Fusion of Heterogeneous Data in Convolutional Networks for Urban Semantic Labeling

Nicolas Audebert[1,2], Bertrand Le Saux[1], Sébastien Lefèvre[2]

[1] ONERA, *The French Aerospace Lab*
[2] Université de Bretagne-Sud/IRISA

March 6, 2017

ONERA
THE FRENCH AEROSPACE LAB

UMR IRISA

TOTAL

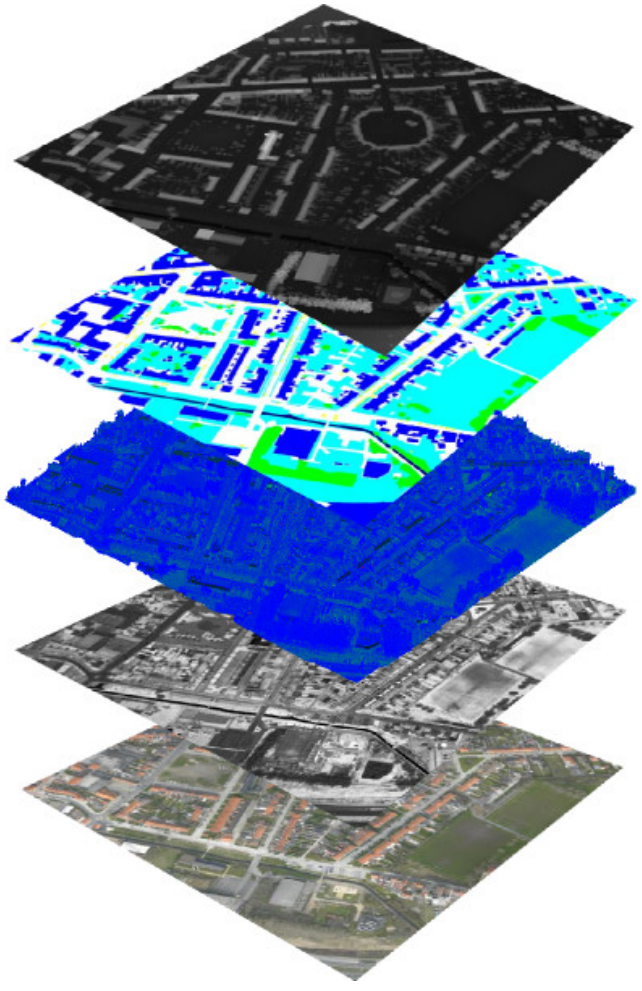
Earth Observation & semantic mapping

Multimodal semantic segmentation with deep nets

Conclusion

Earth Observation & semantic mapping

Earth Observation

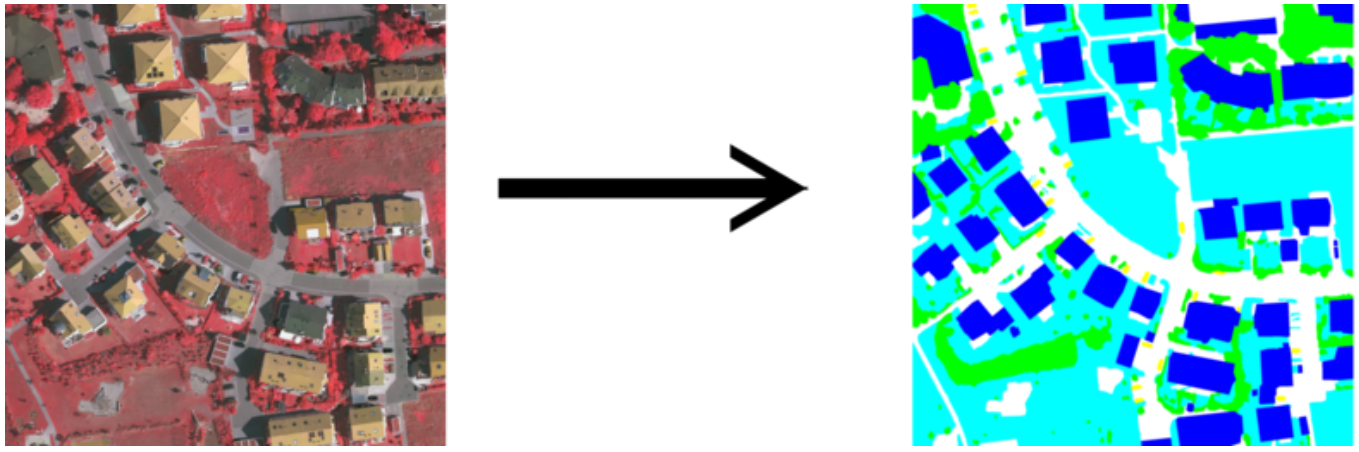


Remote sensing data

Remote sensing data has grown abundant over the years

- Satellite and airborne images...
- with auxiliary products (e.g. DSM)
- ...and annotations !

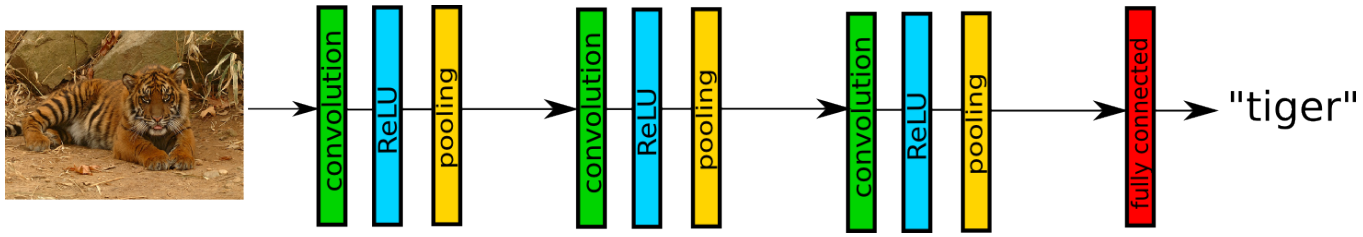
Semantic mapping



Build meaningful maps from data

- **Dense classification** of every pixel
- **Thematic maps**: vegetation type, building/roads extraction, ...
- First step for further analysis: urban planning, biomass estimation, traffic analysis...

Deep learning

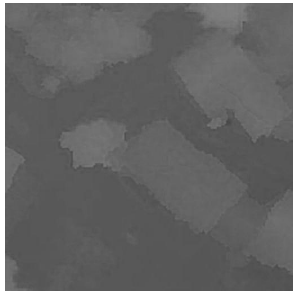


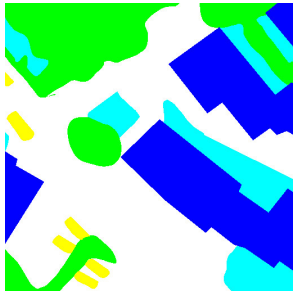
Why deep learning ?

- **Convolutional neural networks** perform significantly better than any previous methods for most vision-based tasks (classification, segmentation, ...)
- Computing power (thanks to GPUs) is now cheap enough to train **very deep models on huge datasets**
- **Annotated remote sensing data** is largely available for supervised learning

Multimodal semantic segmentation with deep nets

Dataset

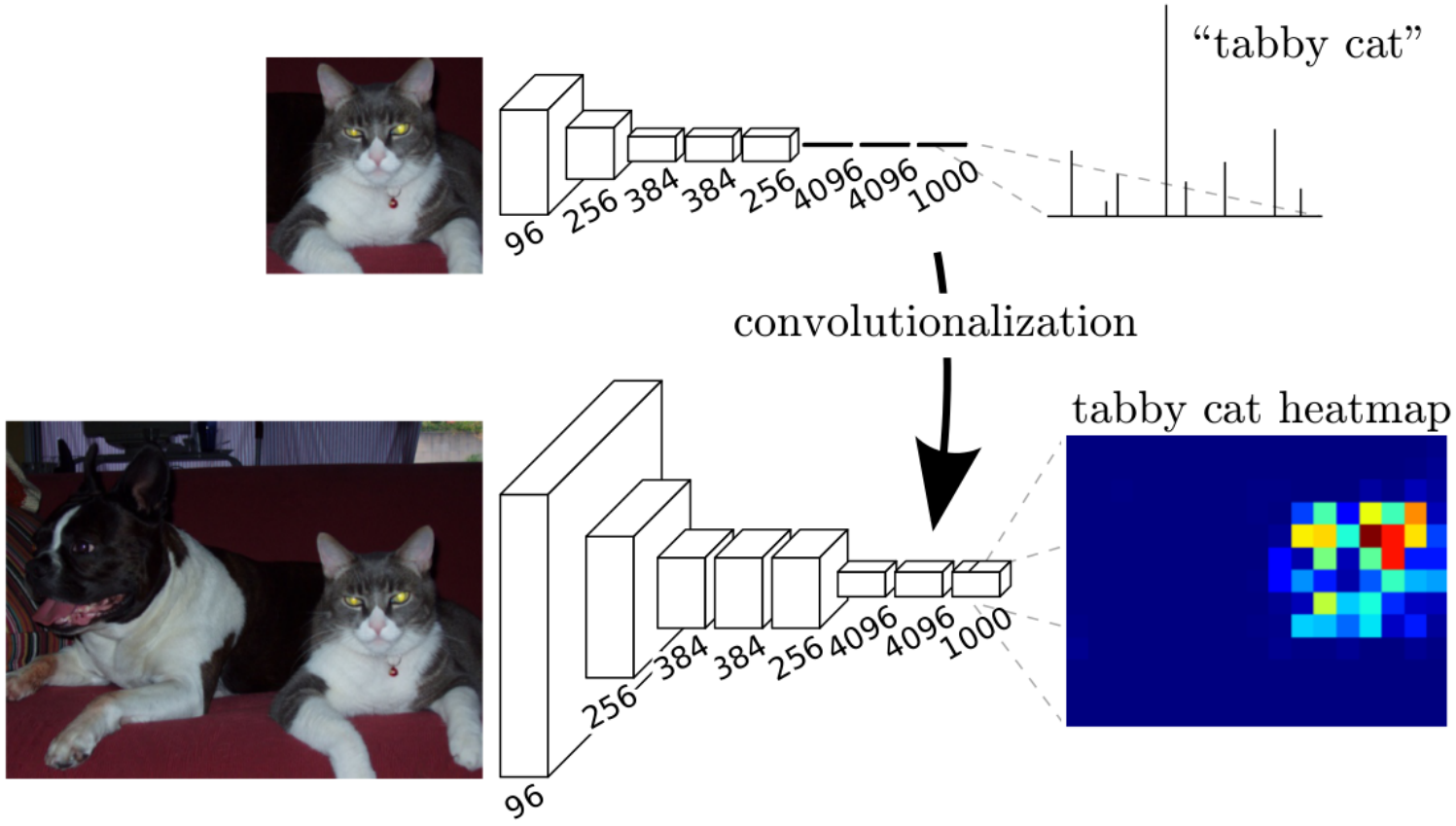




ISPRS Vaihingen

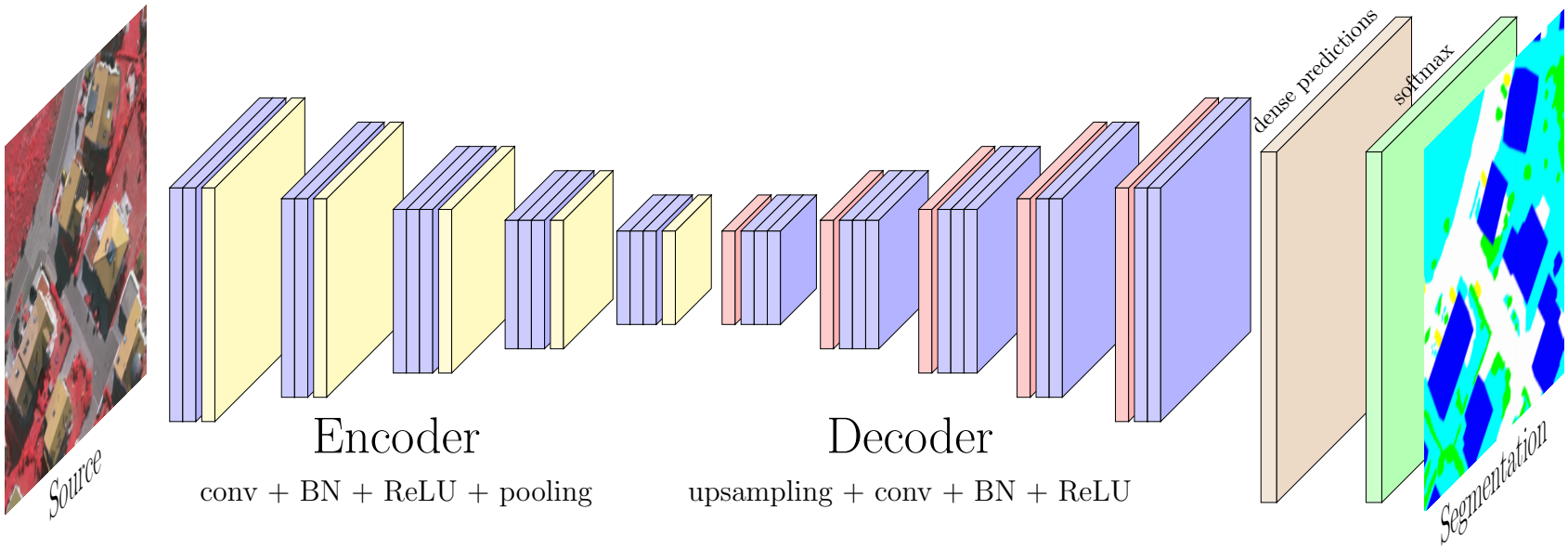
- 9 cm/pixel resolution
- IRRG ortho-rectified images
- Lidar point cloud
  - DSM
  - NDSM
- Dense pixel-wise annotations
  - roads, buildings, low vegetation, trees, vehicles, clutter

Traditional semantic segmentation



Fully Convolutional Networks for Semantic Segmentation, Long et al., CVPR'15.

Traditional semantic segmentation



SegNet, Badrinarayanan et al., 2015.

SegNet for remote sensing data

Preprocessing

HR images are processed by a 128×128 sliding window with a stride of 32*px* (75% overlap).

- \+ Lower GPU memory consumption
- \+ Data augmentation (*training*)
- \+ Averaging on the overlapping pixels (*testing*)
- − Longer processing time

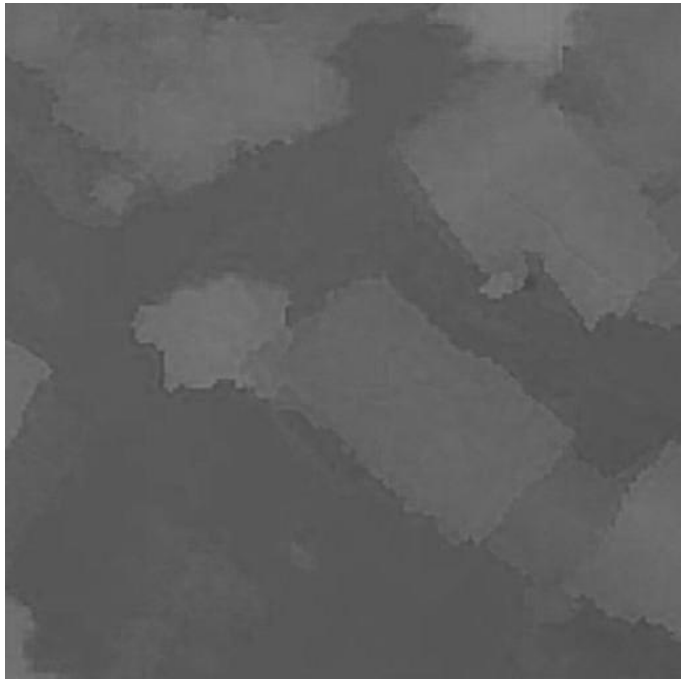
Optimization

SGD with momentum and standard backpropagation

- **Log-loss averaged** on the patch pixels (no spatial regularization)

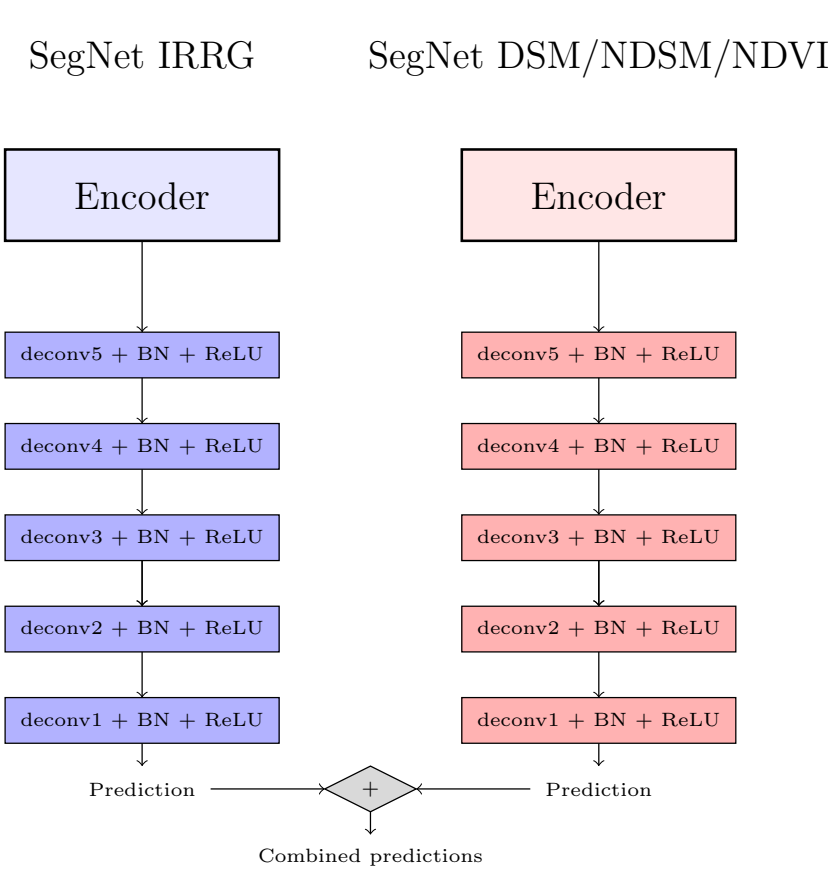
Problem

We have optical data (IRRG) **and** Lidar data (DSM/NDSM). How can we use all this information together ?

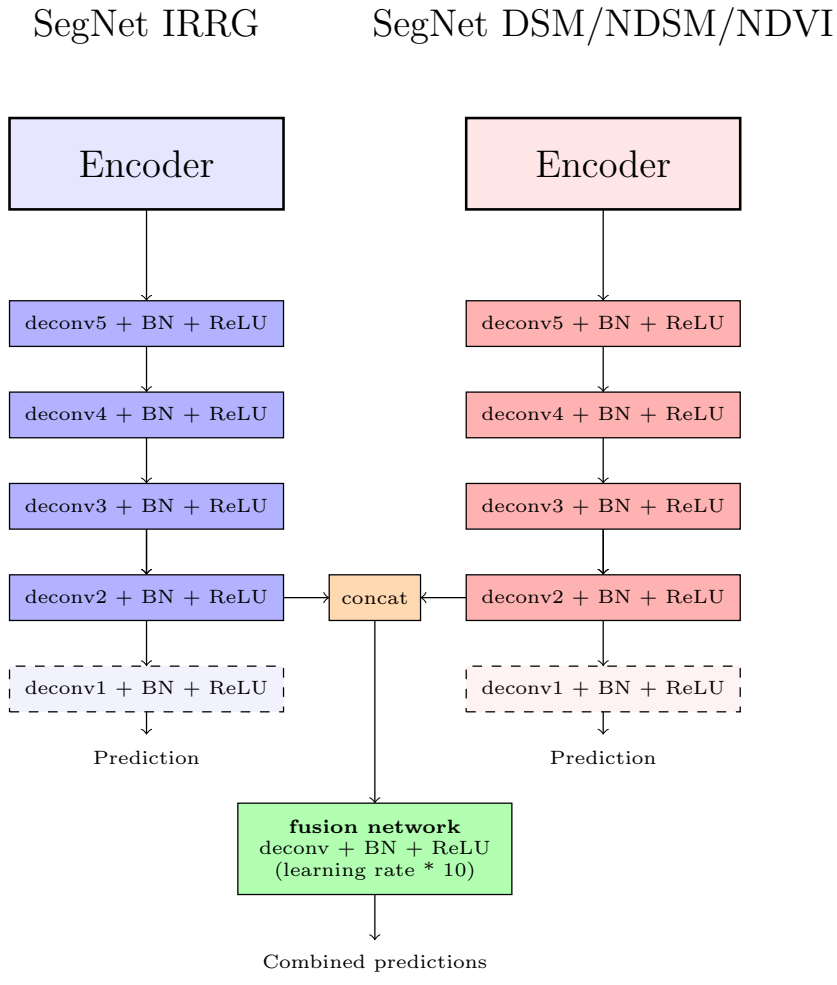


NDSM helps discriminate between roads/buildings, low vegetation/trees...

Multimodal semantic segmentation



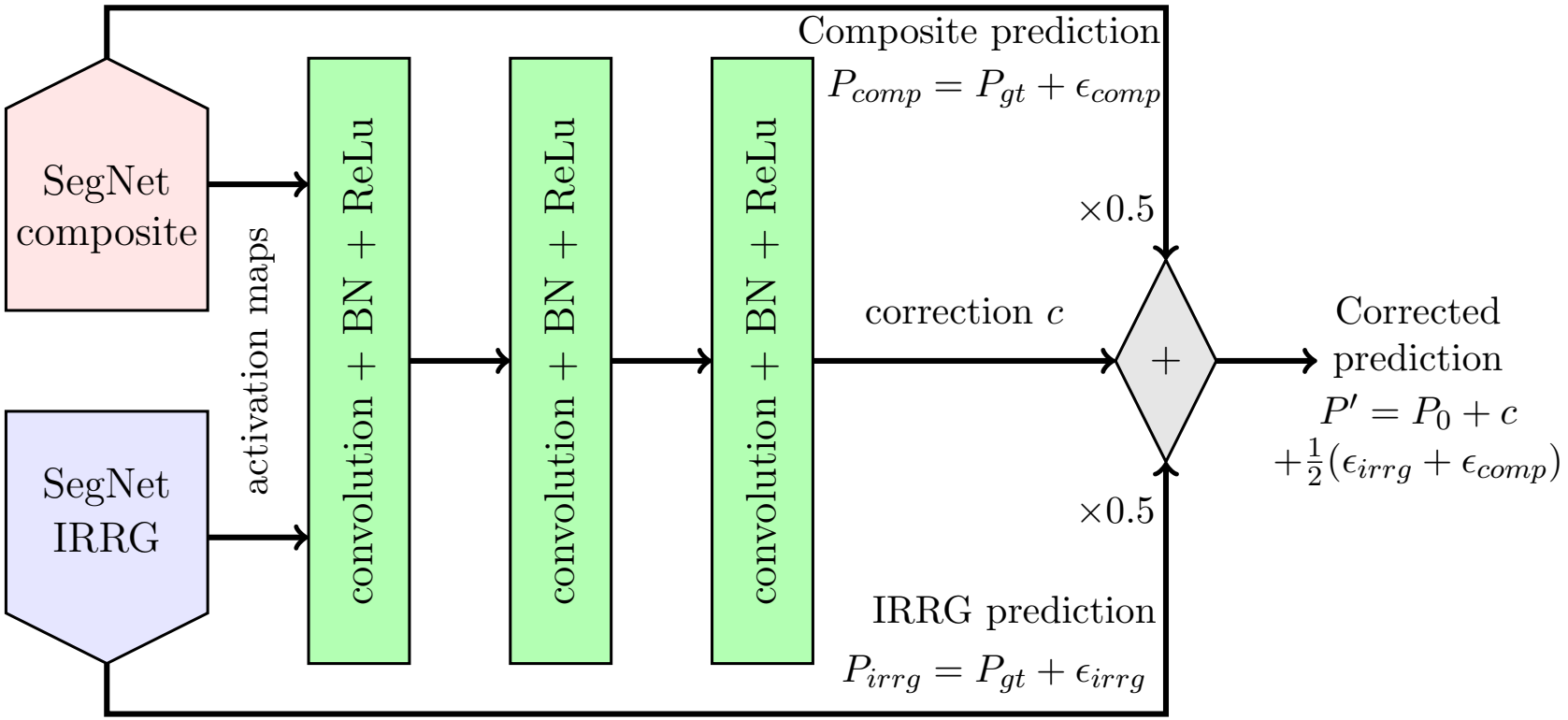
Naive fusion: averaging the two predictions



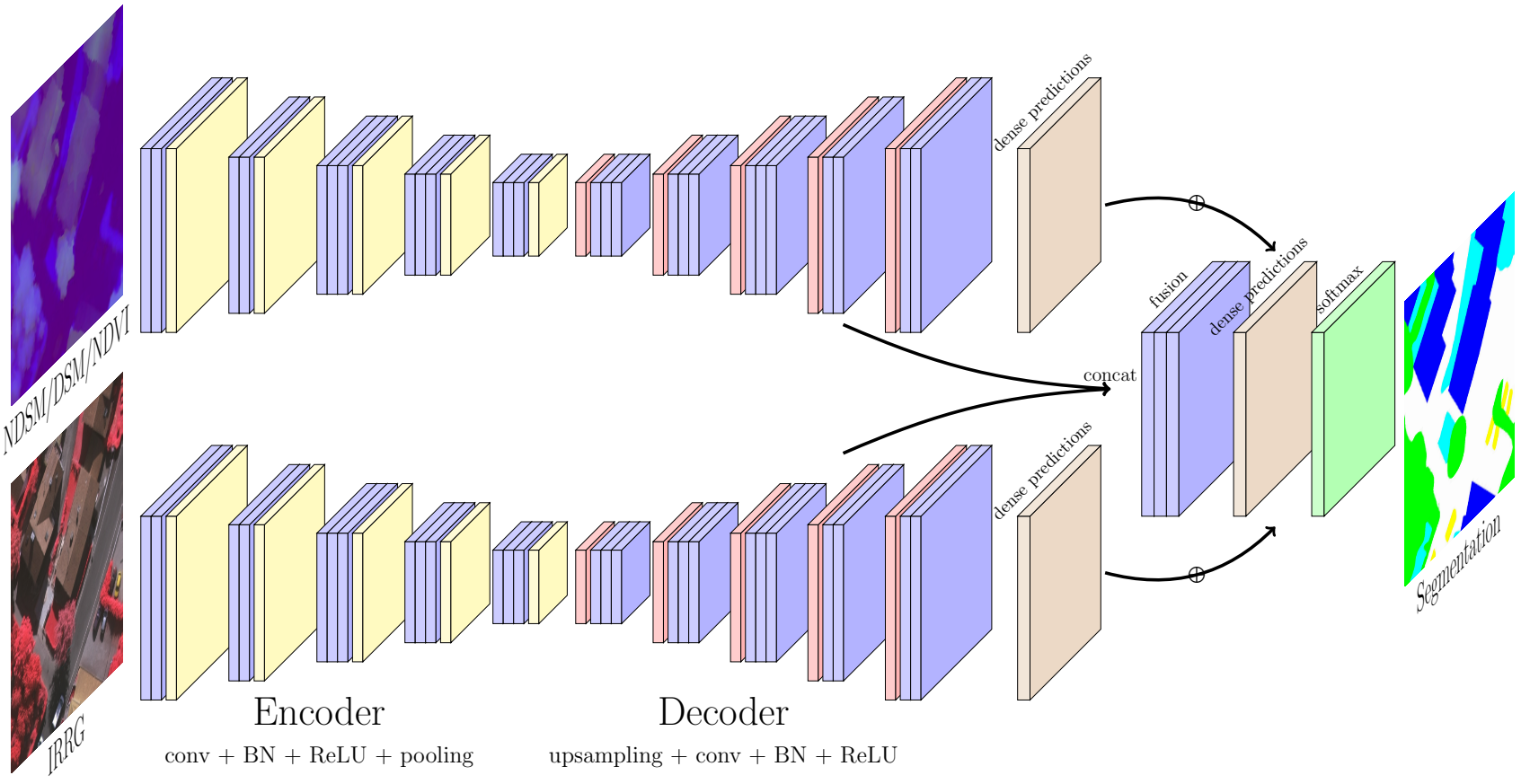
Learning based fusion

Residual correction

Inspired by signal processing and **residual learning** (He et al., 2015), we design a residual correction module for merging the two prediction streams.



The new framework



Results

Table 1: ISPRS 2D Semantic Labelling Challenge Vaihingen results.

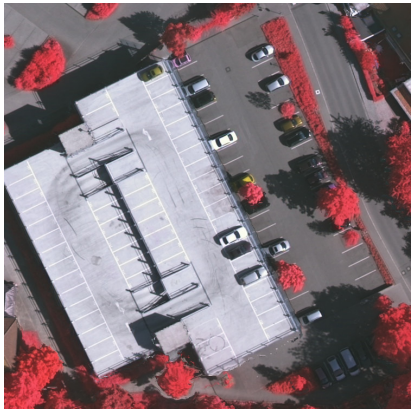
Method	imp surf	building	low veg	tree	car	Accuracy
RF + CRF ("HUST")	86.9%	92.0%	78.3%	86.9%	29.0%	85.9%
CNN ensemble ("ONE_5")	87.8%	92.0%	77.8%	86.2%	50.7%	85.9%
FCN ("DLR_2")	90.3%	92.3%	82.5%	89.5%	76.3%	88.5%
FCN + RF + CRF ("DST_2")	90.5%	93.7%	83.4%	89.2%	72.6%	89.1%
SegNet++	**91.5%**	94.3%	82.7%	89.3%	**85.7%**	89.4%
SegNet++ + fusion	91.0%	**94.5%**	**84.4%**	**89.9%**	77.8%	**89.8%**

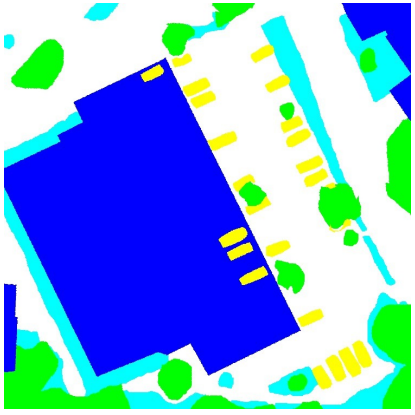
Results

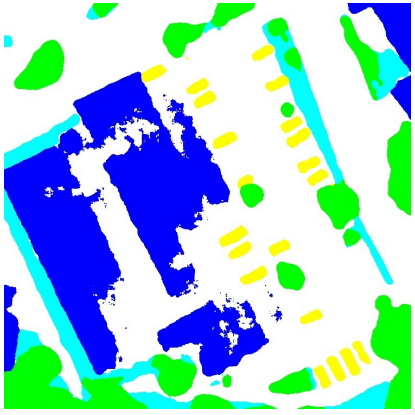
IRRG image | Ground truth | IRRG prediction | Comp. prediction | Fused predictions

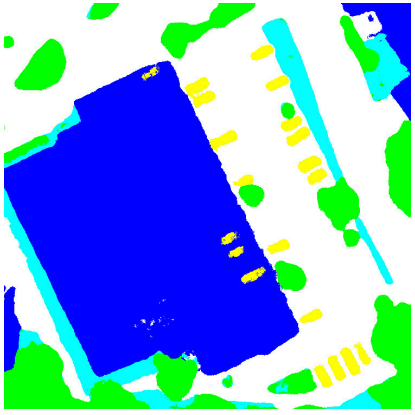
IRRG image | Ground truth | IRRG prediction | Comp. prediction | Fused predictions

Success


IRRG image


Ground truth


IRRG prediction

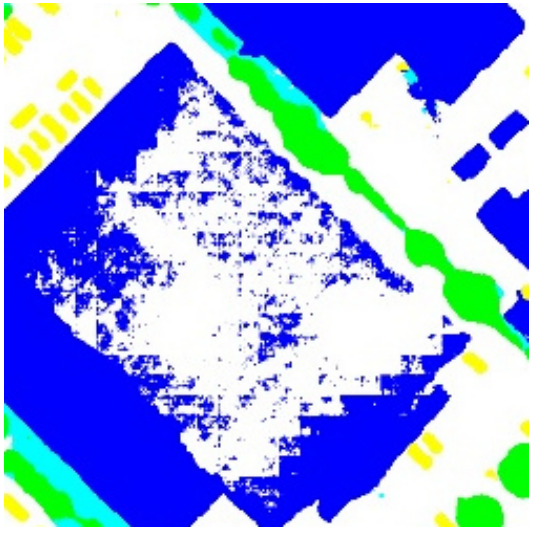

Fused predictions

Building, road or cars ?! Late fusion mixes the prediction and recovers nearly everything.

...and failure

IRRG image

NDSM



Fusion

Missing building in the NDSM: the residual correction fails to prevent the misclassification.

Conclusion

Conclusion

Contribution

- Late fusion strategy for urban remote sensing data
- Deep learning based prediction fusion
- Error correction using residual learning

Code and weights for our version of SegNet

`https://github.com/nshaud/DeepNetsForEO`

Questions ?

Questions ?

Questions are welcomed, now or at

`nicolas.audebert@onera.fr`.

Acknowledgments

The authors thank the ONERA-Total research project NAOMI for the funding of this work.

Code and weights for our version of SegNet

`https://github.com/nshaud/DeepNetsForEO`