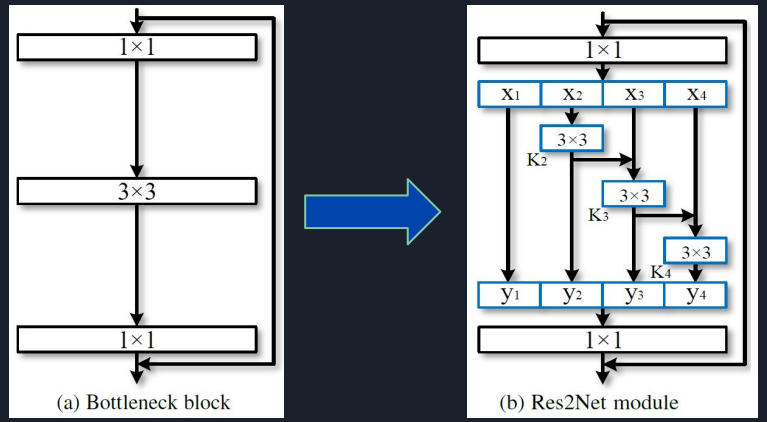
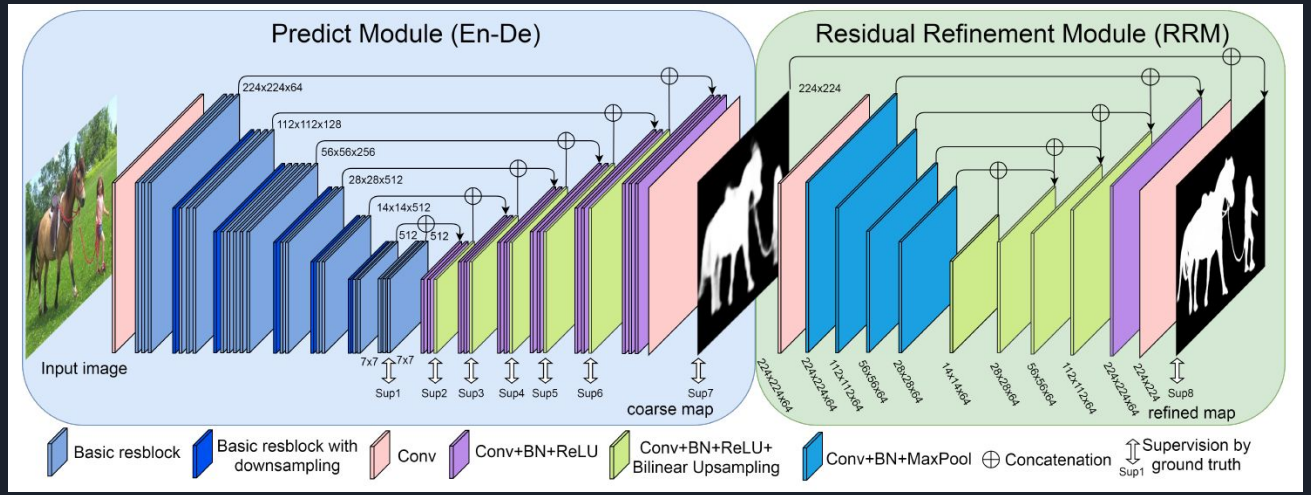
Salient object detection: **Bas2Net**

Irene Cannistraci 1603090
Giovanni Ficarra 1659089

Recap: Bas2Net

We completely rewritten the encoder with `Bottle2Neck` instead of `BasicBlock`, while the decoder is only reorganized, since it didn't use any `BasicBlock`.





The Dataset: DUTS

- DUTS is the dataset used for training and testing by the authors of BasNet's paper;
- It contains 10553 training images representing natural scenes, each one with its ground truth mask.

Evaluation Metrics

We chose the same metrics used in BasNet (Binary Segmentation Evaluation Tool) in order to compare our model with theirs:

1. *Precision-Recall curve*,
2. *Mean F-measure*,
3. *Max F-measure*,
4. *Mean Absolute Error*,
5. Binary Segmentation.

Experiment 1

Environment setup:

- Google Colab;
- GPU Tesla T4 15079MiB;
- PyTorch 1.5.0+cu101.

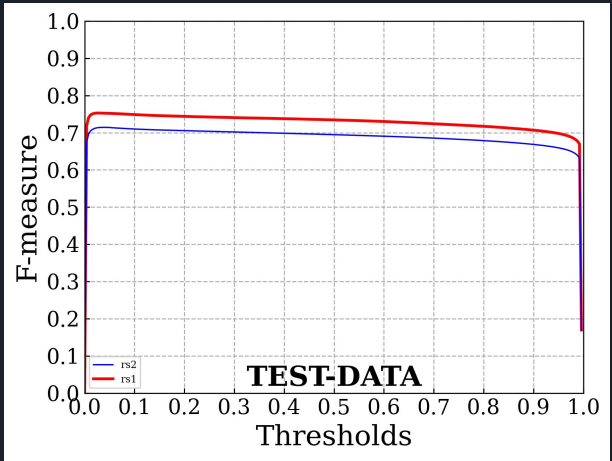
Configuration:

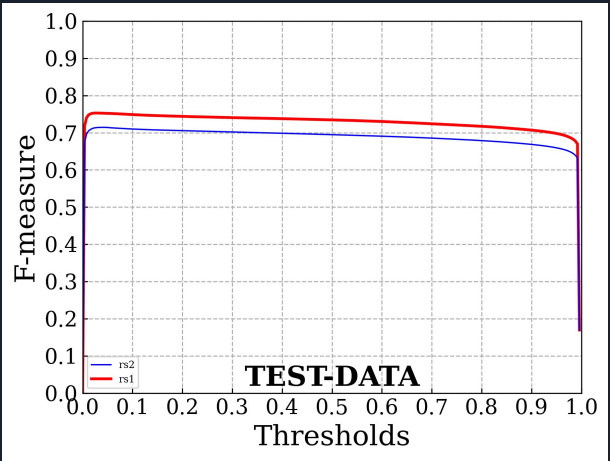
- 1000 training images;
- 1000 test images;
- 100 epochs;
- other hyper-parameters remained the same of BasNet's paper.

Experiment 1: quantitative results

- **BasNet**:
  - *Max F-score*: 0.753,
  - *Mean F-score*: 0.725,
  - *MAE*: 0.084,
  - *Time*: ~5h 12m.

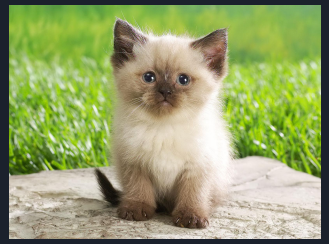
- **Bas2Net**:
  - *Max F-score*: 0.714,
  - *Mean F-score*: 0.687,
  - *MAE*: 0.101,
  - *Time*: ~3h 40m.





Experiment 1: qualitative results

Test images

BasNet results




Bas2Net results




Experiment 2

Environment setup:

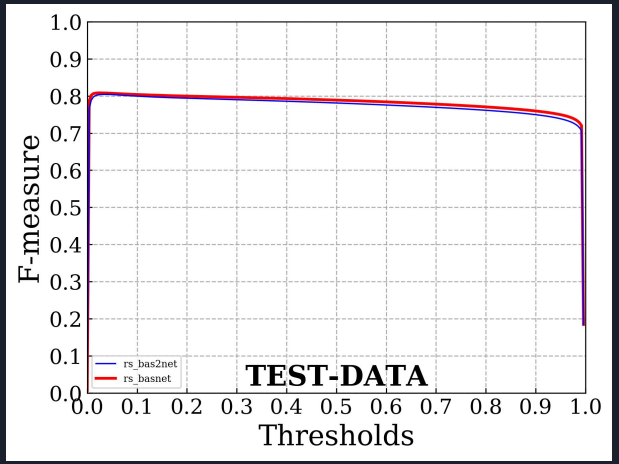
- Google Colab with GPU Tesla T4 15079MiB;
- Microsoft Azure nc6 VM with GPU Tesla K80;
- PyTorch 1.5.0+cu101.

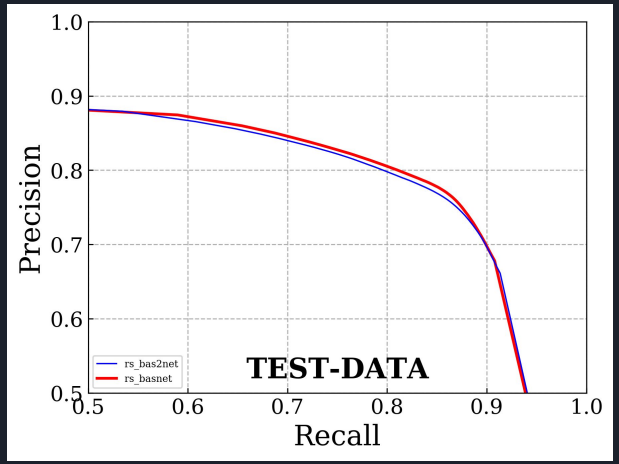
Configuration:

- 10553 training images;
- 5019 test images;
- 50 epochs;
- other hyper-parameters remained the same of BasNet's paper.

Experiment 2: results

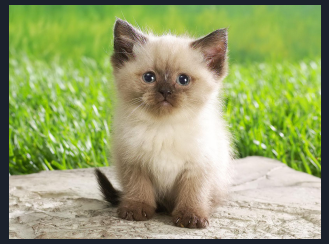
- **BasNet**:
  - *Max F-score*: 0.808,
  - *Mean F-score*: 0.779,
  - *MAE*: 0.068,
  - *Time*: ~66h 30m.
- **Bas2Net**:
  - *Max F-score*: 0.804,
  - *Mean F-score*: 0.772,
  - *MAE*: 0.070,
  - *Time*: ~40h 0m.





Experiment 2: qualitative results

Test images

BasNet results




Bas2Net results



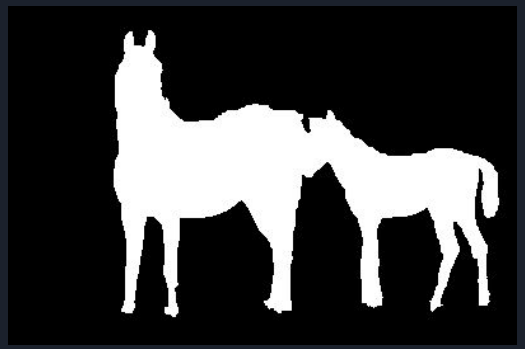

Experiment 3

Environment setup:

- Using Google Colab to execute test;
- Using the model produced in the *experiment 2*.

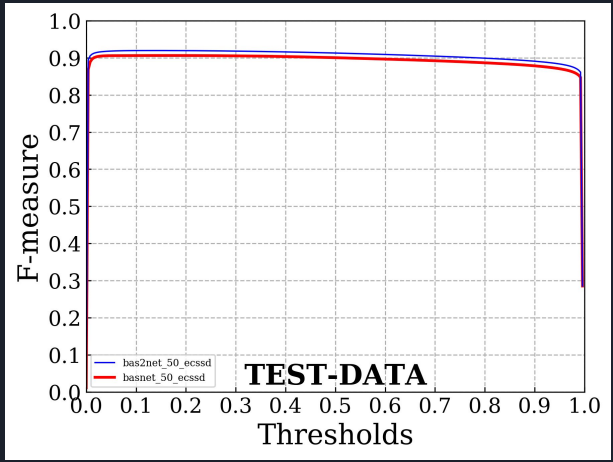
Configuration:

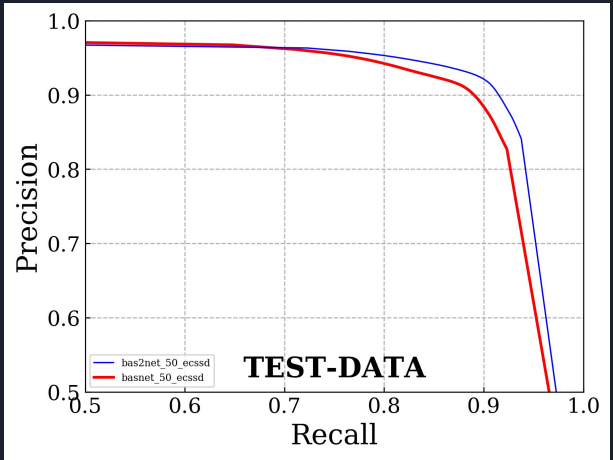
- 1000 test images from Extended Complex Scene Saliency Dataset (ECSSD)



Experiment 3: results

- **BasNet**:
  - *Max F-score*: 0.906,
  - *Mean F-score*: 0.890,
  - *MAE*: 0.055.
- **Bas2Net**:
  - *Max F-score*: 0.920,
  - *Mean F-score*: 0.903,
  - *MAE*: 0.049.





Experiment 3: qualitative results

Test images

BasNet results

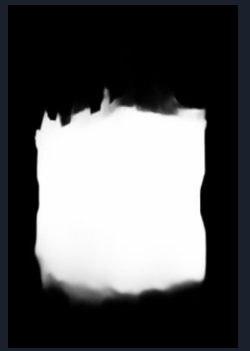
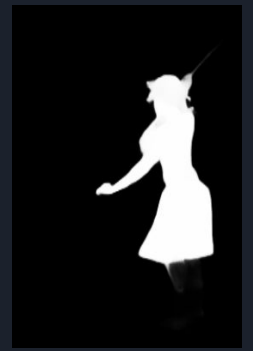
Bas2Net results



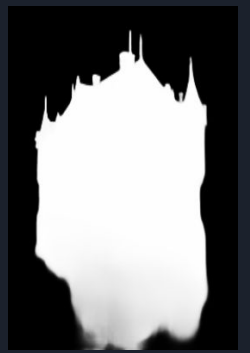
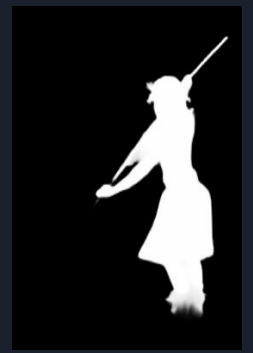








Experiment 4: WIP

Environment setup:

- Google Colab with GPU Tesla T4 15079MiB;
- Microsoft Azure nc6 VM with GPU Tesla K80;
- PyTorch 1.5.0+cu101.

Configuration:

- 10553 training images;
- 5019 test images;
- 100 epochs;
- other hyper-parameters remained the same of BasNet's paper.

Status:

- Currently at epoch 70.

Conclusions

- Bas2Net performances didn't improve with respect to BasNet on DUTS;
- Bas2Net performances improve with respect to BasNet on ECSSD;
- Bas2Net duration is ~50m for an epoch, while BasNet is ~1h30m:
  - we obtained a ~40% faster network;
- Bas2Net saved models occupy ~470MB, while BasNet models occupy ~1GB of memory when serialized:
  - we obtained a >50% save in terms of memory usage.

Future works

- Compare using other datasets;
- Repeat the experiment with different scales;
- Compare Bas2Net with PoolNet and Pool2Net;
- Implement Transfer Learning from Res2Net in order to see if performances get better.

Thank you!
Questions?