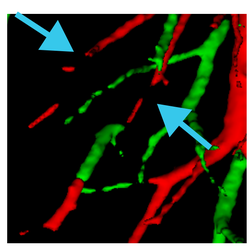
TopoMortar: A dataset to evaluate image segmentation methods focused on topology accuracy

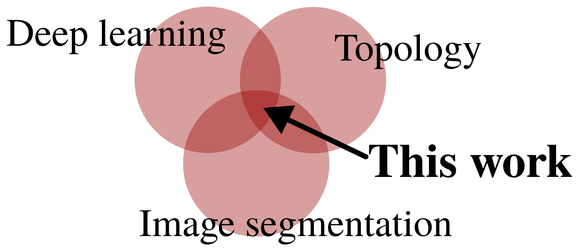
Juan Miguel Valverde[1,2] **Motoya Koga**[3] **Nijihiko Otsuka**[3] **Anders Bjorholm Dahl**[1]

[1]*Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, Denmark*
[2]*A.I. Virtanen Institute, University of Eastern Finland, Finland*
[3]*Department of Architecture, Faculty of Engineering, Sojo University, Japan*

INTRODUCTION

- **Topology accuracy** (connectivity) is crucial for quantification
  ↳ *Number of connected components and holes*





Are topology loss functions really useful? 🤔

Their evaluation was constrained due to:

Dataset challenges not being accounted for — We tackle →
- Accurate labels
- In-distr. test set
- Large training set

Imperfect labels | Corrupted images | Small datasets

We tackle — Limited experimental setting

No data augmentation | One random seed | Old architectures

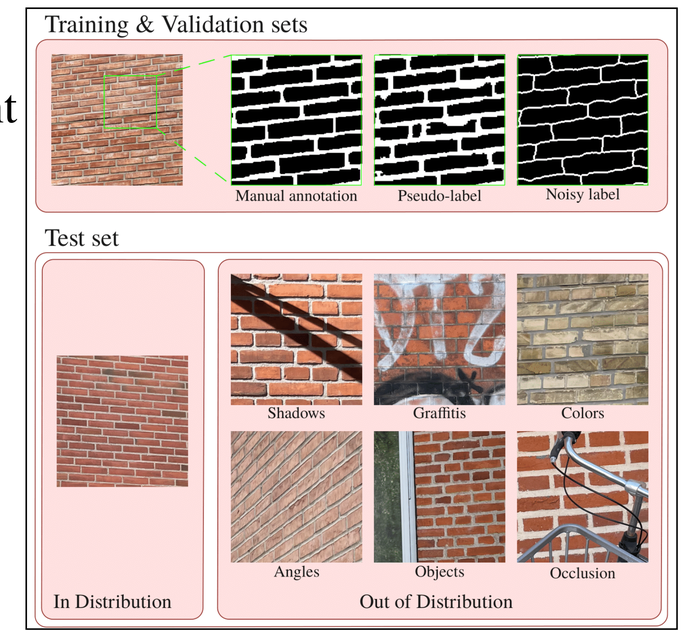
TOPOMORTAR

Task: Mortar segmentation in red brick wall images

Characteristics:

- Fixed train/val/test splits ⟶ Avoids arbitrary splits

- Compact ⟶ Convenient
  Pixel dims: 512x512
  Dataset size: 420 images



- Small training set
- Noisy and pseudo-labels
- Out-of-distr. test set

} Allows to study performance under different dataset challenges

EXPERIMENTS, RESULTS & FINDINGS

⚙️ **Strong baselines:** Ten data augmentations | Ten random seeds | DynUNet | Long training time

🧪 **Research question** 💻 Experimental setup 📊 Results 🎯 **Conclusion**

🧪 1. Are there cheaper alternatives to topology losses?

💻 We looked at:

A) Applying Data Augmentation
B) Increasing the dataset size
C) Tackling noisy labels

📊 Cheaper alternatives (A,B,C) can make CE+Dice loss more effective than many topology loss functions.

🎯 **Tackling dataset challenges with well-known strategies (A,B,C) improves topology accuracy**

🧪 2. Are topology losses useful in datasets with no challenges? (i.e., large enough datasets with good labels)

💻 We compared CE+Dice loss vs. several topology losses

- TopoMortar
- Accurate labels for training
- In distr. test set
- Large training set

📊 Only 1/7 topology losses was better than CE+Dice loss

🎯 **Most topology losses offer no advantage when training a model 1) for long enough, 2) on a sufficiently-large dataset, 3) with accurate labels and 4) strong data augmentation.**

🧪 3. Does the effectiveness of topology losses depend on the dataset challenges?

💻 We evaluated several losses under dataset challenges:

- Very small training set
- Out-of-distr. test set
- Pseudo-labels
- Noisy labels

📊 Some topology losses performed better than others under the presence of specific dataset challenges

🎯 **The effectiveness of topology losses depends on the dataset challenges.**

🧪 4. Can cheaper alternatives improve topology losses?

💻 We employed topology losses with methods to tackle:

- Out-of-distr test set ⟶ Extra DA transform
- Noisy/Pseudo-labels ⟶ Self-training

📊 All losses improved topology accuracy.
Often, CE+Dice loss outperformed topology losses.

🎯 **The effectiveness of topology losses can be improved with strategies to tackle dataset challenges.**

Acknowledgements. This work was supported by Villum Foundation and NordForsk.