TWO-STAGE SEGMENTATION IN NEURAL NETWORKS

Diana Teixeira e Silva
up201805131@fe.up.pt

Ricardo Cruz
rpcruz@fe.up.pt

Tiago Gonçalves
tiago.f.goncalves@inesctec.pt

Diogo Carneiro
Diogo.Carneiro@pt.bosch.com

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

The 15th International Conference on Machine Vision (ICMV 2022)
November 18-20, Rome, Italy

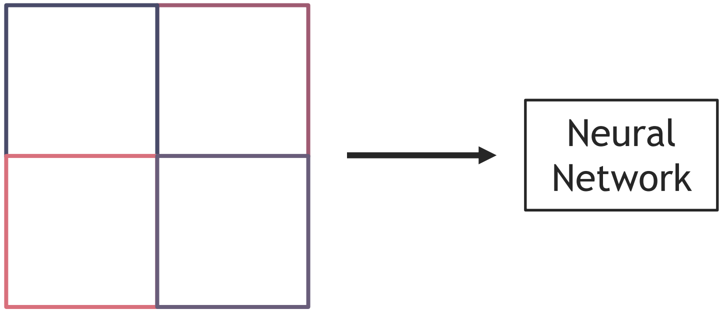
INESCTEC

Presentation Outline

Context

Segmentation of high-resolution images

- ⊗ Time Consuming
- ⊗ Requires a lot of memory

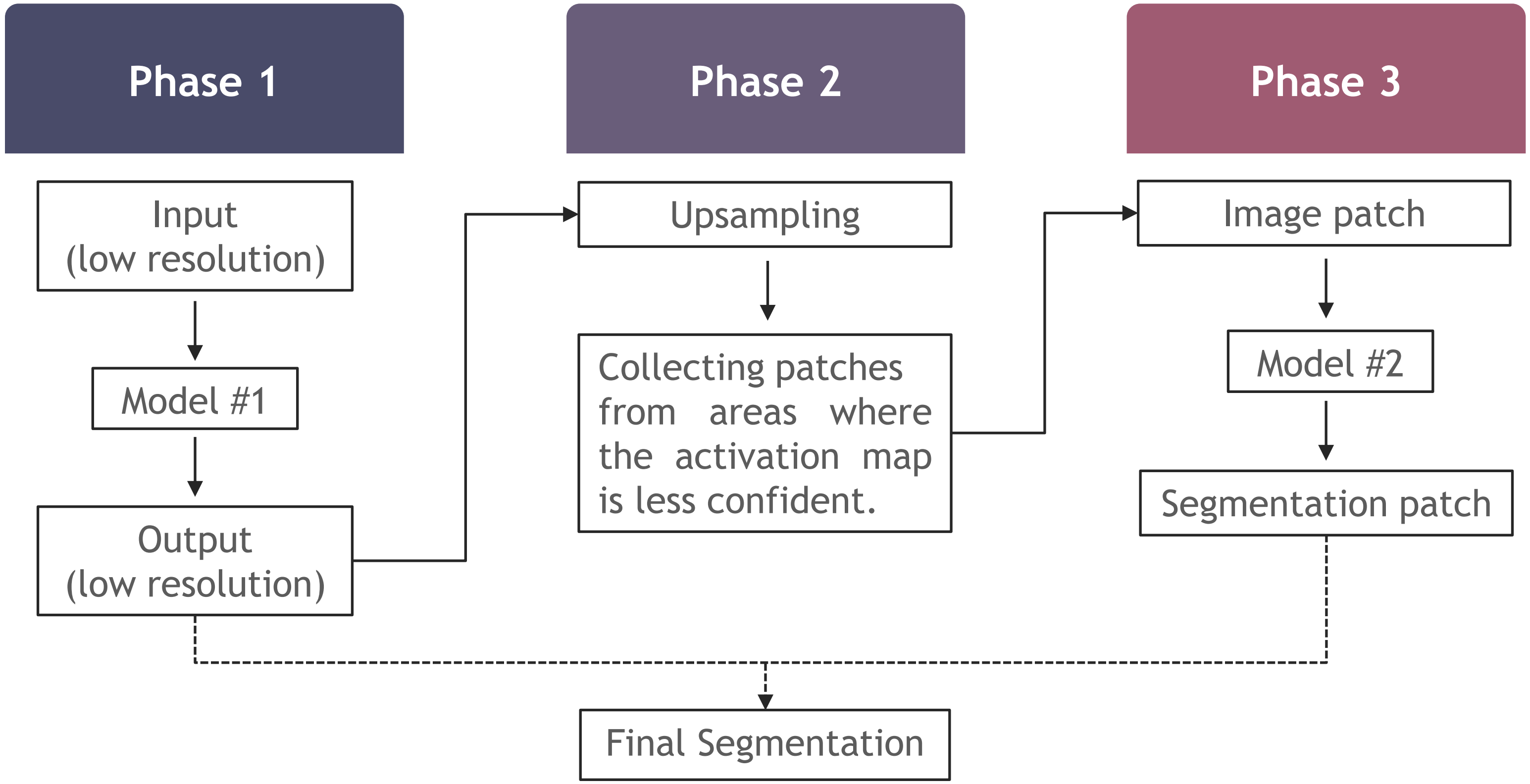


Segmentation of individual patches

- ⊗ Equal time spent in hard-to- and easy-to-segment regions
- ⊗ The boundaries of patches

Iterative Segmentation

Iterative Segmentation



Datasets

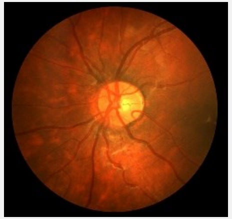
Dataset: Average Resolution → $hi_{size}{}^2$

BIOMEDICAL DATA

Dermoscopic images

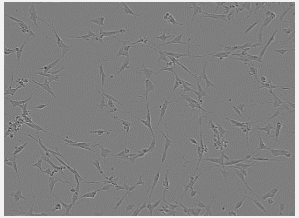
PH2: 575×766→ 768^2

Retinal fundus images

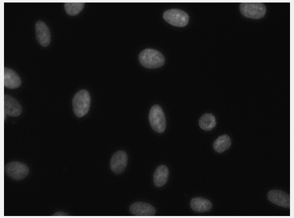


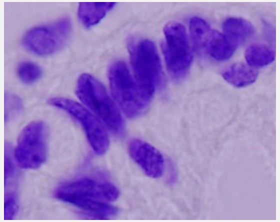
RETINA*: 745×782 → 768^2

Microscopic images



SARTORIUS:
520×704 → 512^2





BOWL2018:
328×369 → 256^2

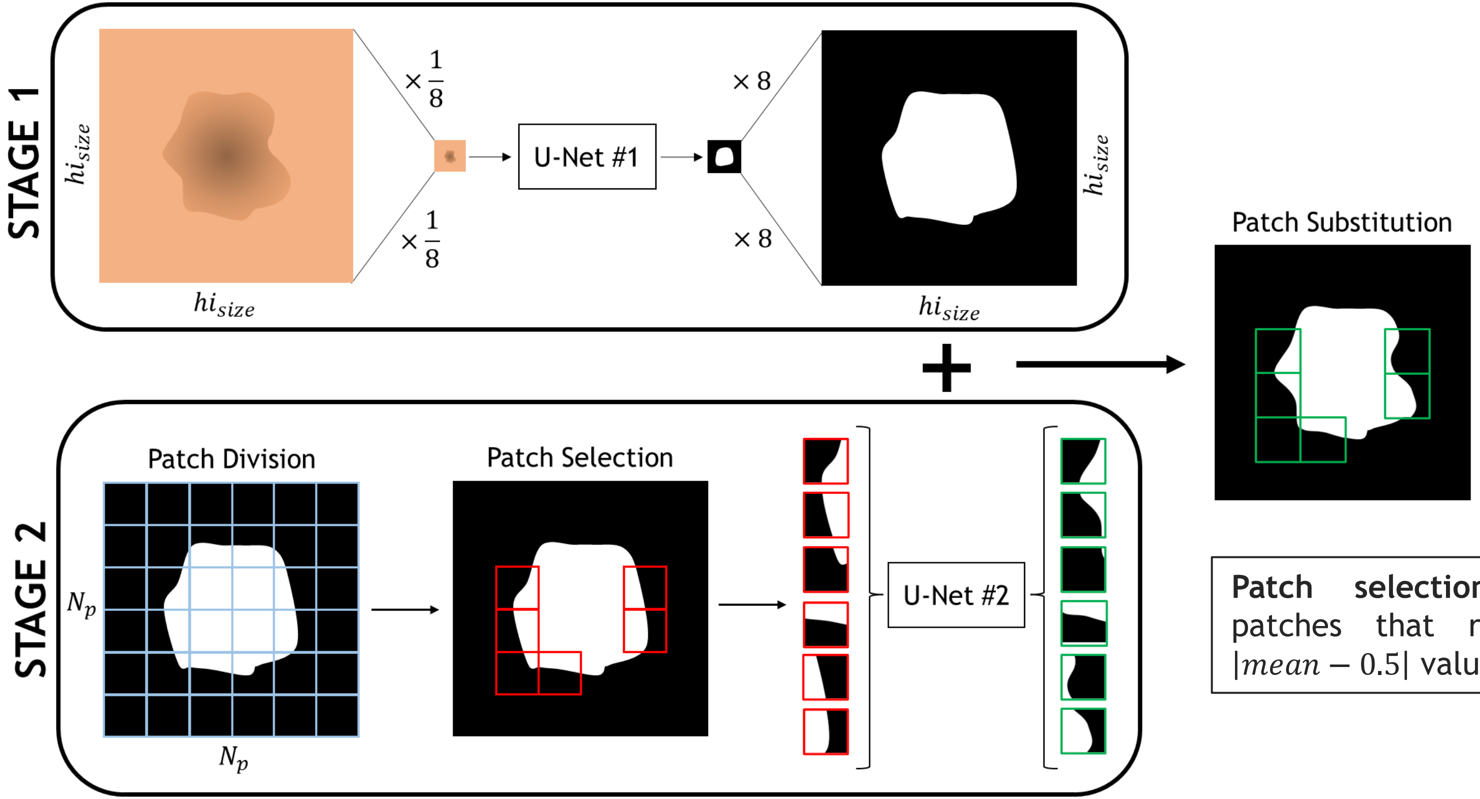
AUTONOMOUS DRIVING

KITTI: 375×1271 → 512^2

BDD100K: 720×1280 → 768^2

* Composition of three datasets: CHASE_DB1, DRIVE and STARE

Two-Stage Segmentation



Patch selection criteria: Np patches that reported lowest $|mean - 0.5|$ value.

Training

Data Augmentation

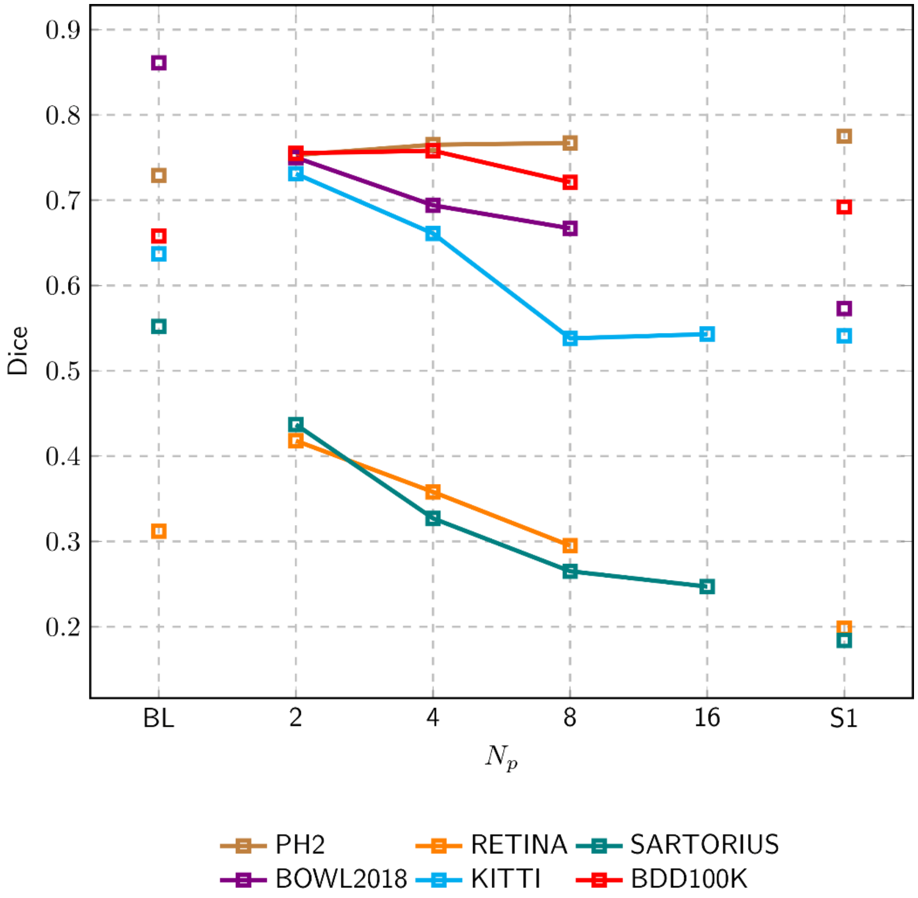
1. Horizontal flip;
2. Contrast/Brightness modification;
3. Random rotation (only applied to the biomedical datasets).

	Training	
	U-Net #1 (low res)	U-Net #2 (patches)
Resizing	$\left(\frac{1}{8} + 5\%\right) \cdot hi_{size}$	hi_{size}
Random Cropping	$\frac{1}{8} \cdot hi_{size}$	$\frac{1}{Np} \cdot hi_{size}$, $Np = \{2,4,8,16\}$
Loss Function	$\mathcal{L}(y, \hat{y}) = \mathcal{L}_f\ (y, \hat{y}) + (1 - D(y, \hat{y}))$, where $\mathcal{L}_f$ is the focal loss and $D(y, \hat{y})$ is the Dice coefficient	
Learning Rate	0.0001	
Epochs	200	1000
Batch Size	64	

Results

3 different evolutions:

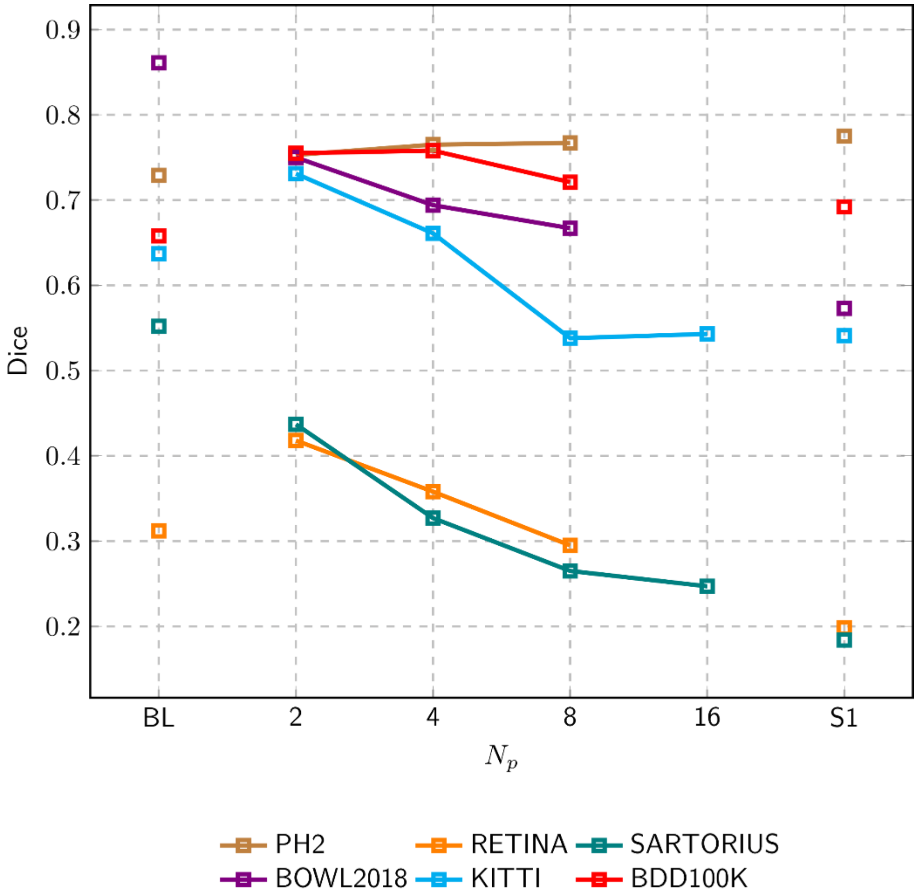
1. PH2 and BDD100K
2. SARTORIUS and BOWL2018
3. KITTI and RETINA



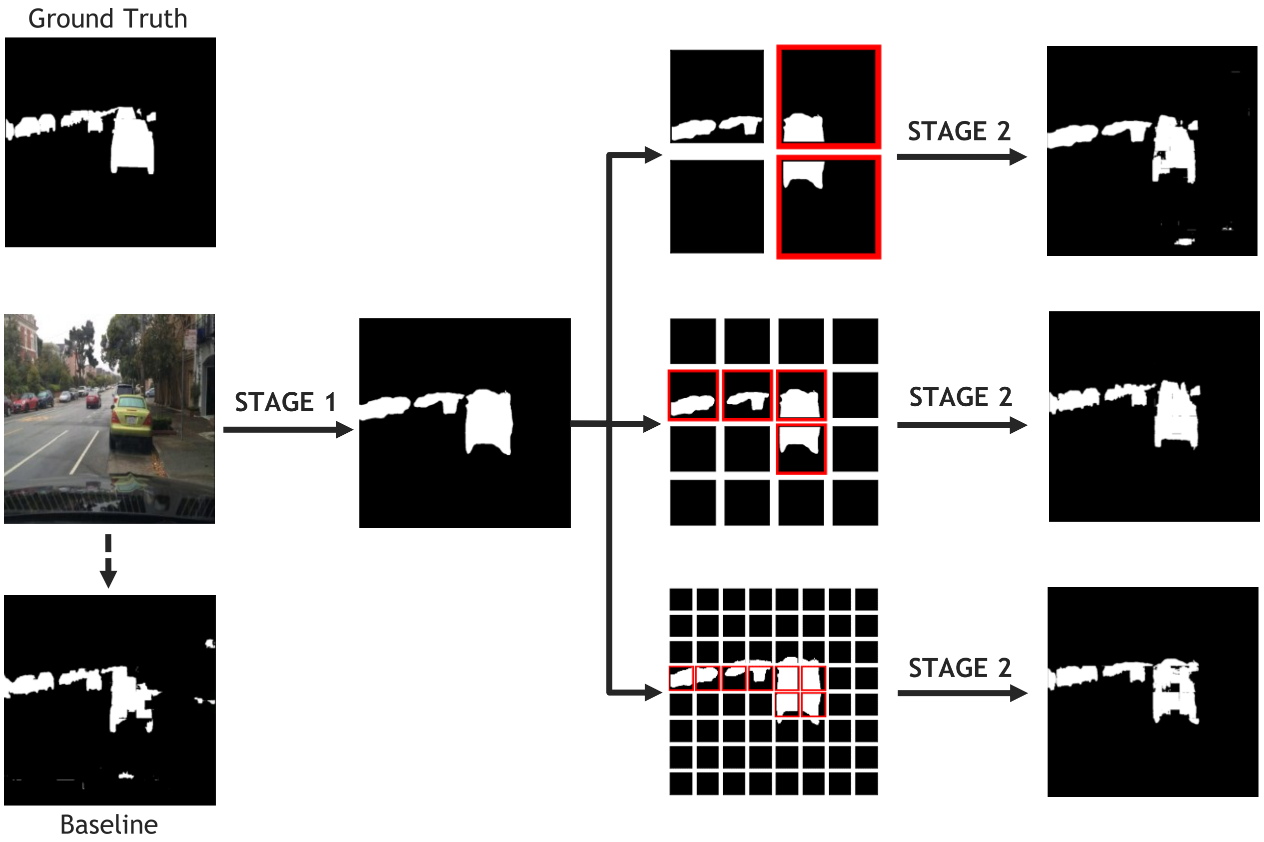
Results

1 PH2 and BDD100K

- PH2 registers the best Dice coefficient in Stage 1 (S1);
- BDD100K registers the best Dice coefficient for Np=4 (Stage 2 – S2);
- All stages from our method register a higher Dice coefficient when compared to the baseline (BL).



Results

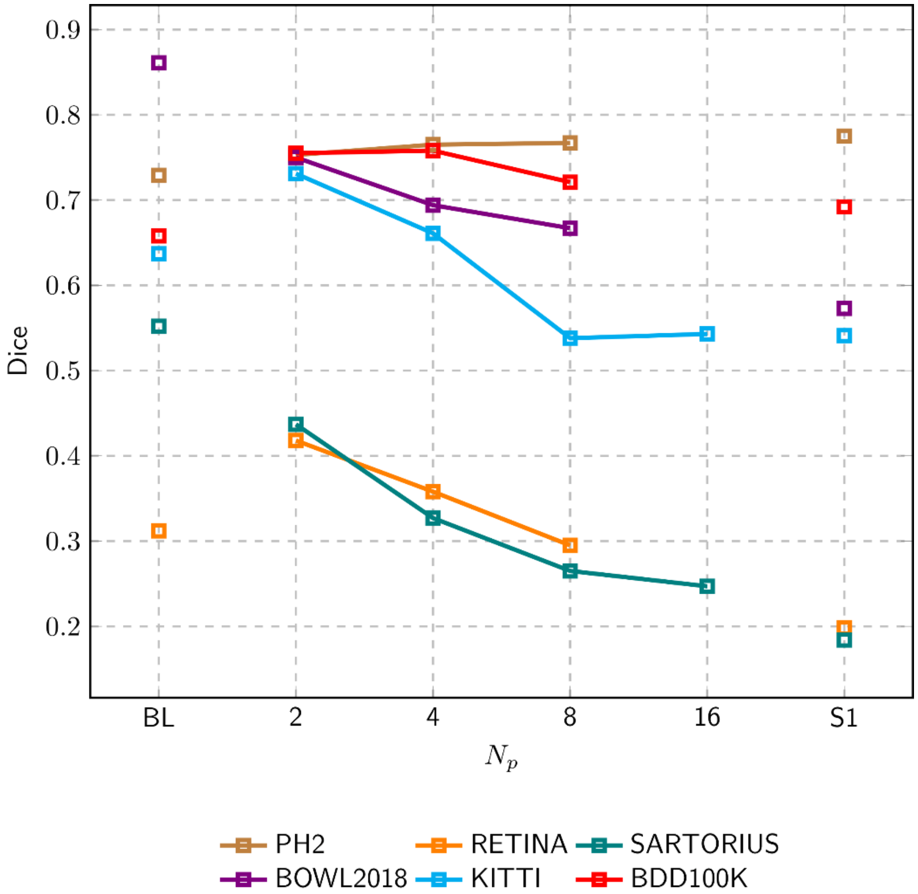


Example 1: Segmentation results for an image from the BDD100K dataset.

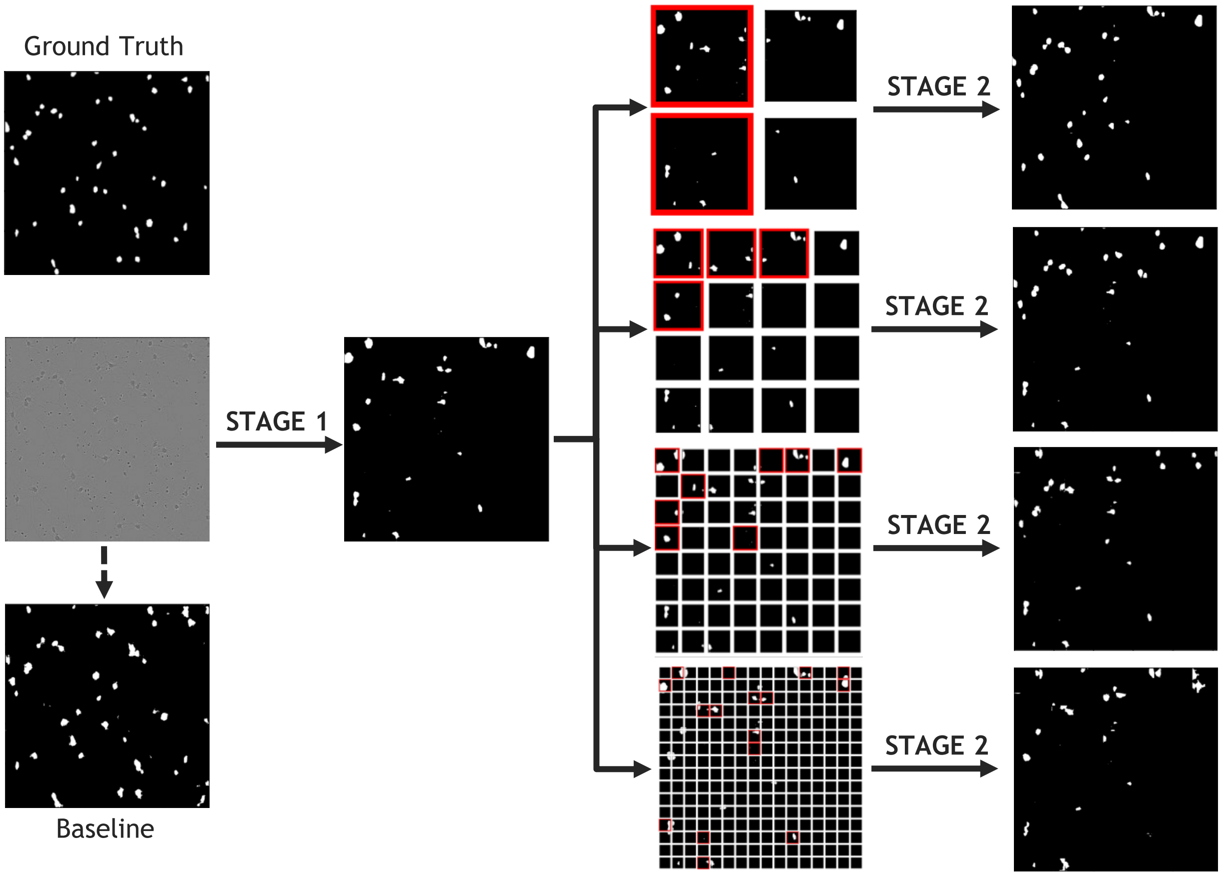
Results

2 SARTORIUS and BOWL2018

- BL registers the best result;
- Poor results from S1 compromise S2 result.



Results

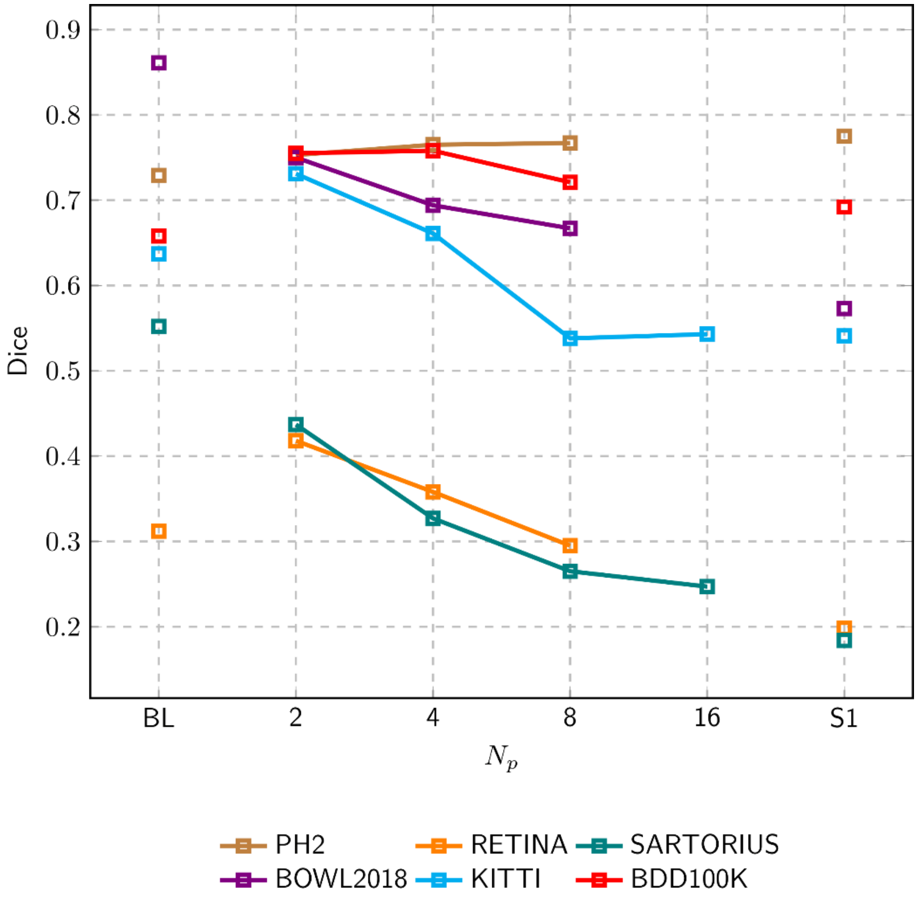


Example 2: Segmentation results for an image from the SARTORIUS dataset.

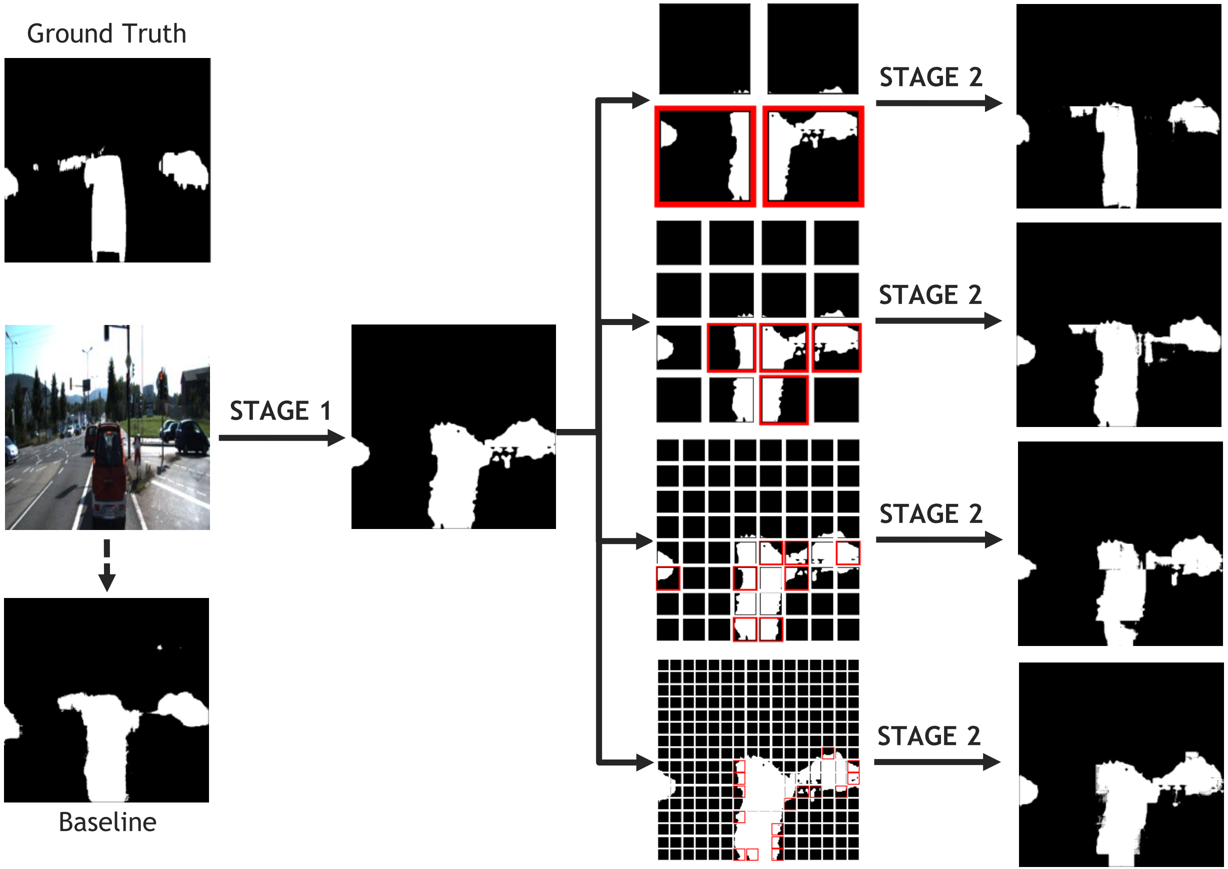
Results

3 KITTI and RETINA

- Our method performs better than the baseline only for the bigger patches (Np={2,4}), worsening for smaller patches;



Results



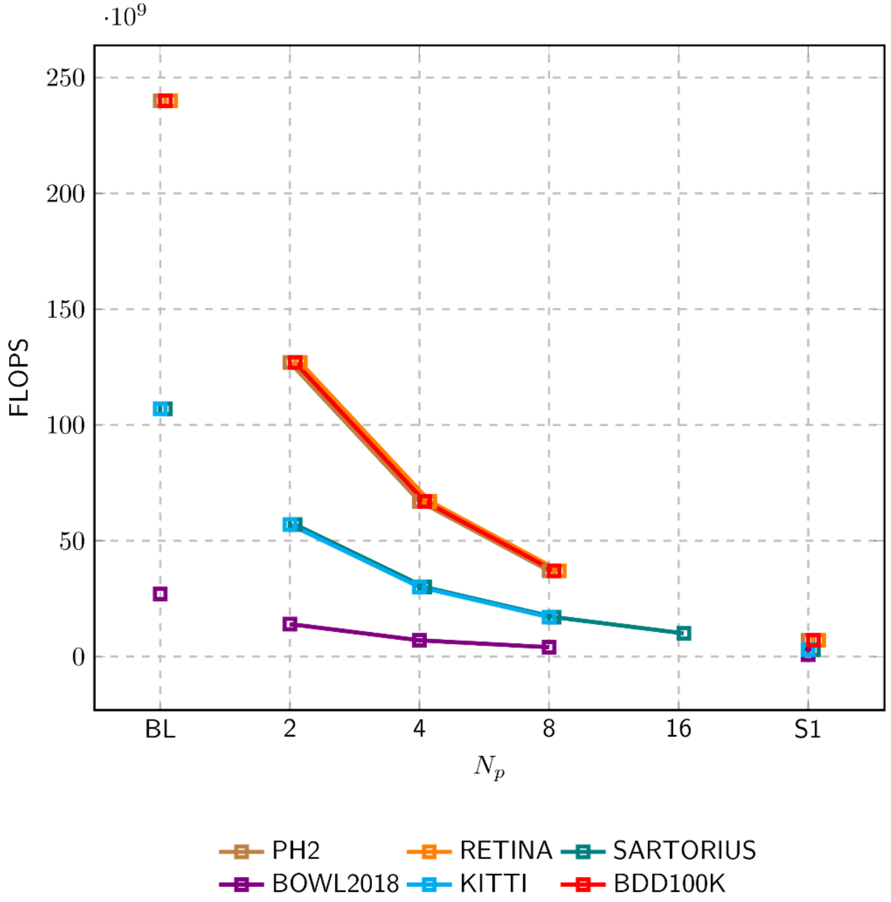
Example 3: Segmentation results for an image from the KITTI dataset.

Results

- Increasing the number of patches, the number of operations reduces;

- In all cases, our method requires fewer operations than the baseline.

Saves at least 50% and up to 80% of the total number of operations.



Conclusion and Future Work

Compared to the baseline model, our method showed:

1. Generally similar Dice coefficients;
2. Fewer arithmetic operations.

Use datasets with higher resolution;

Use a fine-grain sliding window.

Identify regions where the neural network is undecided and use these regions with the same scale (autonomous driving datasets).

Thank you for your attention.

TWO-STAGE SEGMENTATION IN NEURAL NETWORKS

Diana Teixeira e Silva
up201805131@fe.up.pt

Ricardo Cruz
rpcruz@fe.up.pt

Tiago Gonçalves
tiago.f.goncalves@inesctec.pt

Diogo Carneiro
Diogo.Carneiro@pt.bosch.com

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

The 15th International Conference on Machine Vision (ICMV 2022)
November 18-20, Rome, Italy

INESCTEC