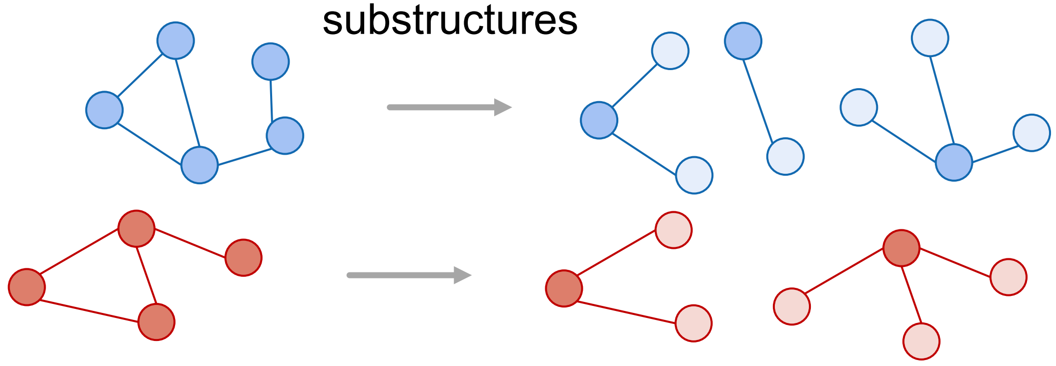
Wasserstein Weisfeiler-Lehman Graph Kernels

Matteo Togninalli*, Elisabetta Ghisu*, Bastian Rieck, Felipe Llinares-López, and Karsten Borgwardt
Machine Learning and Computational Biology Lab
ETH Zurich @Basel

$\mathcal{R}$-Convolution kernels aggregate node representations

(1) Decompose graphs in substructures



(2) Compute substructure similarities

$$k_{base}(a, b)$$

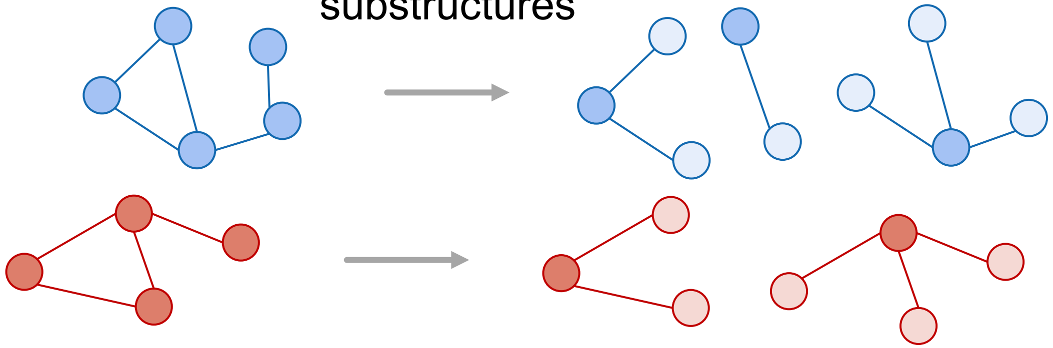
(3) Average the similarities

$$\frac{1}{n_G n_{G'}} \sum_{a \in G} \sum_{b \in G'} k_{base}(a, b)$$

Wasserstein Weisfeiler-Lehman kernel (WWL)

OLD

(1) Decompose graphs in substructures



(2) Compute substructure similarities

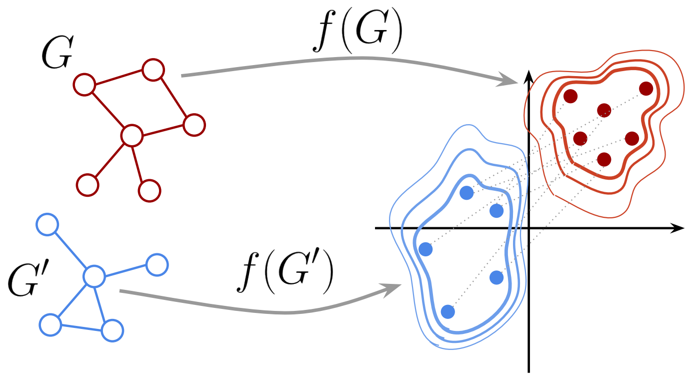
$$k_{base}(a, b)$$

(3) Average the similarities

$$\frac{1}{n_G n_{G'}} \sum_{a \in G} \sum_{b \in G'} k_{base}(a, b)$$

NEW

(1) Embed the graphs' nodes



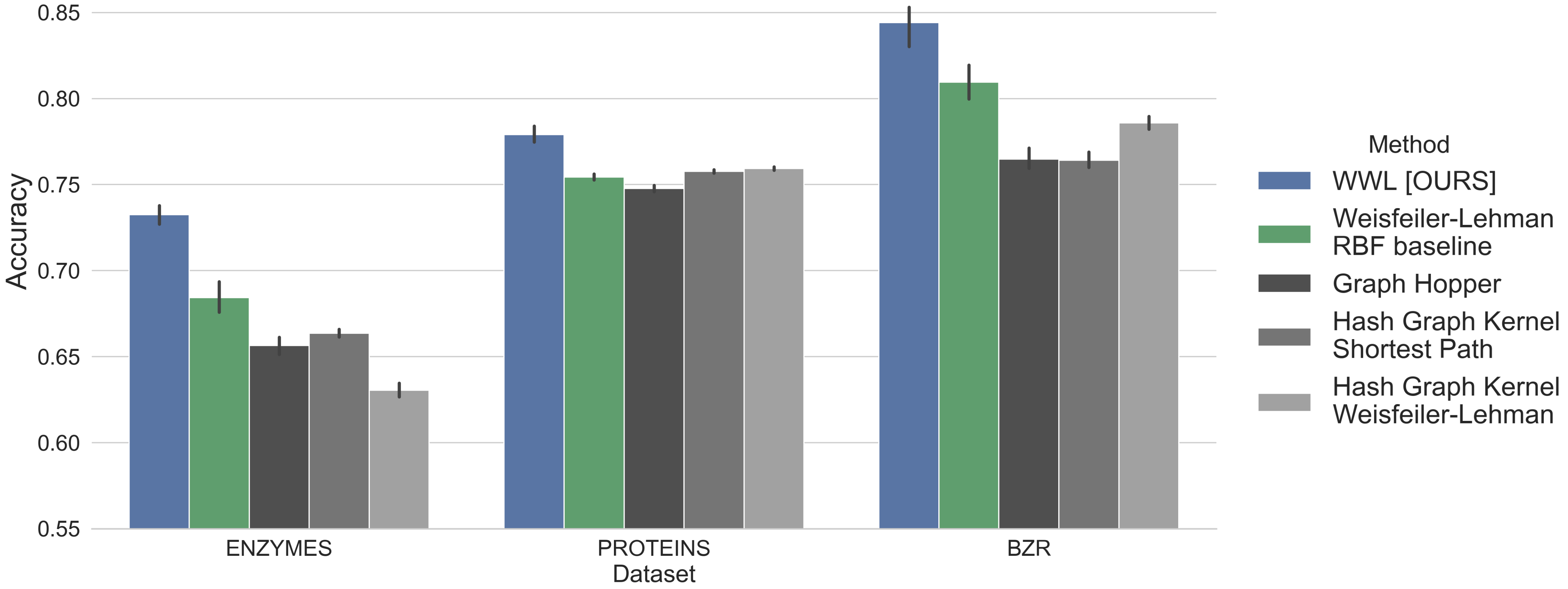
(2) Compute the Wasserstein Distance

$$D_W^f(G, G')$$

(3) Apply the Laplacian kernel

$$K_{\text{WWL}}(G, G') = e^{-\lambda D_W^f(G, G')}$$

WWL outperforms the state-of-the-art



Morris, Christopher, et al. "Faster kernels for graphs with continuous attributes via hashing." *2016 IEEE 16th International Conference on Data Mining (ICDM)*. IEEE, 2016.
Feragen, Aasa, et al. "Scalable kernels for graphs with continuous attributes." *Advances in Neural Information Processing Systems*. 2013.

Take home messages

- We present a **novel similarity measure** between graphs
- It can be used to accurately classify
  - categorically labelled graphs
  - continuously attributed and weighted graphs

- Come and see us tonight (5:00 pm – 7:00 pm) at **poster #12**
- matteo.togninalli@bsse.ethz.ch
- elisabetta.ghisu@bsse.ethz.ch