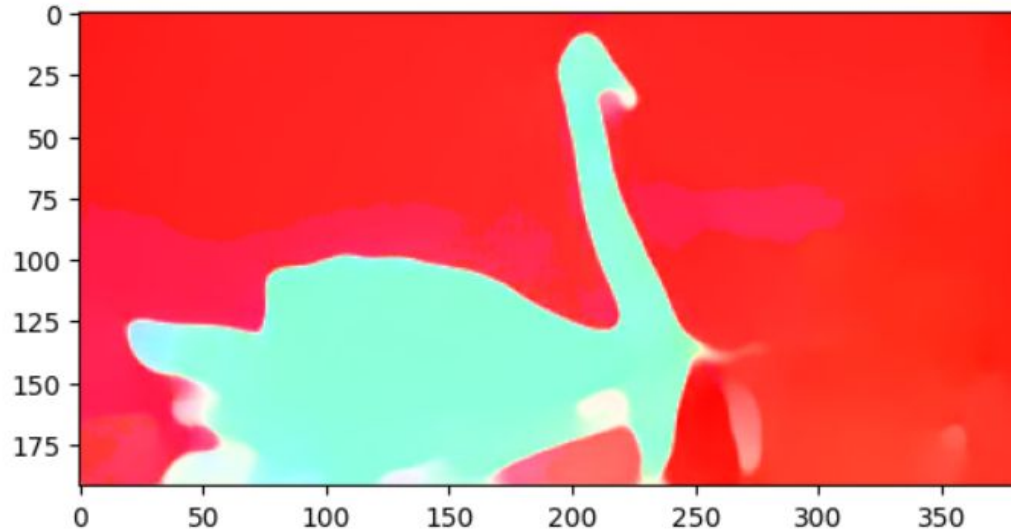
METHODS AND RESULTS ON TRAINING SEQUENCE



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

SECTION 3 :

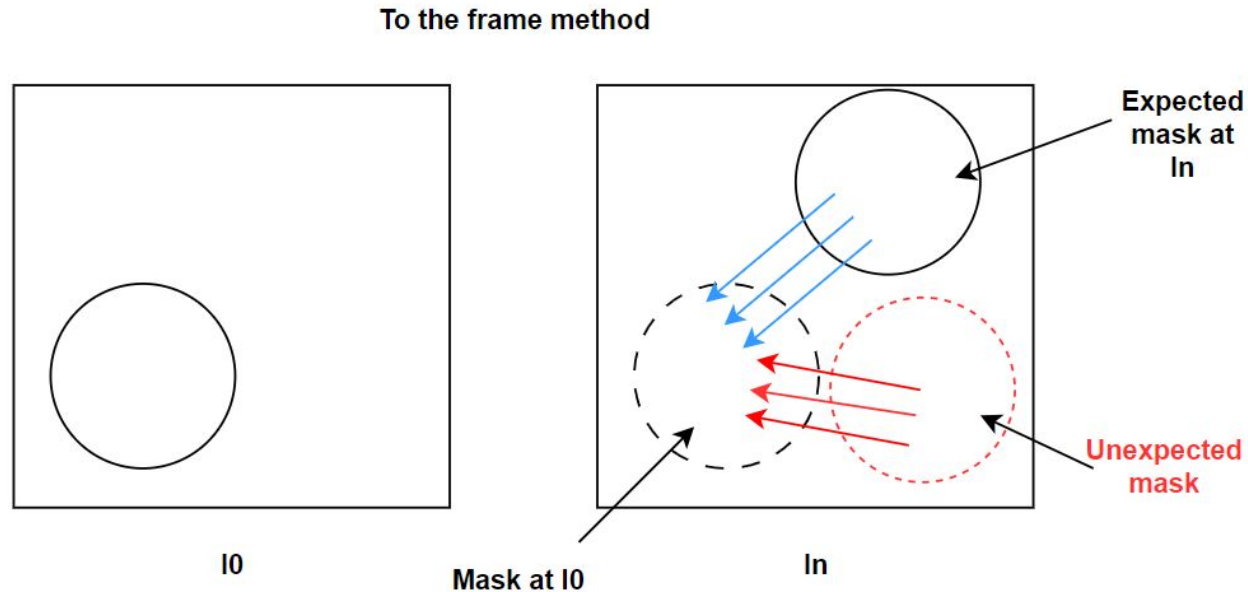
3.1. Flownet 2 output and our adaptation technique



- Competitive outcomes (Retaining the shape of the object)
- Problem of moving camera → Affect negatively mask propagation process

SECTION 3 :

3.1. Flownet 2 output and our adaptation technique

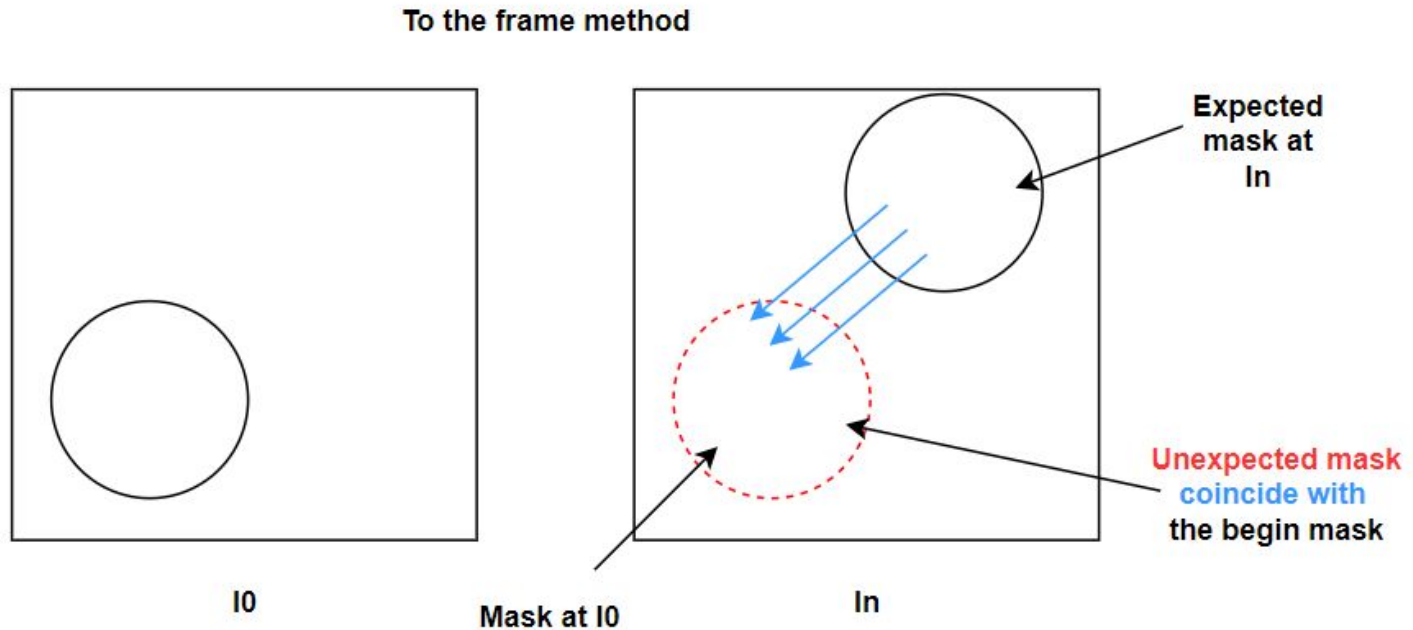


Our solution is to **eliminate** all the background flow

Another problem if removing all background flow

SECTION 3 :

3.1. Flownet 2 output and our adaptation technique

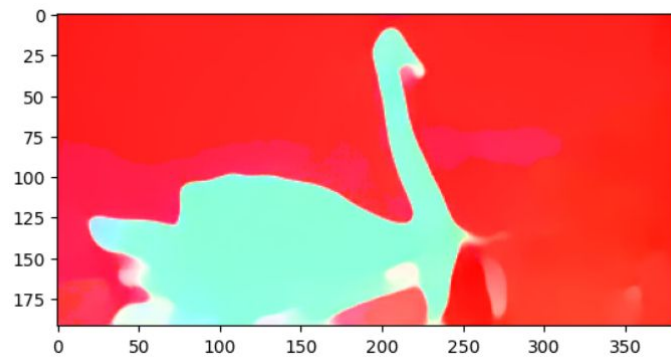


Our solution is to **eliminate** all the background flow **except** the **begin mask area**

If the model is **robust enough**, it should predict this area to **move to other places not stationary**

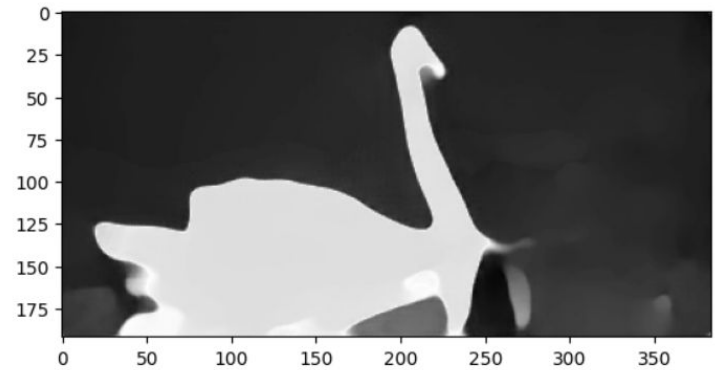
SECTION 3 :

3.1. Flownet 2 output and our adaptation technique



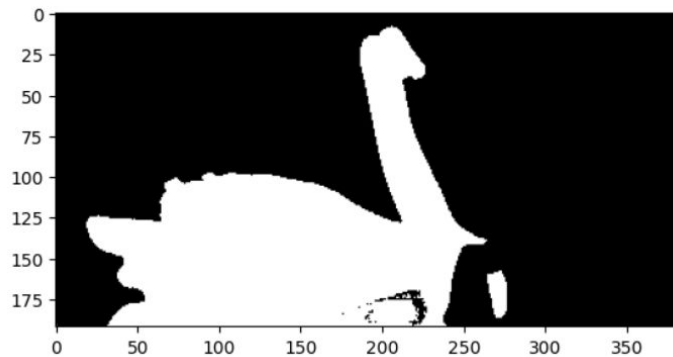
Output

RGB2Gray

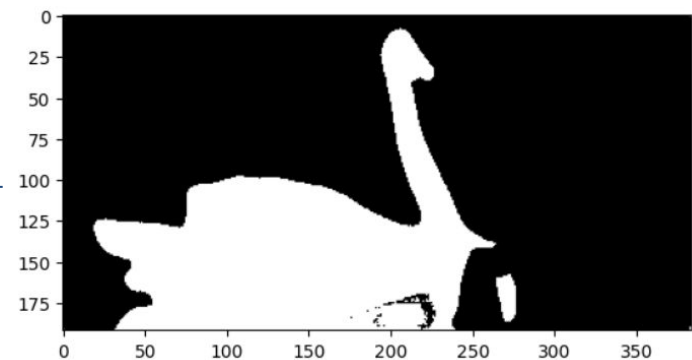


Grayscale

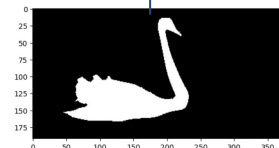
Li's iterative Minimum
Cross Entropy thresholding



Retained area of optical flows



OR

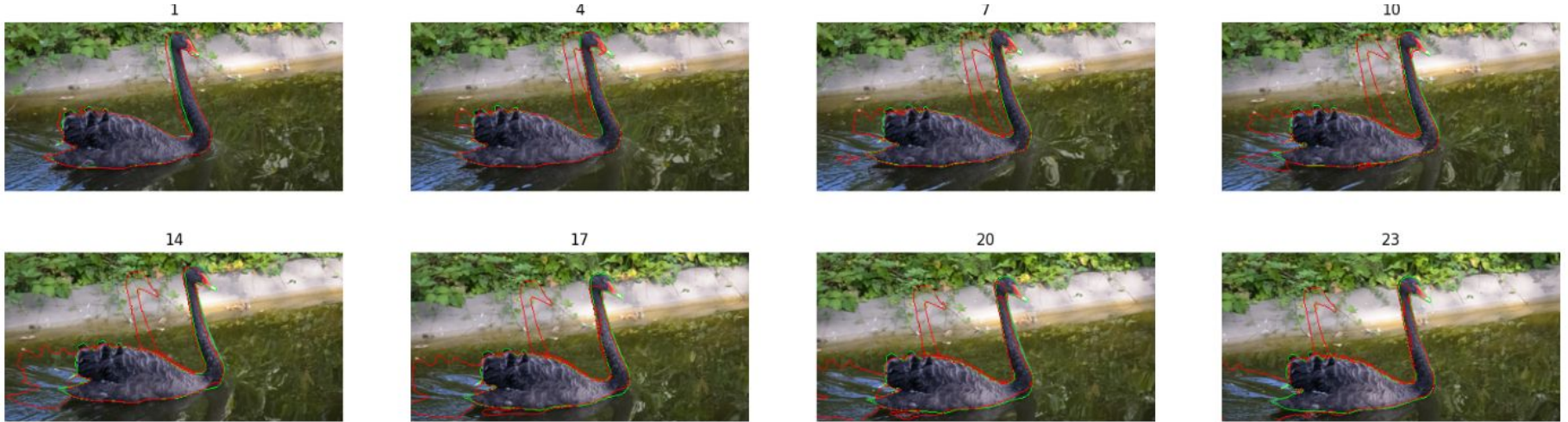


Begin mask

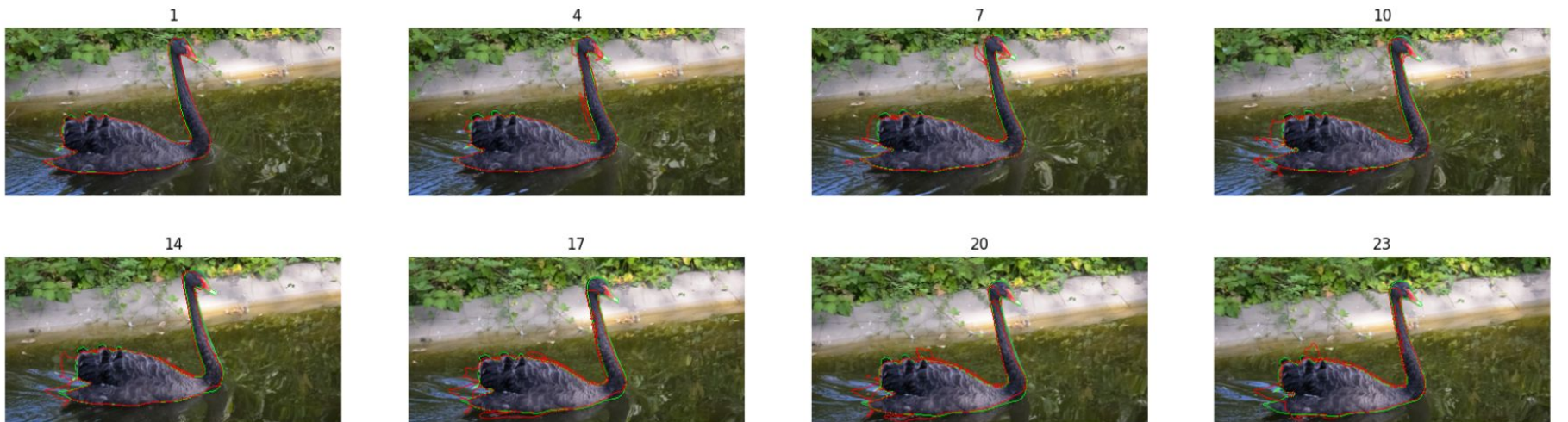
SECTION 3 :

3.1. Results on Training sequence (Direct integration, Swan sequence)

Before our method (Bring Flownet output directly to mask propagation):



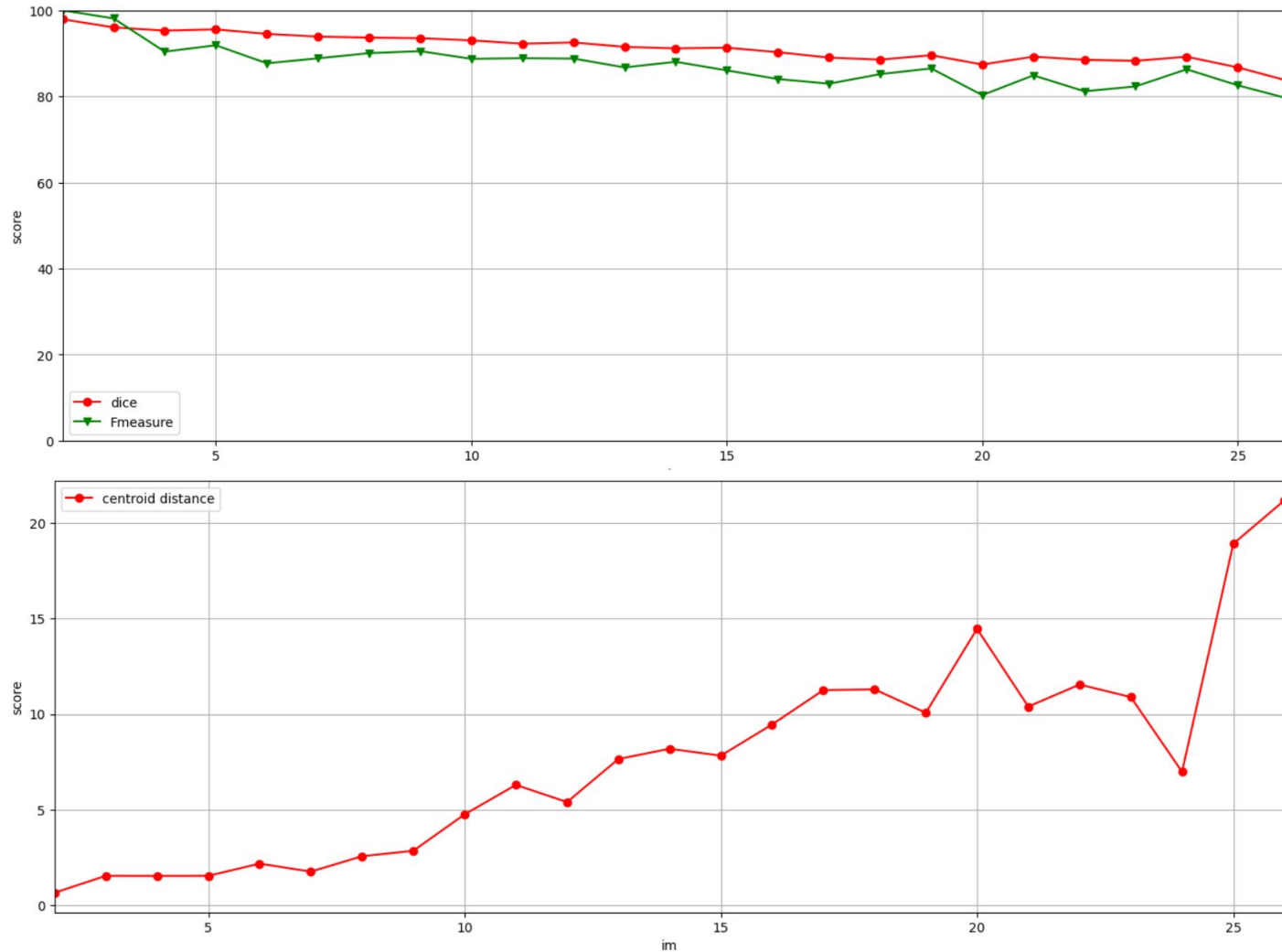
After our method:



SECTION 3 :

3.1. Results on Training sequence (Direct integration, Swan sequence)

After our method:



SECTION 4

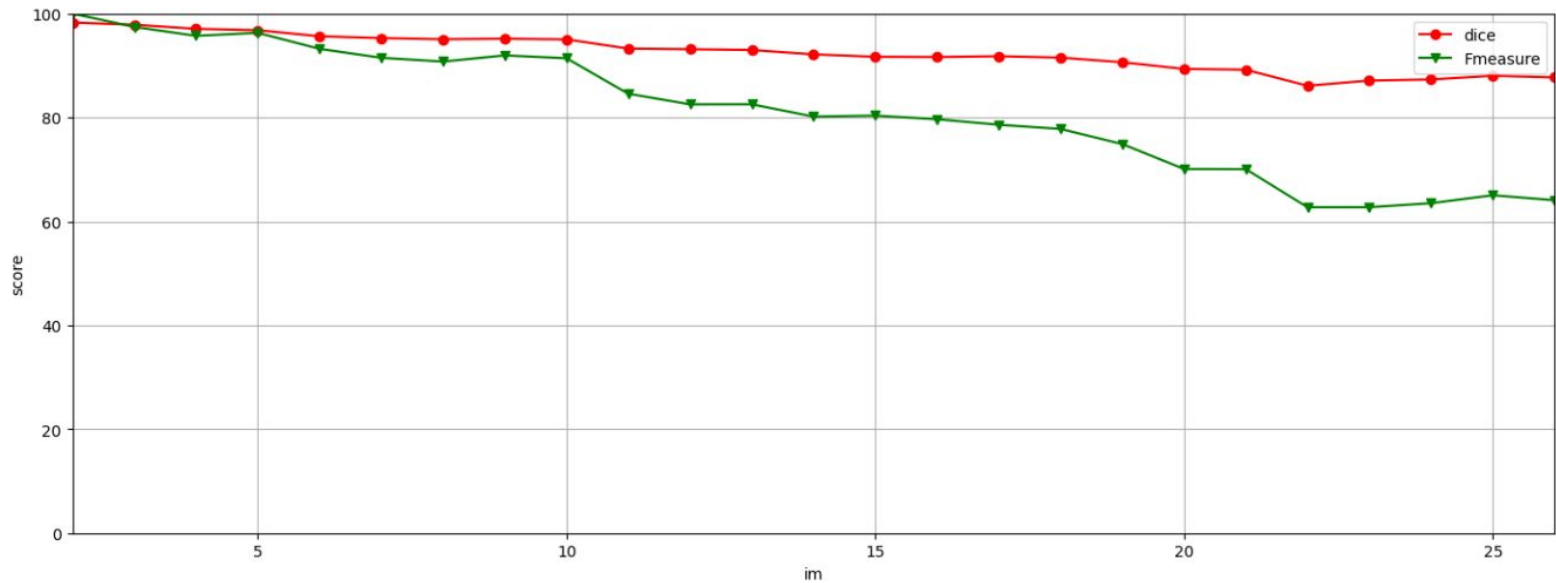
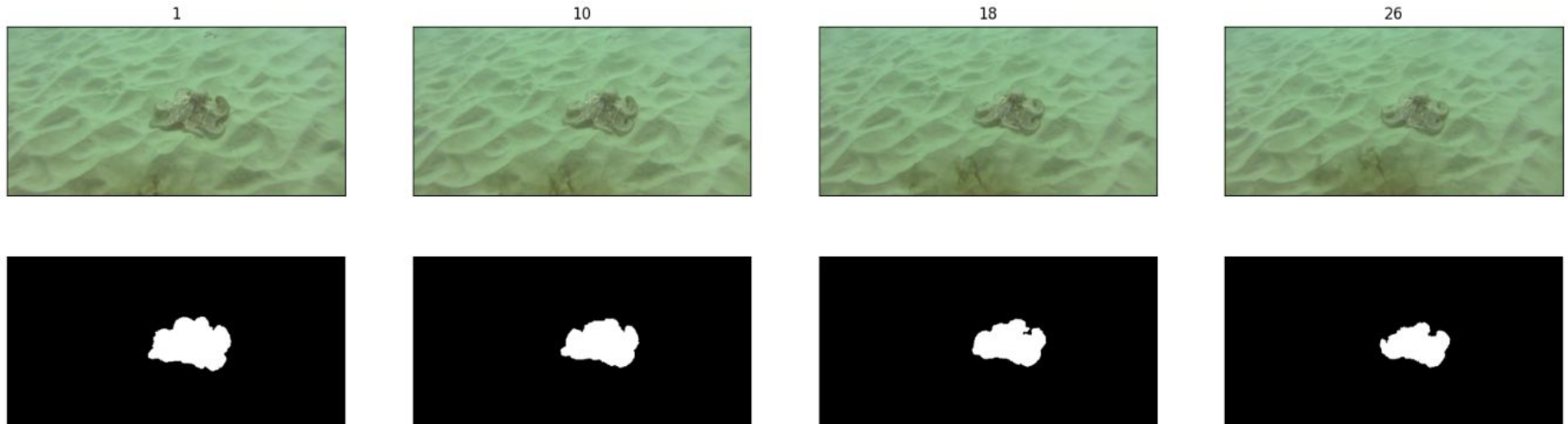
RESULTS ON TESTING SEQUENCE



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

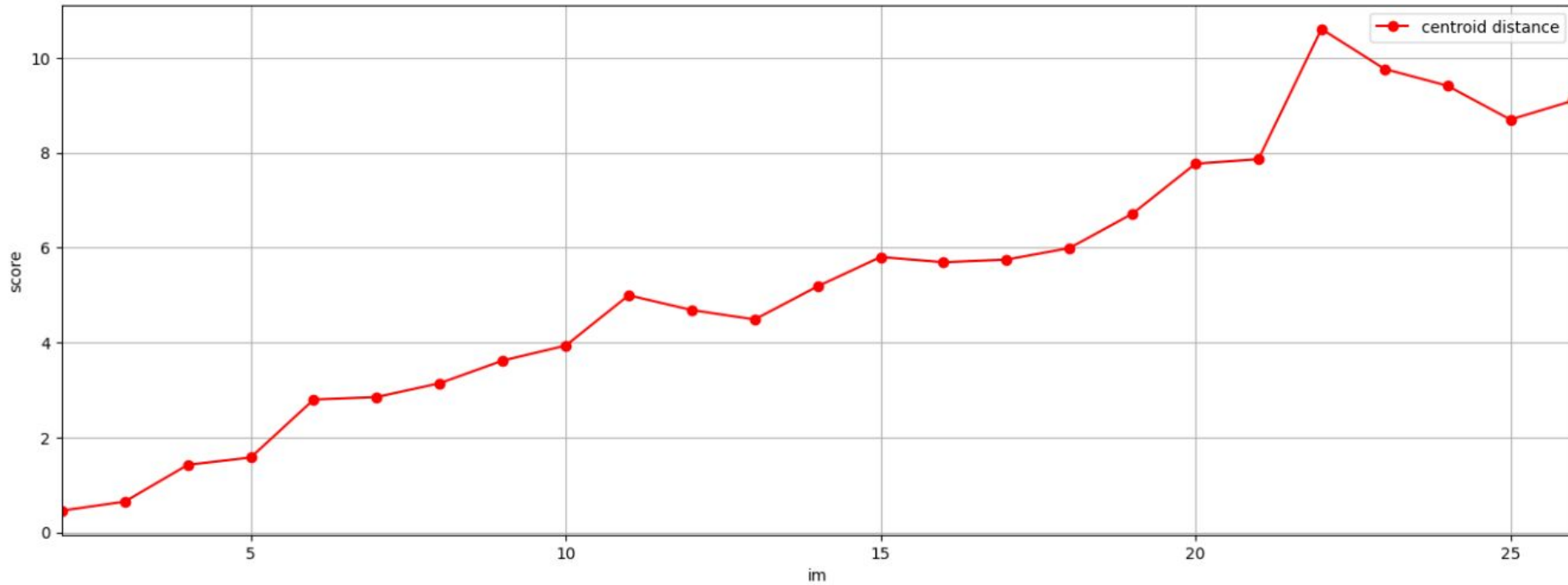
SECTION 4 : Results on test sequence

3.1. Results on Test sequence - Octopus



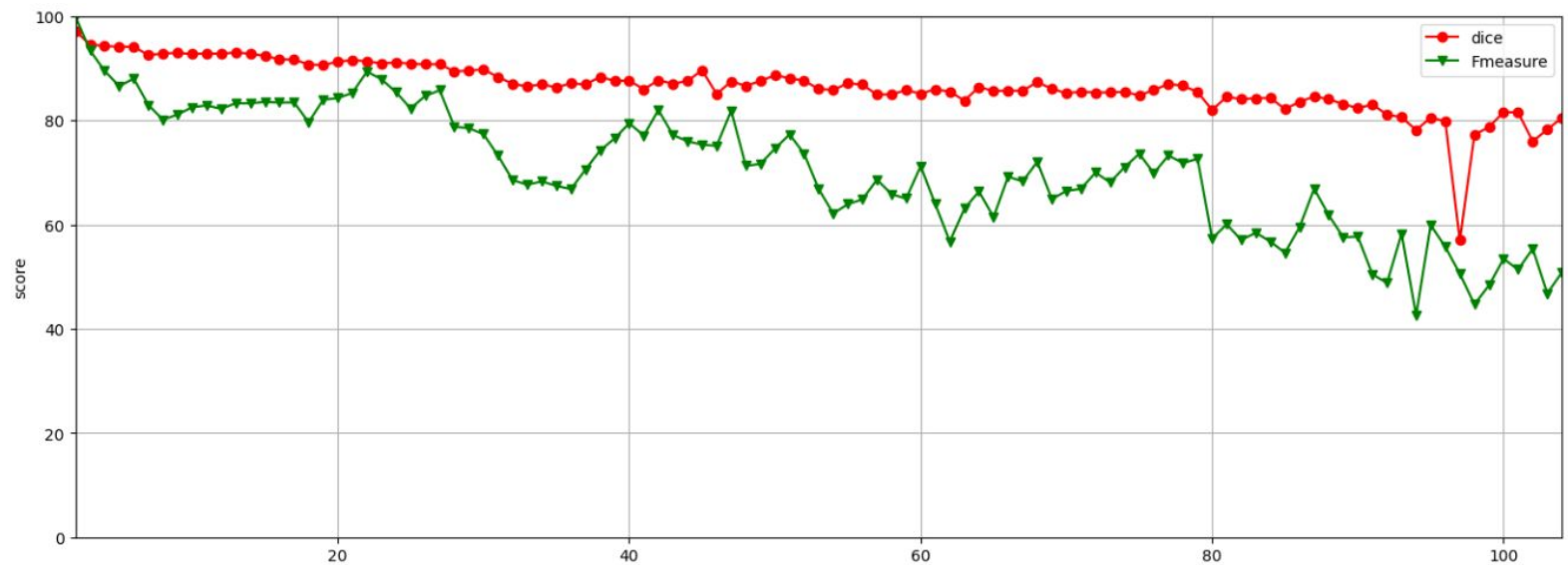
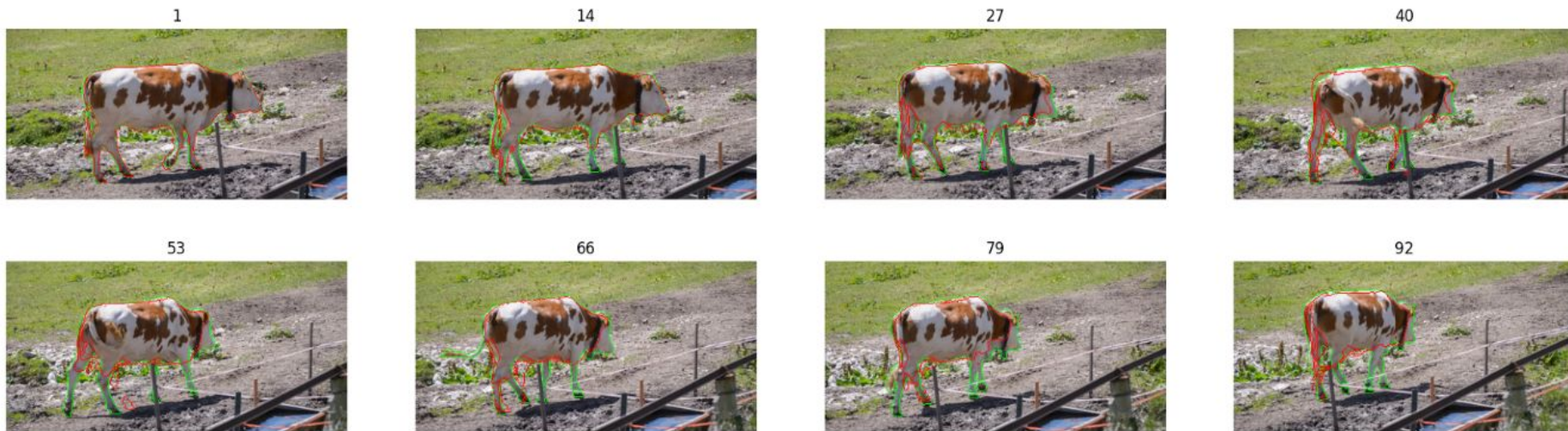
SECTION 4 : Results on test sequence

3.1. Results on Test sequence - Octopus



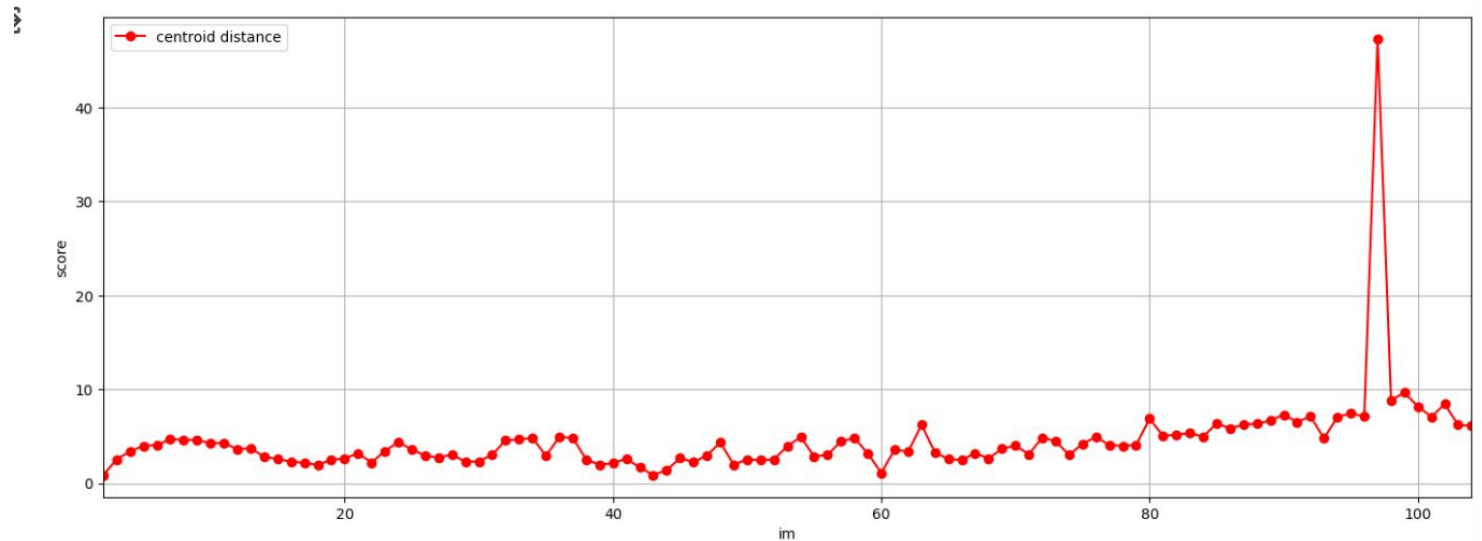
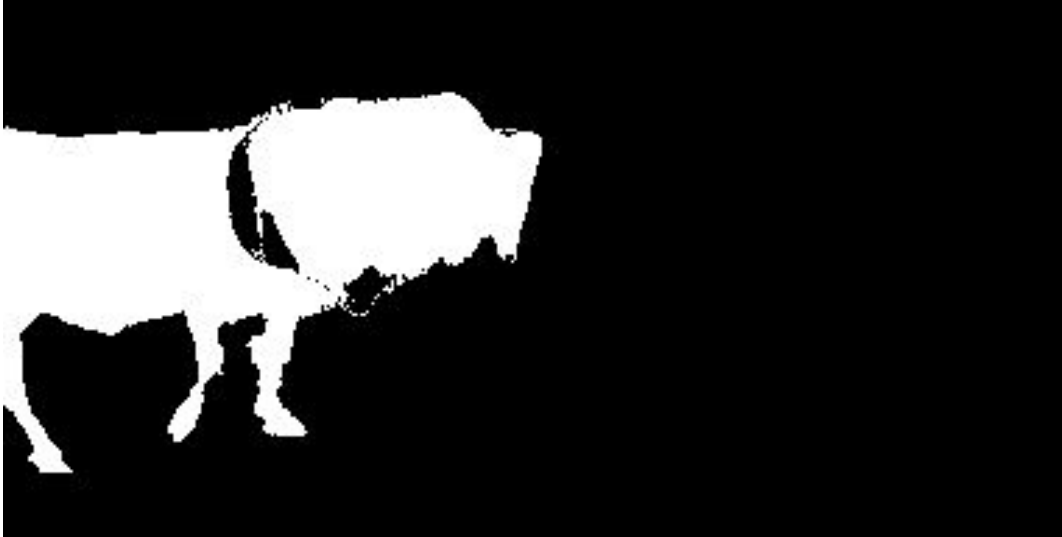
SECTION 4 : Results on test sequence

3.1. Results on Test sequence - Cow



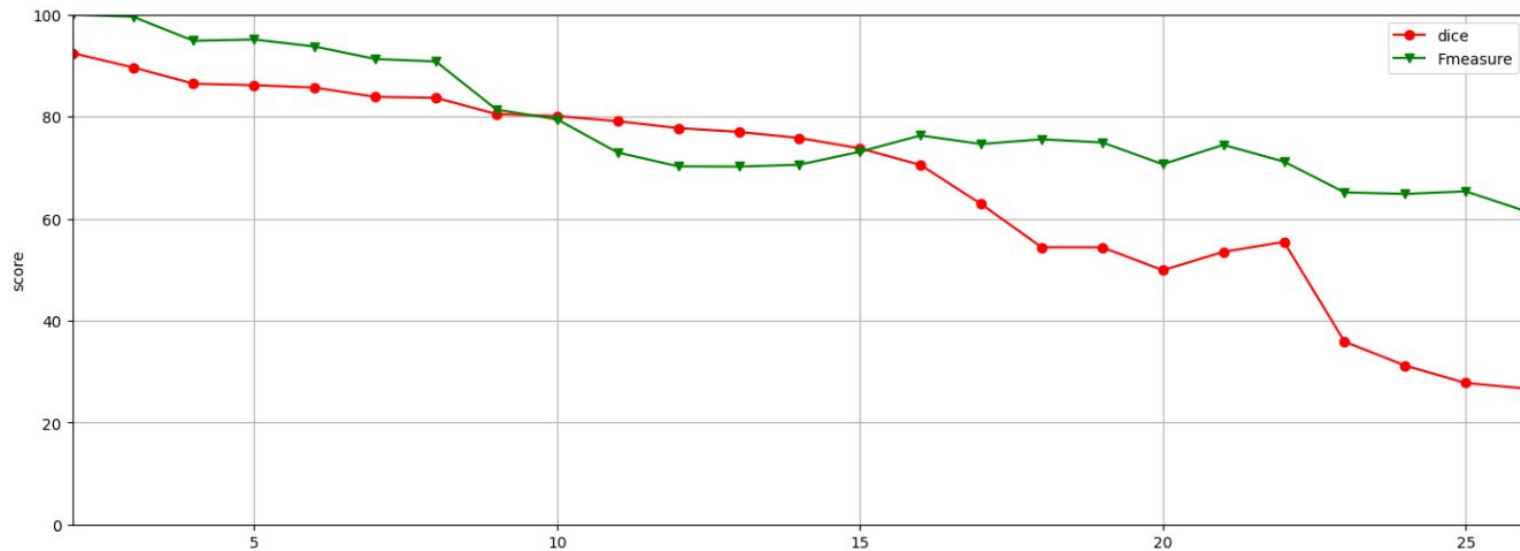
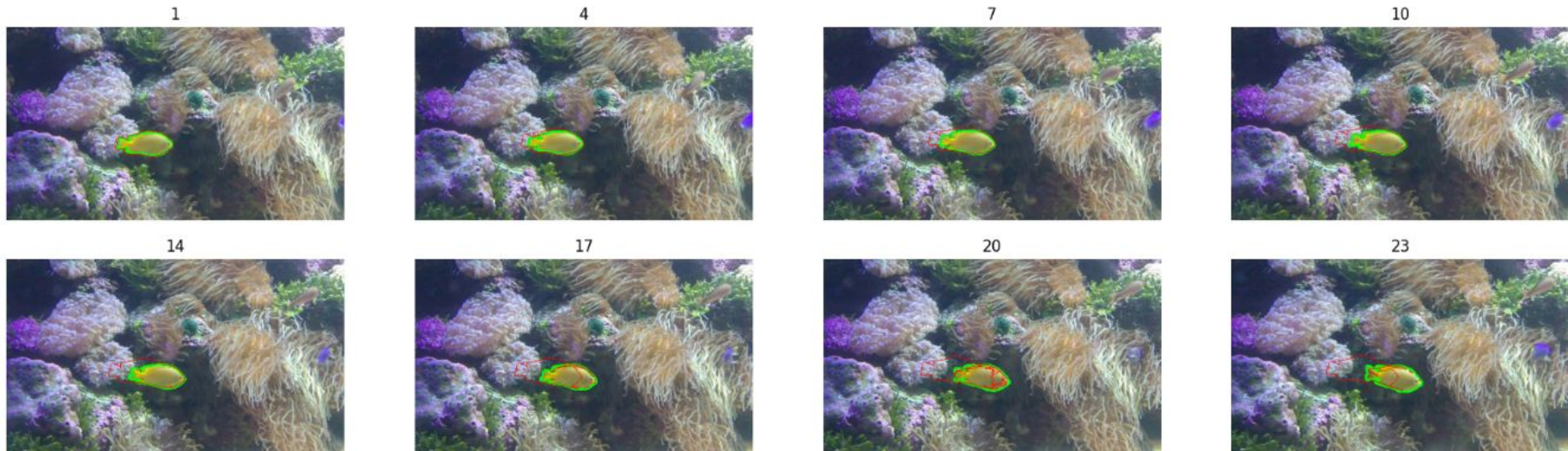
SECTION 4 : Results on test sequence

3.1. Results on Test sequence - Cow



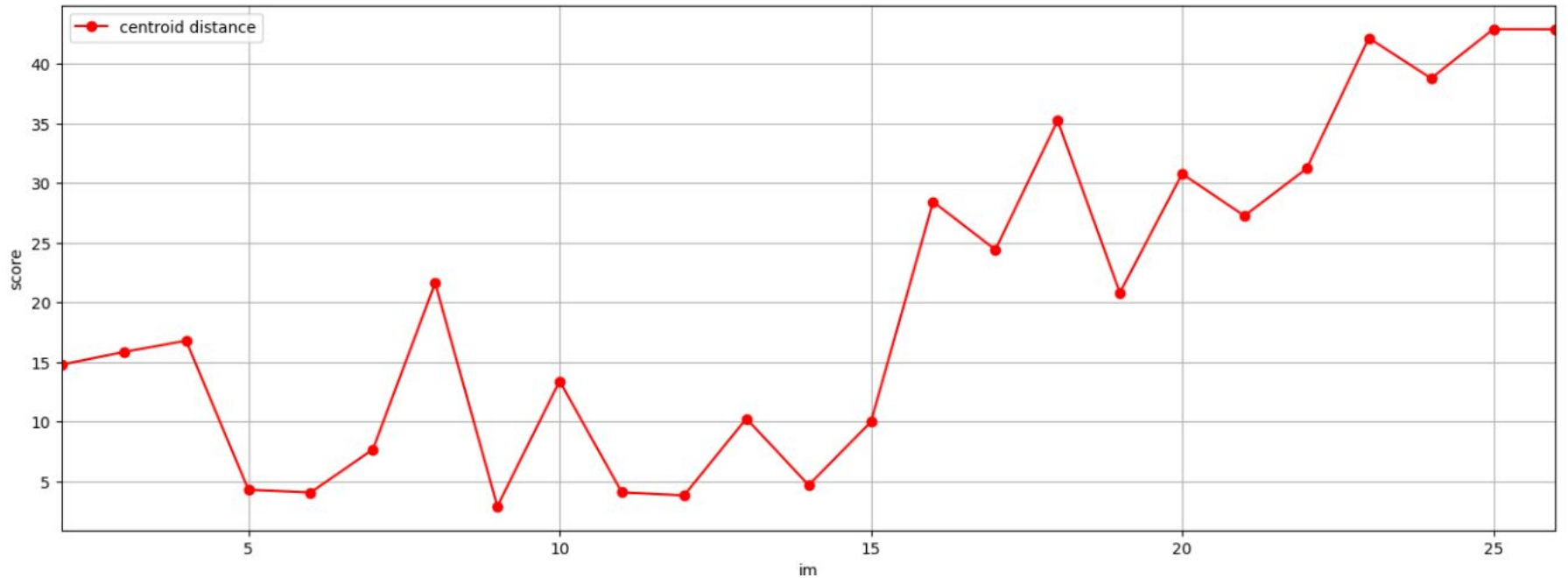
SECTION 4 : Results on test sequence

3.1. Results on Test sequence - Fish



SECTION 4 : Results on test sequence

3.1. Results on Test sequence - Fish



SECTION 5

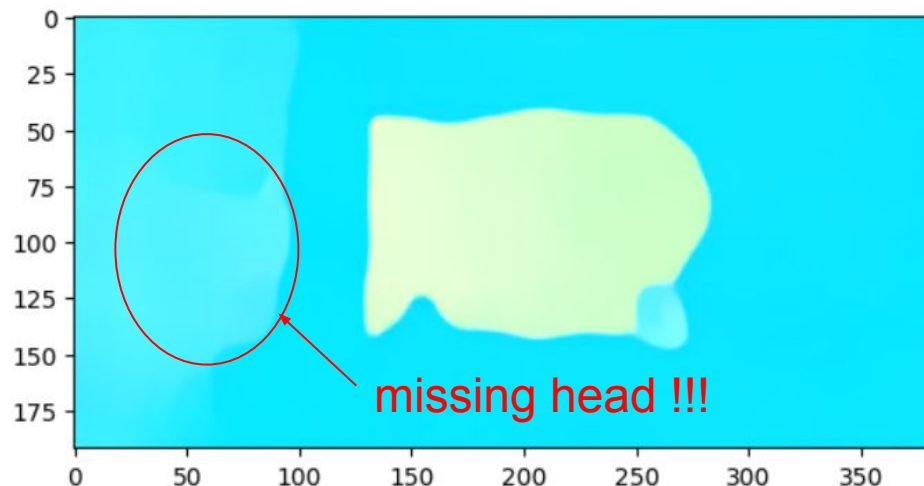
CONCLUSION



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

SECTION 5 :

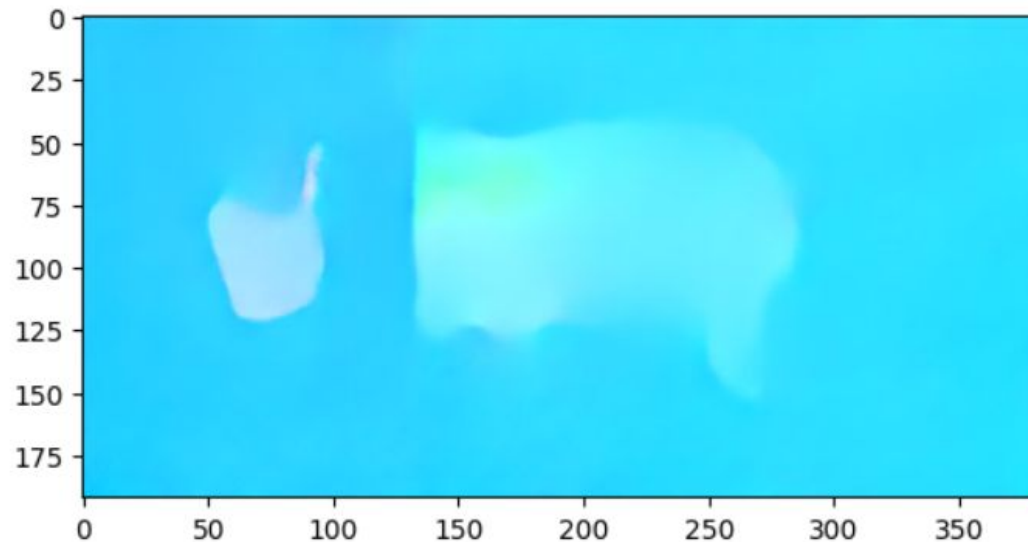
- Flownet 2 performs pretty well for large displacement (High accuracy between very far frames → good performance in direct integration)
- Our technique work well **only** on direct integration → Depend on the begin mask → Not too robust to occlusion/ when the object changes its pose to the camera
- Occlusion weakness may come from the model as well (relies on **local pixel information** and **motion consistency**)



Direct integration

SECTION 5 :

- However, by examining 2 consecutive frames, the model can also produce reliable results → We can try hybrid approach to see if it would be better or not



THANK FOR YOUR ATTENTION



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

APPENDIX



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

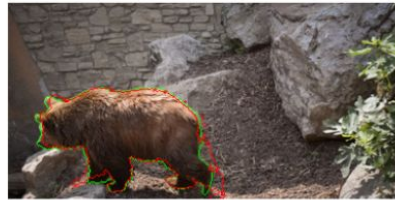
APPENDIX :

Results on Training sequence (Direct integration, bear sequence)

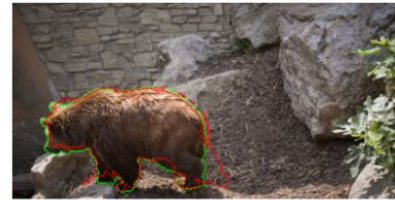
1



4



7



10



14



17



20

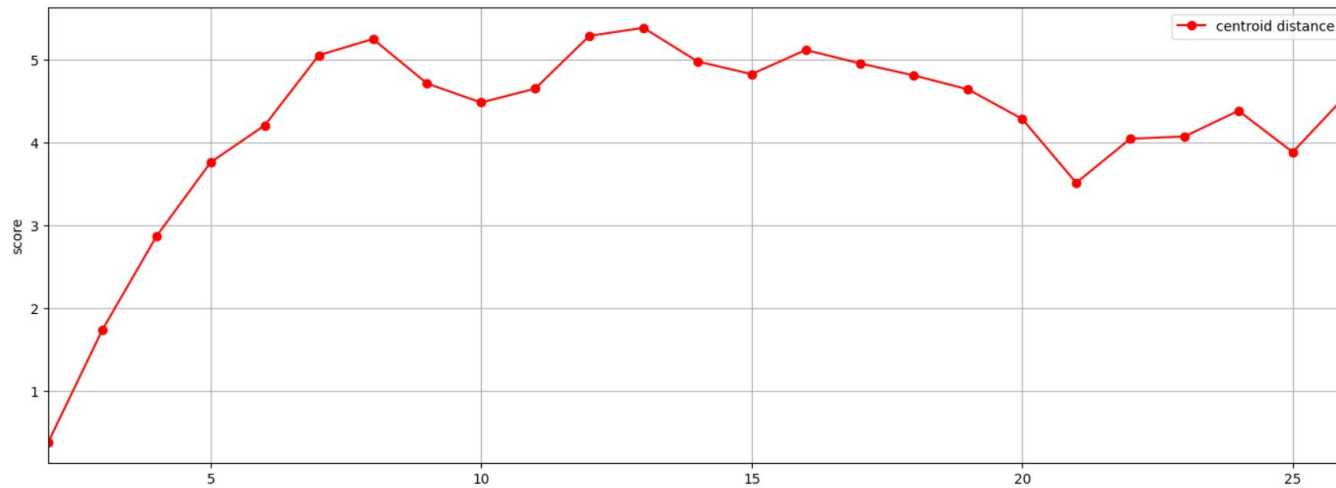
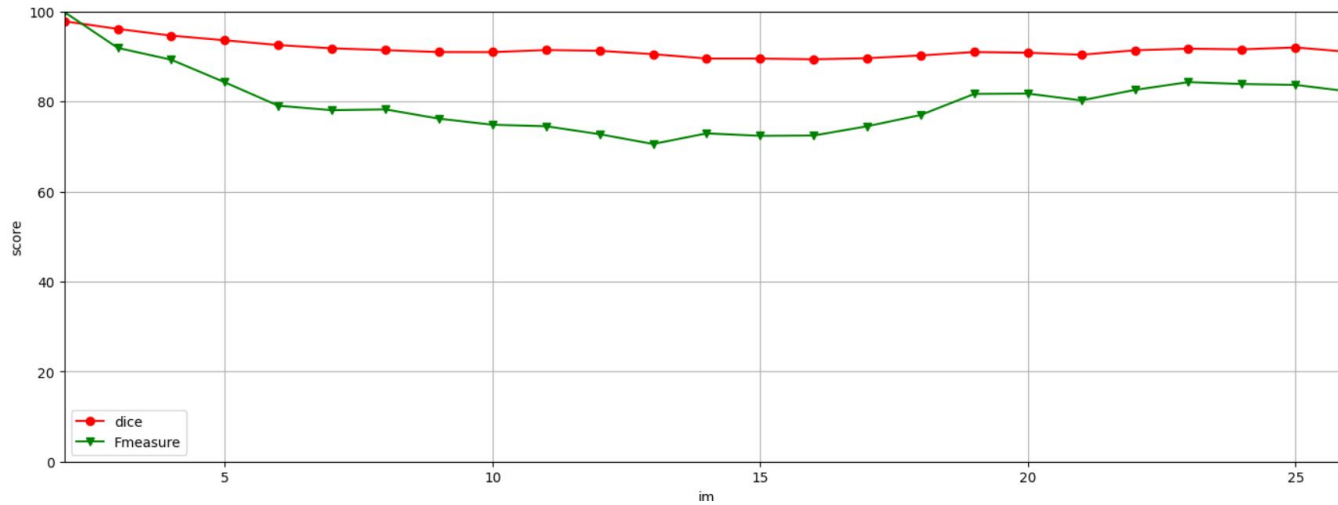


23



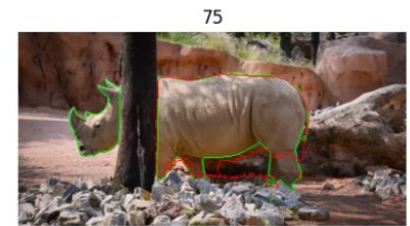
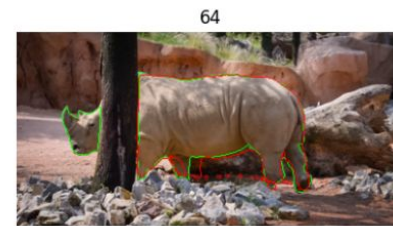
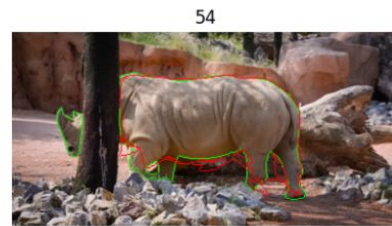
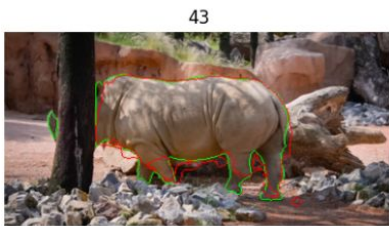
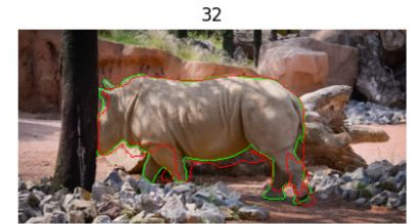
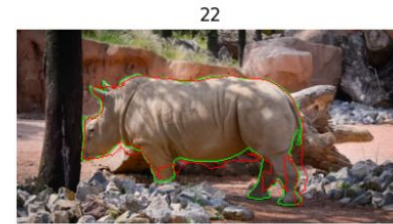
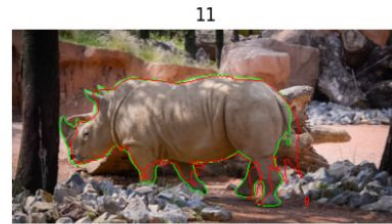
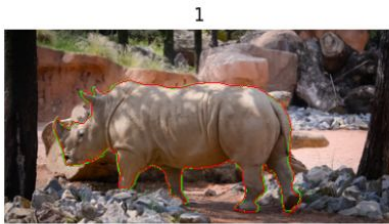
APPENDIX :

Results on Training sequence (Direct integration, bear sequence)



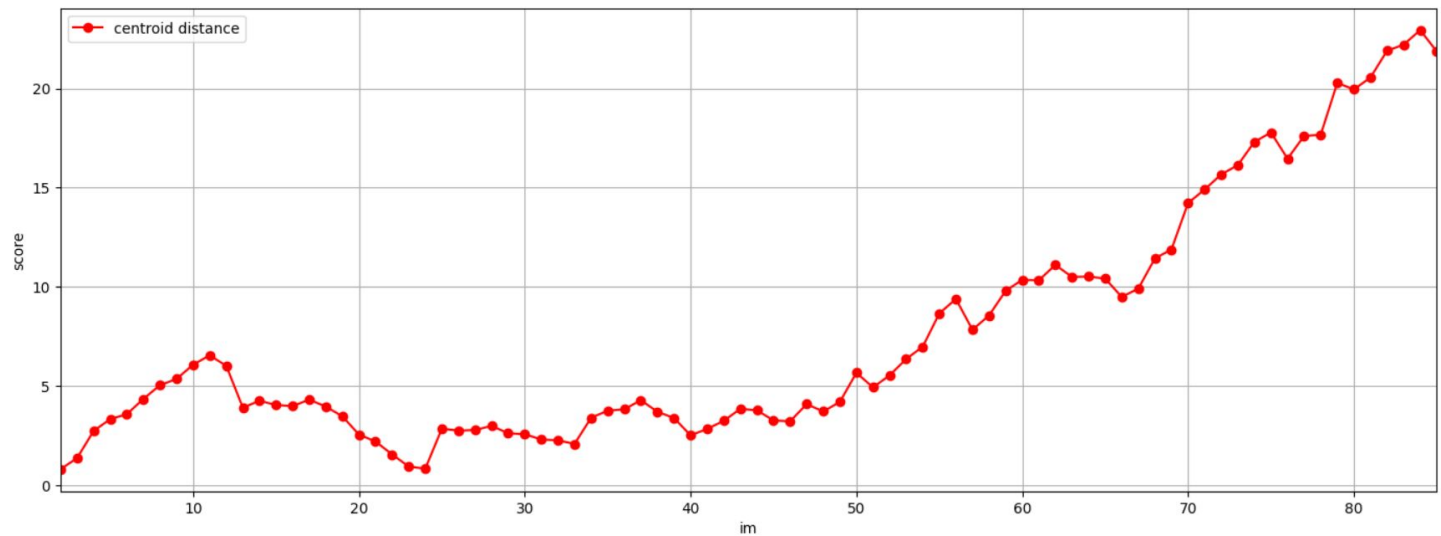
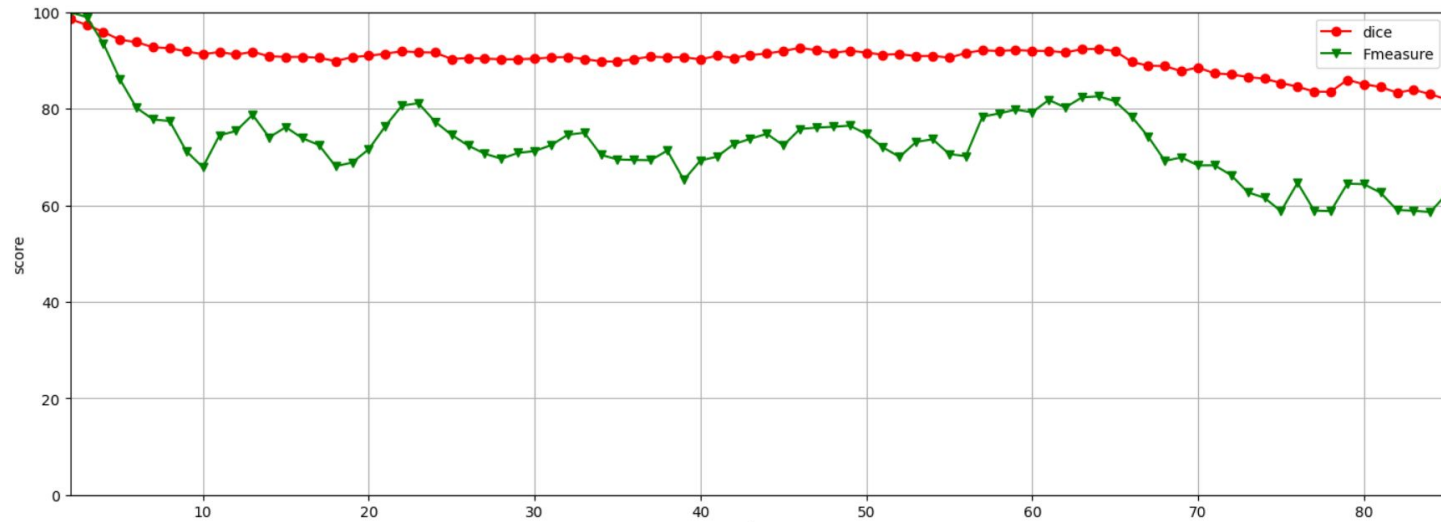
APPENDIX :

Results on Training sequence (Direct integration, rhino sequence)



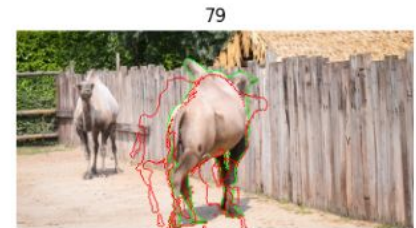
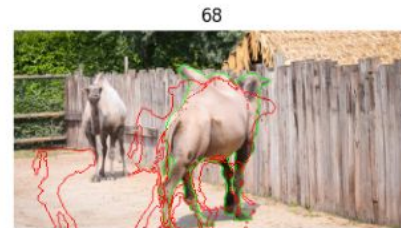
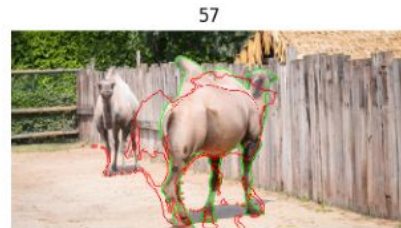
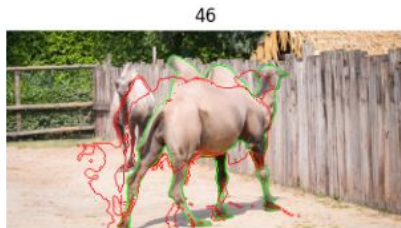
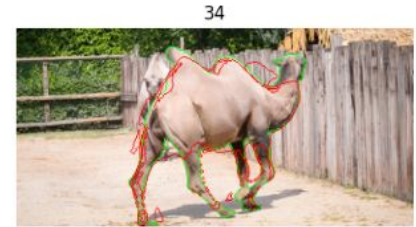
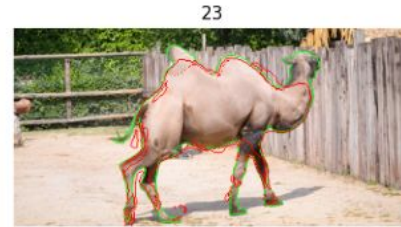
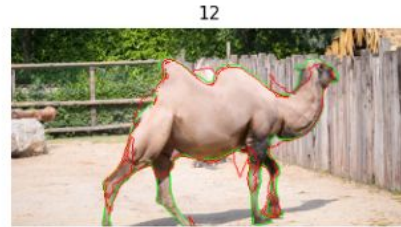
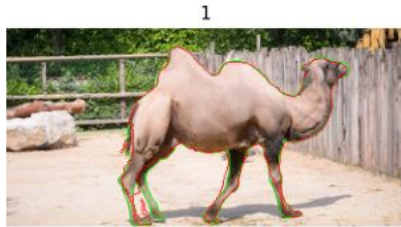
APPENDIX :

Results on Training sequence (Direct integration, rhino sequence)



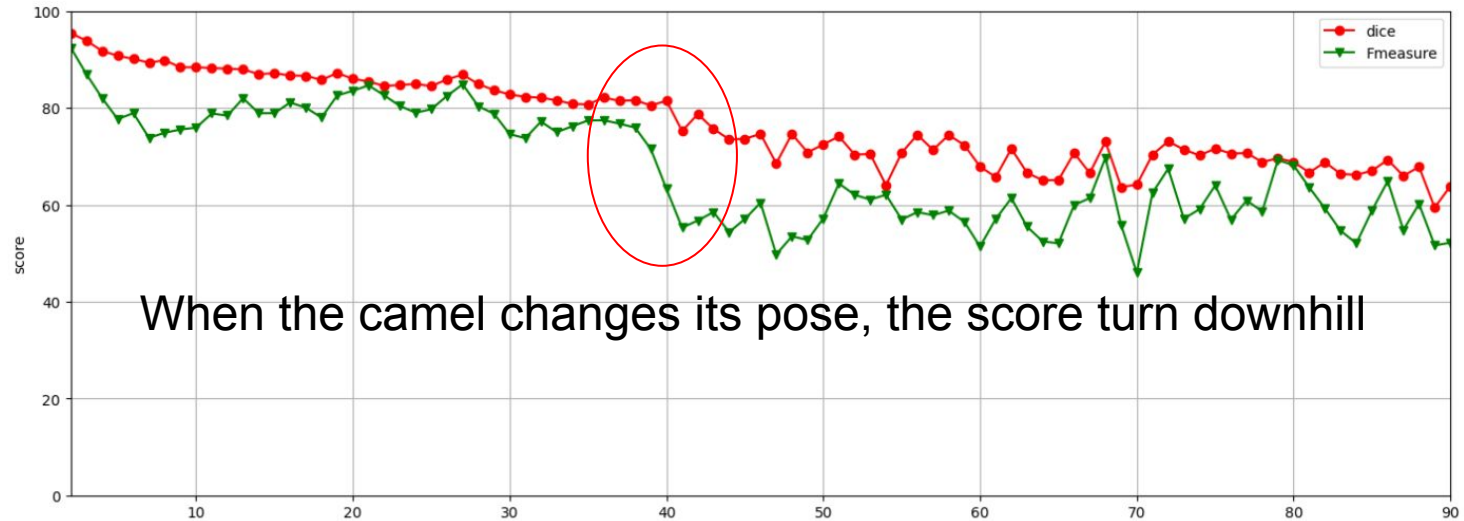
APPENDIX :

Results on Training sequence (Direct integration, camel sequence)



APPENDIX :

Results on Training sequence (Direct integration, camel sequence)



When the camel changes its pose, the score turn downhill

