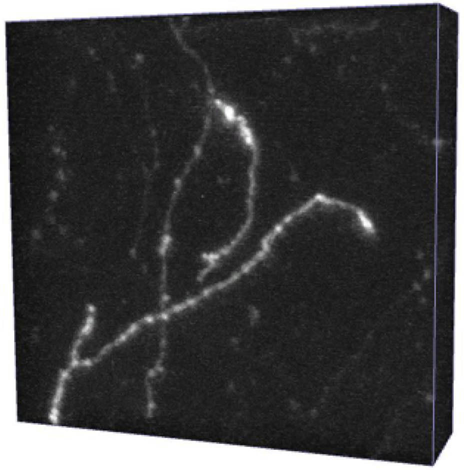
Active Learning and Focused Proofreading for Delineation

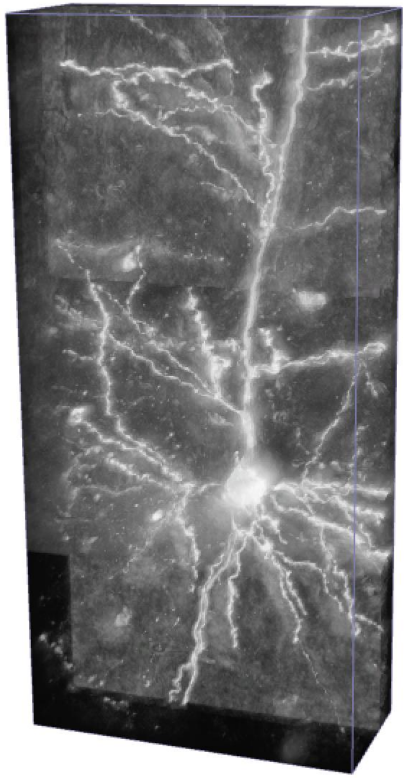
AGATA MOSINSKA, JAKUB TARNAWSKI, PASCAL FUA

EPFL

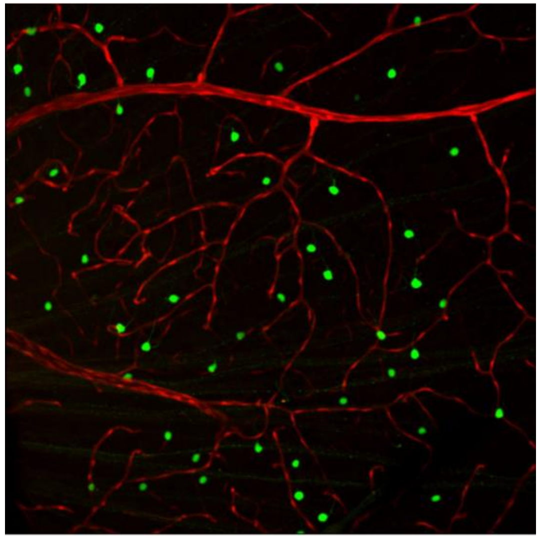
Linear Structures



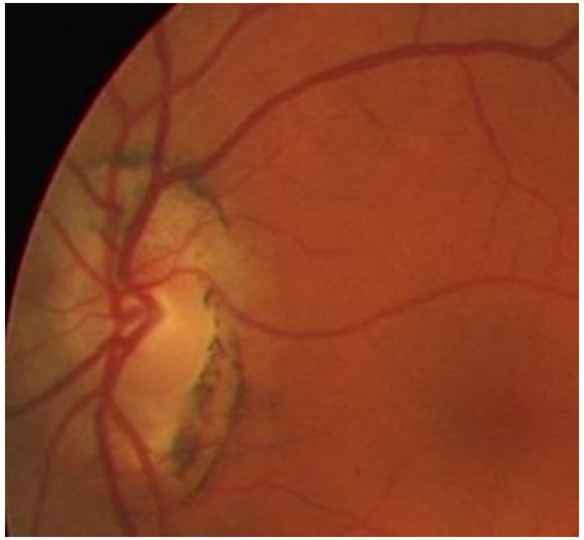
~20 μm



~200μm



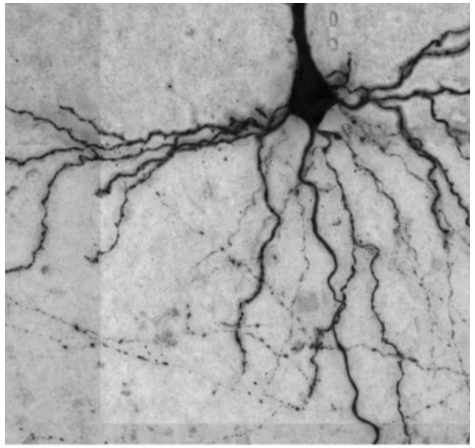
~1mm



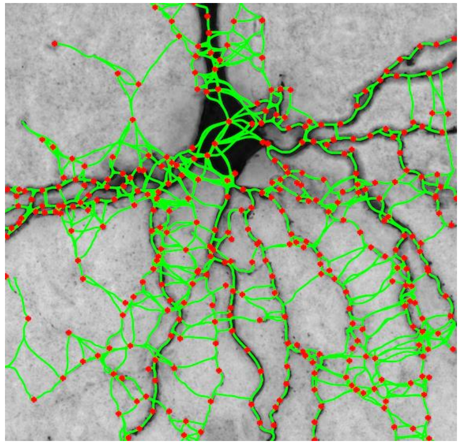
~1cm

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

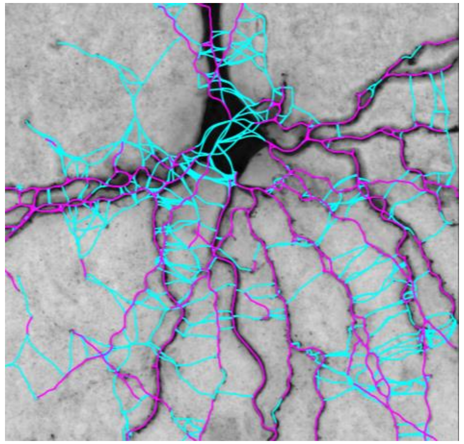
Graph-based Reconstruction



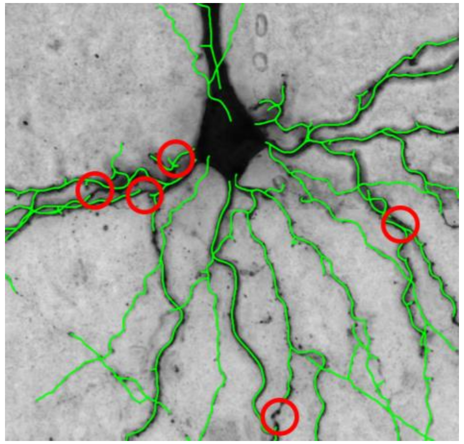
Image



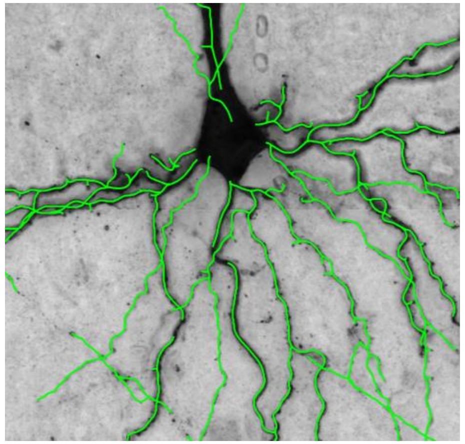
Overcomplete graph



Weighted graph after edge classification



Initial reconstruction - MinSubGraph

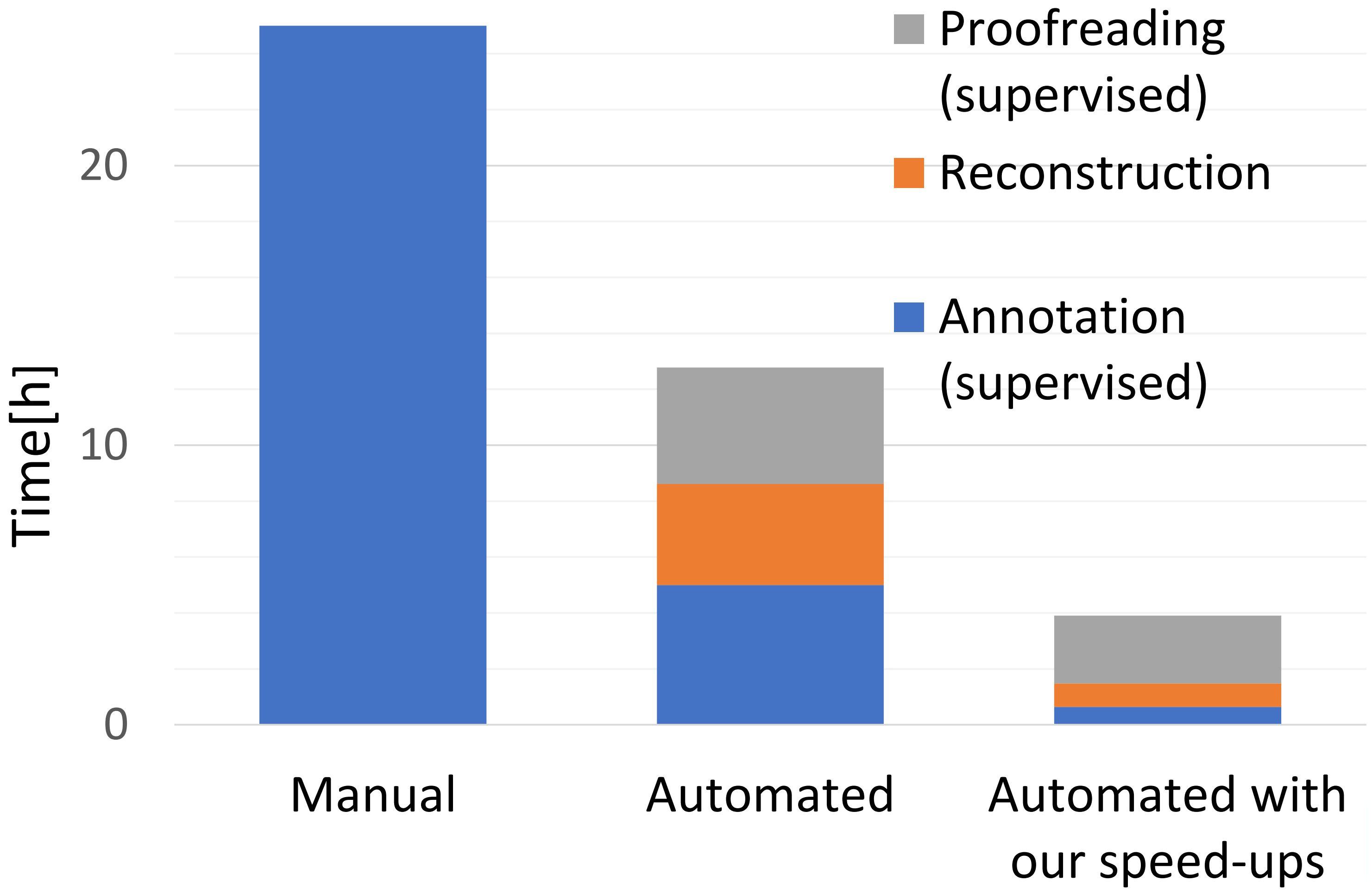


Final result after proofreading

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

How much time do we save with automation?

To reconstruct
25 out of 70 000 000
neurons in the mouse
brain we need*:



* estimated times

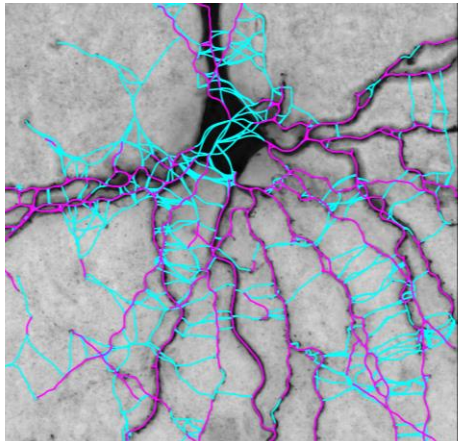
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

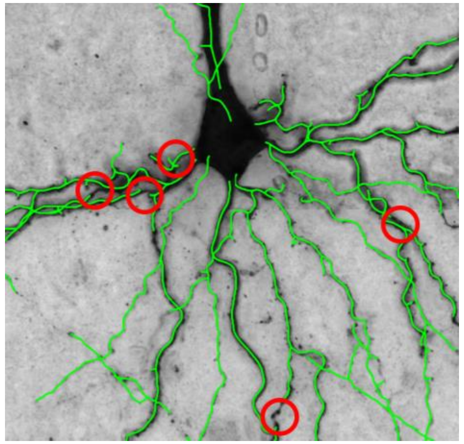
Our Contribution

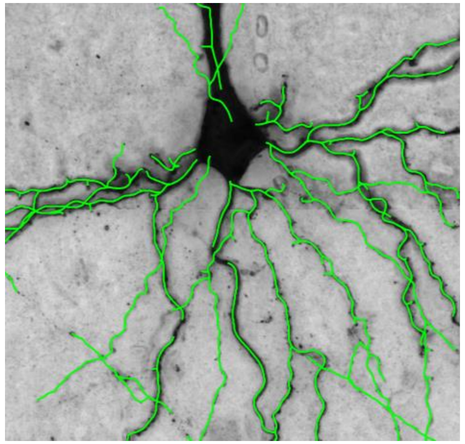
Accelerating 3 steps of reconstruction

Geometry-based AL strategy

Efficient MIP formulation







Annotation

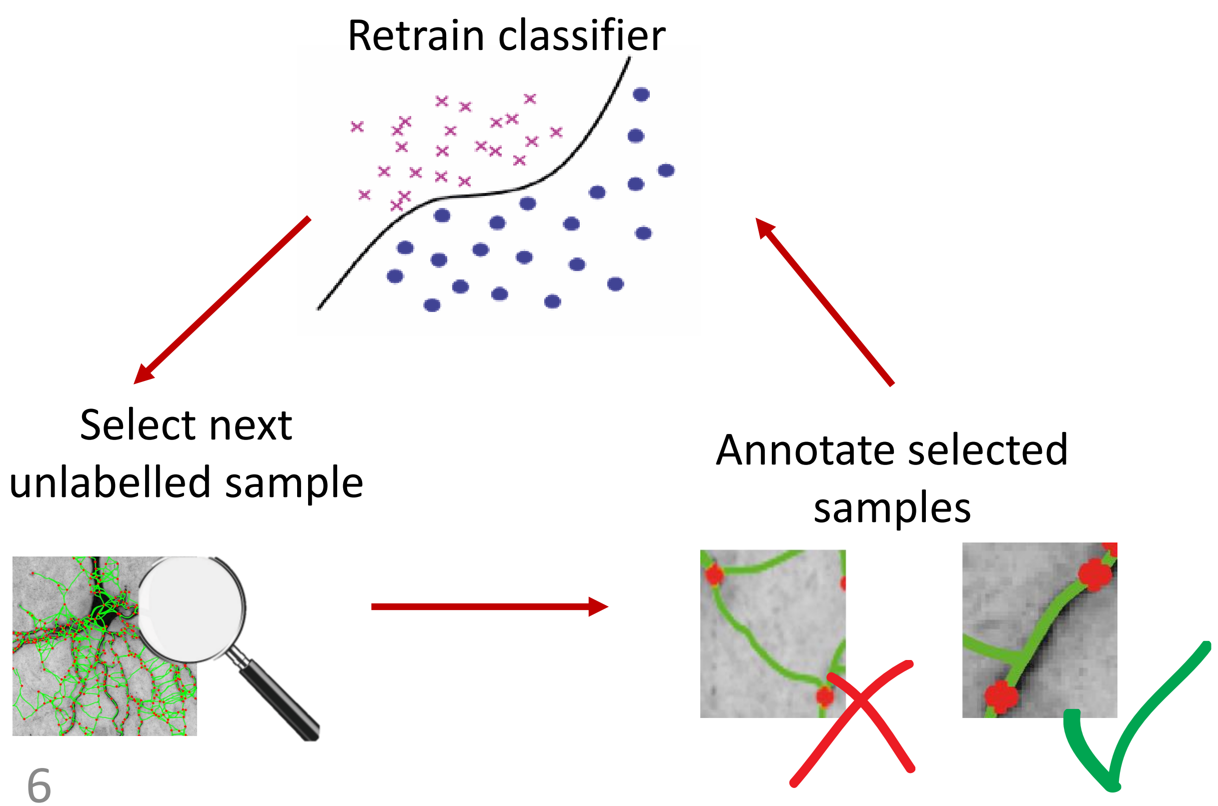
Reconstruction

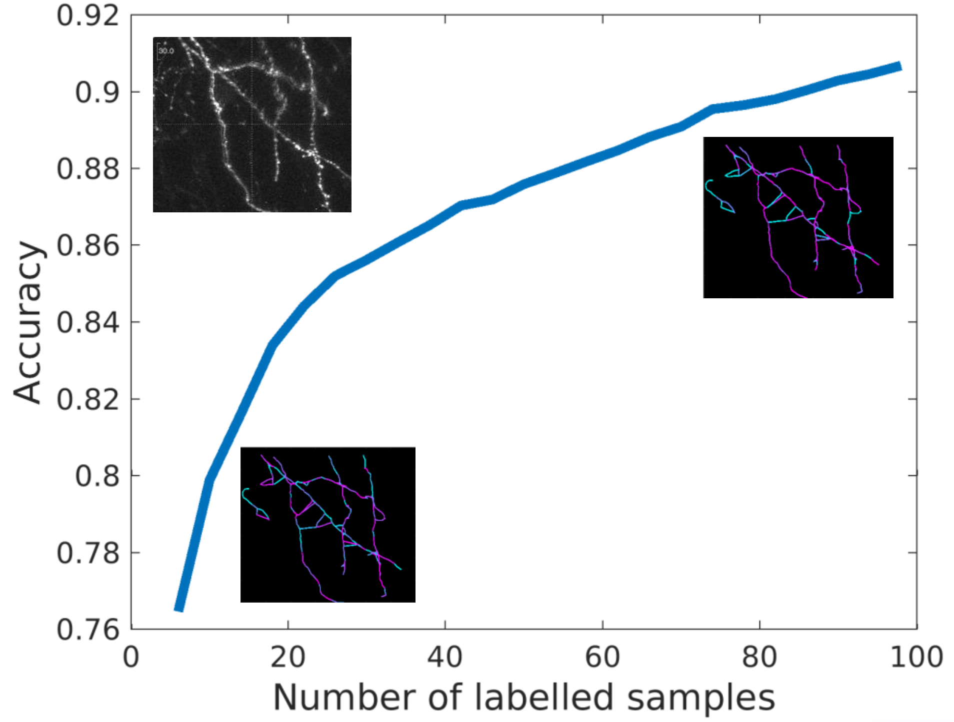
Proofreading

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Active Learning for Faster Annotation

Goal: reduce the annotation effort by labelling only the most informative examples





ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Active Learning for Faster Annotation

Question: how to select informative samples?

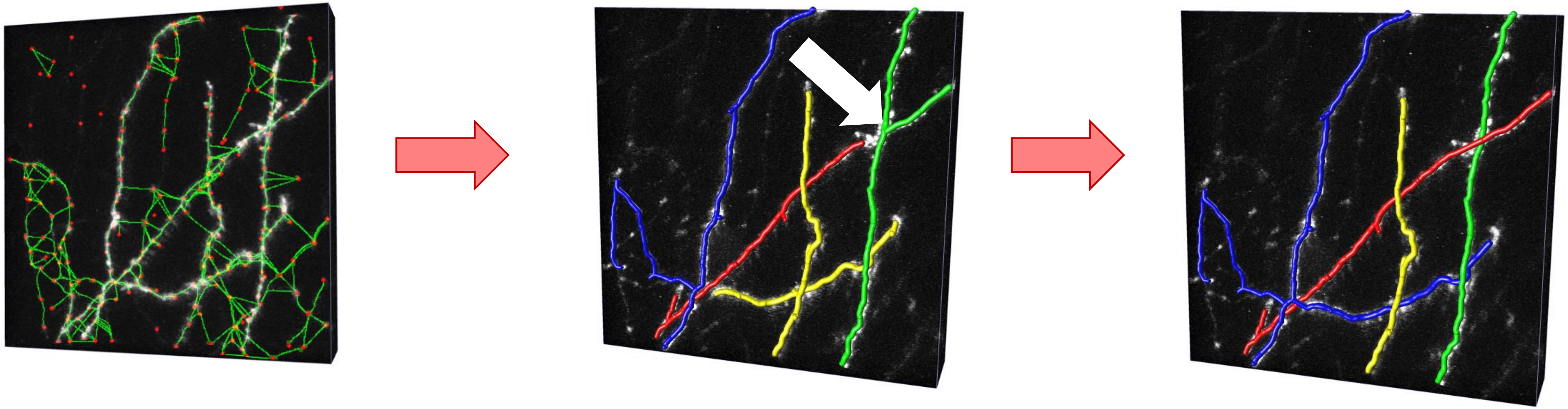
1. Uncertainty-based Sampling
2. Density-based Sampling
3. Expected Error Reduction

BUT most methods do not take geometry into account.

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Exploiting Geometry to Speed Up Annotation

Alter the weight of one edge at a time and see how the reconstruction changes.

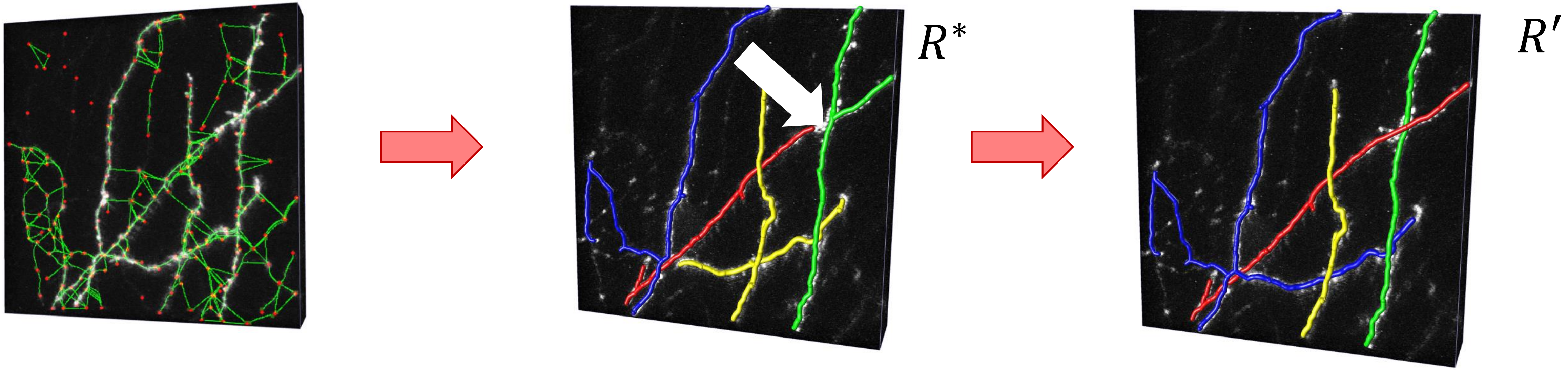


Select the sample that causes the greatest improvement in reconstruction.

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Exploiting Geometry to Speed Up Annotation

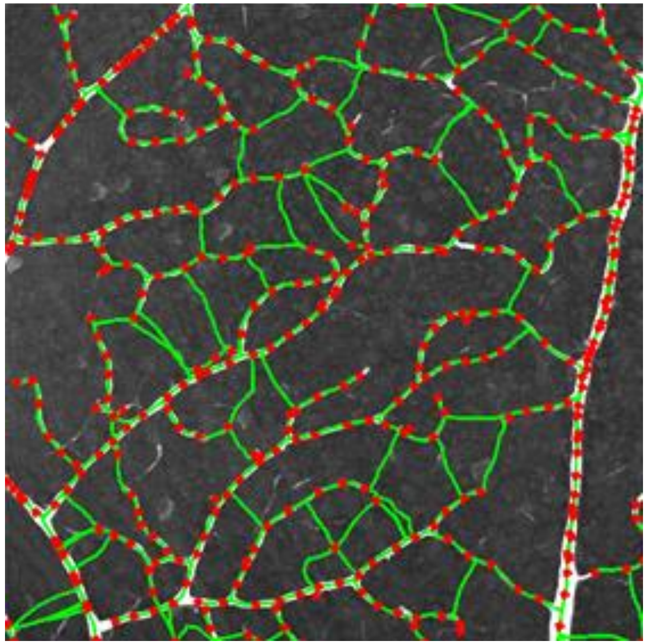
Improvement measured by the change in cost: $\Delta c = c(R^*) - c(R')$

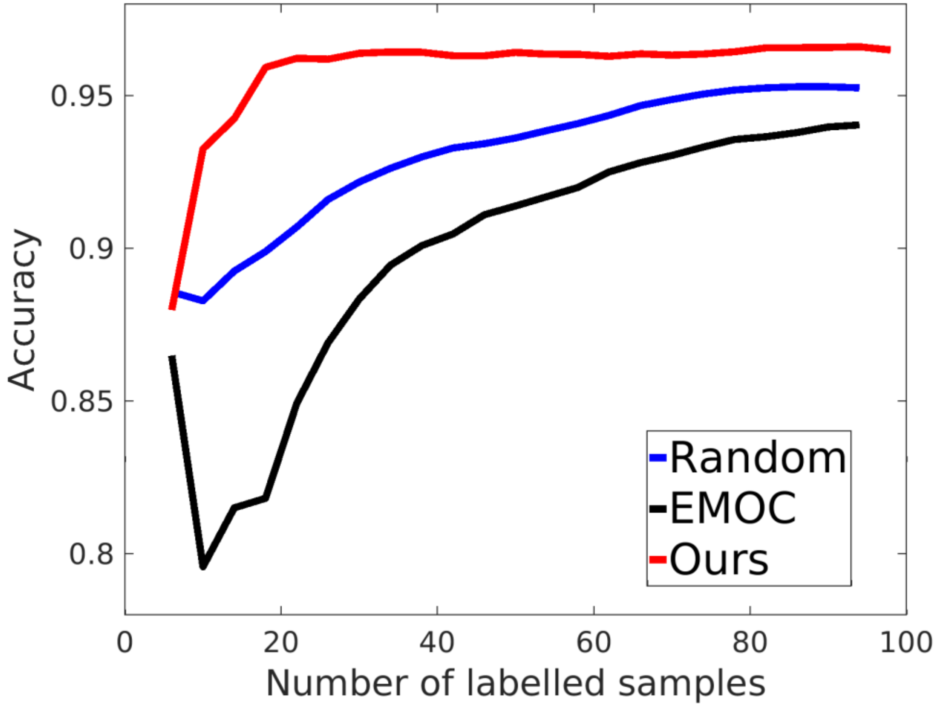


Intuition: if after changing the edge's weight the cost decreases, then the edge might have been misclassified.

Exploiting Geometry to Speed Up Annotation

Sometimes random annotation performs better than Active Learning strategies...

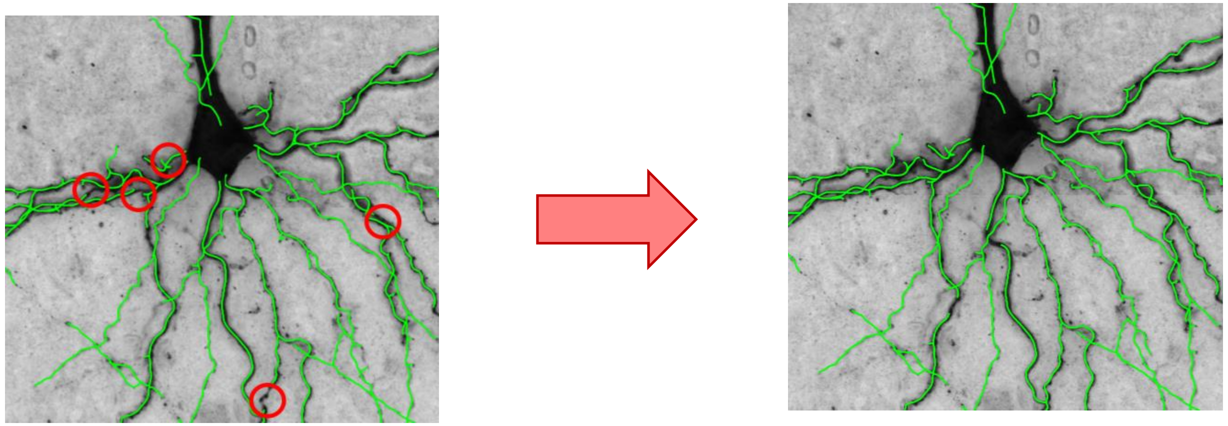




But exploiting geometry helps to avoid some common pitfalls

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Focused Proofreading



Favor samples that not only decrease the cost, but also change the topology.

Combine cost and topology and score edges by:

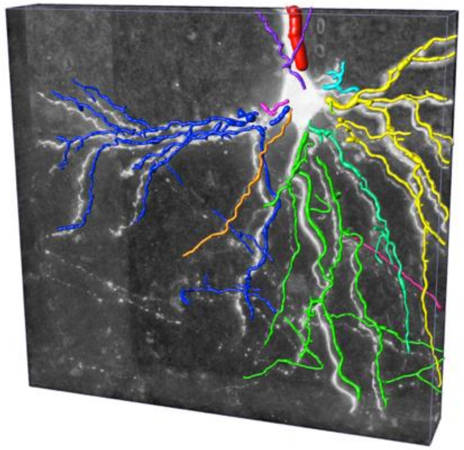
$$\frac{c(R^*) - c(R')}{DIADEM(R^*, R')}$$

DIADEM score captures similarity between two graphs

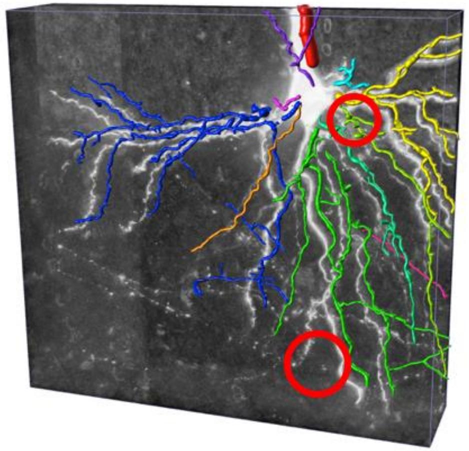
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Combining Active Learning and Proofreading

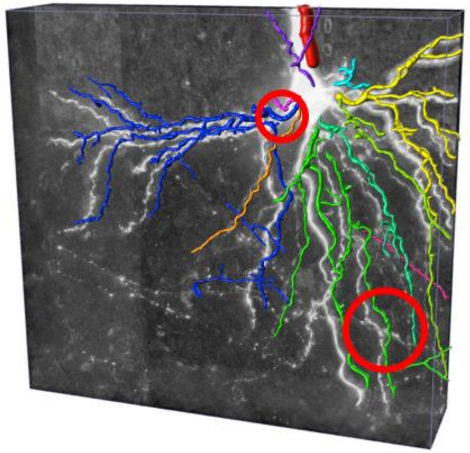
50 samples to train classifier (AL) and
25 to fast-proofread the reconstruction



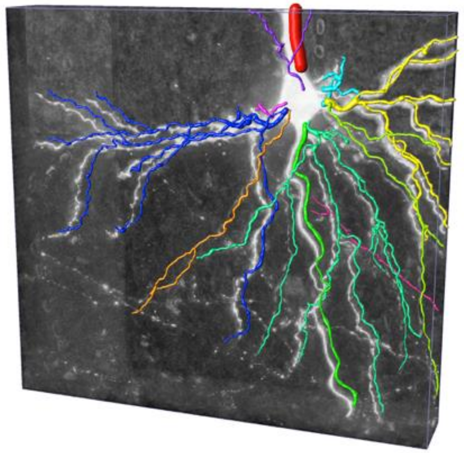
Initial



After 15 edits



After 25 edits



Ground truth

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Efficient MIP Formulation

Turetken et al. PAMI'16: $O(|V||\mathcal{E}|)$ size
Our equivalent formulation: $O(|\mathcal{E}|)$ size

	Axons1	*Axons2*	*Axons3*	*Axons4*	*Axons5*	*Axons6*
# edges	164	223	224	265	932	2638
MIP [3]	0.91	1.04	1.19	1.45	78.3	393.7
MIP ours	0.03	0.10	0.04	0.23	0.10	5.23
speedup	26.1x	10.1x	27.3x	6.3x	743.5x	75.2x

	BFNeuron1	*BFNeuron2*	*OPF1*	*OPF2*	*BFNeuron3*	*BFNeuron4*
# edges	120	338	363	380	645	2826
MIP [3]	0.48	2.25	1.53	1.65	2.13	308.23
MIP ours	0.02	0.12	0.05	0.08	0.26	2.30
speedup	18.2x	17.7x	29.4x	19.9x	8.1x	134.0x

[3] Turetken et al., PAMI'16

Conclusion

- Reduction in annotation and proofreading effort by employing geometry-based strategy
- Efficient formulation of optimization problem that allows for interactive applications
- Next steps:
  - edge sampling strategies
  - other definitions of change in reconstruction