Algebraic Multigrid Acceleration for Efficient Label Spreading

Antonia van Betteray[1], Jonathan Klees[1], Miriam Schäfers[2], Matthias Rottmann[1]
[1]University of Osnabrück, [2] Ruhr University Bochum

ECCV 2026 | MALMÖ | SEPT 8–12

Overview

Motivation

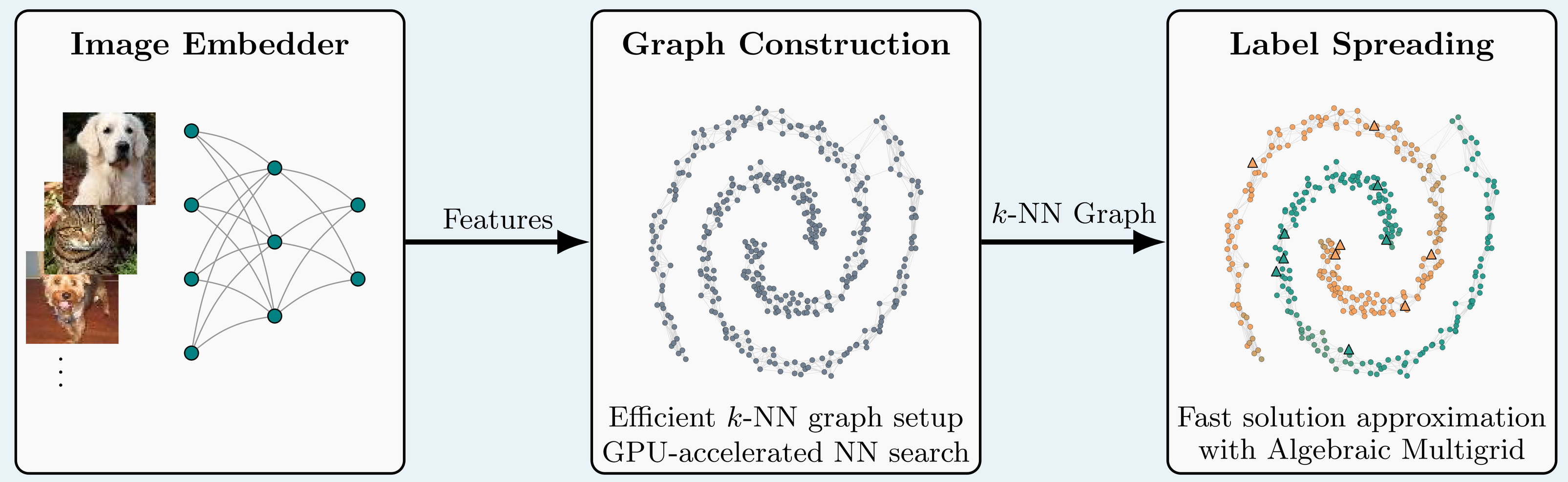
- Label Spreading is a useful technique in Semi-Supervised Learning
- Label Spreading can be used for semi-automated labeling but it has computational limitations
  - Graph Construction
  - Matrix Inversion
- Current implementations are not suitable for large datasets

Our Approach

- We efficiently construct k-NN graphs using GPU parallelization
- We leverage Algebraic Multigrid (AMG) for efficient solving
- AMG solvers have a multilevel structure and can approximate both direct and long range information exchange in label spreading

Main Results

- Our implementation significantly reduces runtime of label spreading
- It enables the application on large image datasets
- With AMG solvers, label spreading is less sensitive to hyperparameters



Framework Overview

- We embed images using CLIP and UMAP.
- Efficient neighborhood search by Faiss for graph construction
- Label spreading in embedding space with efficiency through AMG

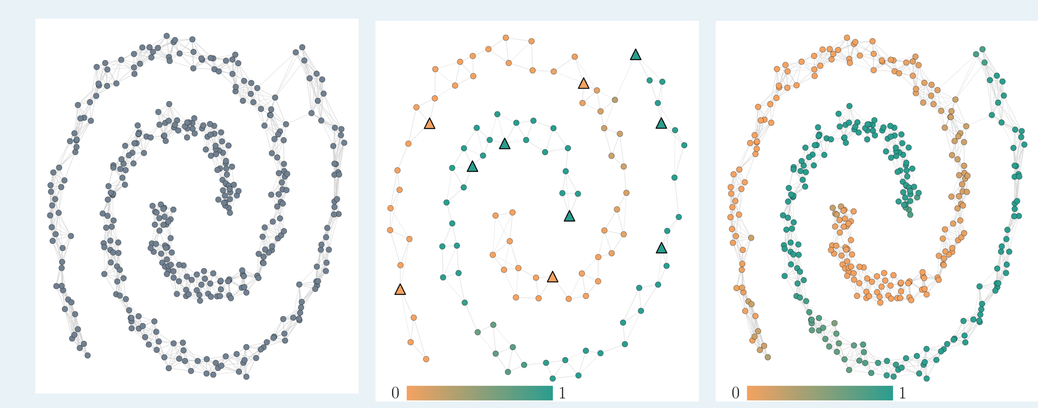
Good To Know

Label Spreading

- For label spreading, initial labels are encoded in a matrix $Y \in \mathbb{R}^{n\times c}$
- Label spreading builds on Laplacian regularization using the affinity matrix $W \in \mathbb{R}^{n\times n}$ and its normalized version $S = D^{-1/2}WD^{-1/2}$
- Estimated labels are then derived by $F = (I - \alpha S)^{-1}Y \in \mathbb{R}^{n\times c}$
- α controls the diffusion strength, a larger α yields a stiffer solution
- The condition number of the system $I - \alpha S$ is $\kappa = \frac{1+\alpha}{1-\alpha}$
- Large α is often required and ensures consistency in the large data limit

Algebraic Multigrid Methods

- AMG solvers iteratively coarsen the large problem
- The system is solved on the coarsest grid with little effort
- The solution is then smoothened to finer grids
- AMG approaches yield near-linear complexity for large sparse systems

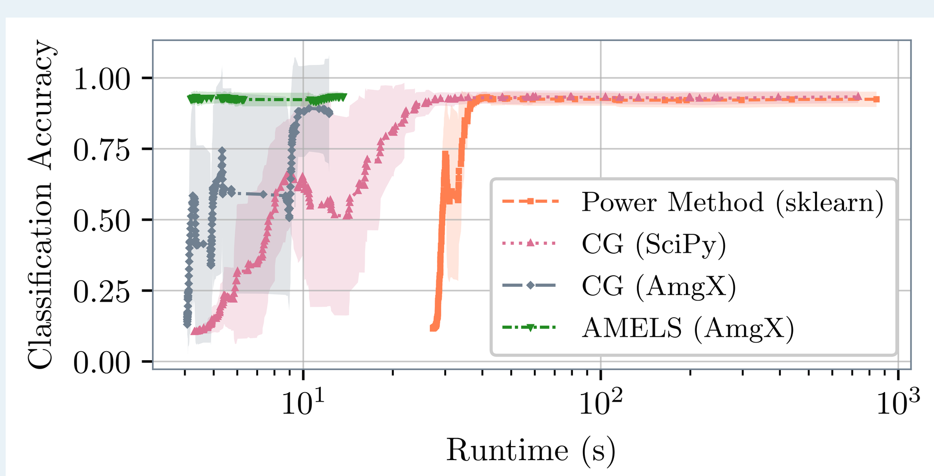


Full Graph | Coarse Level | Full Solution

Efficiency Results

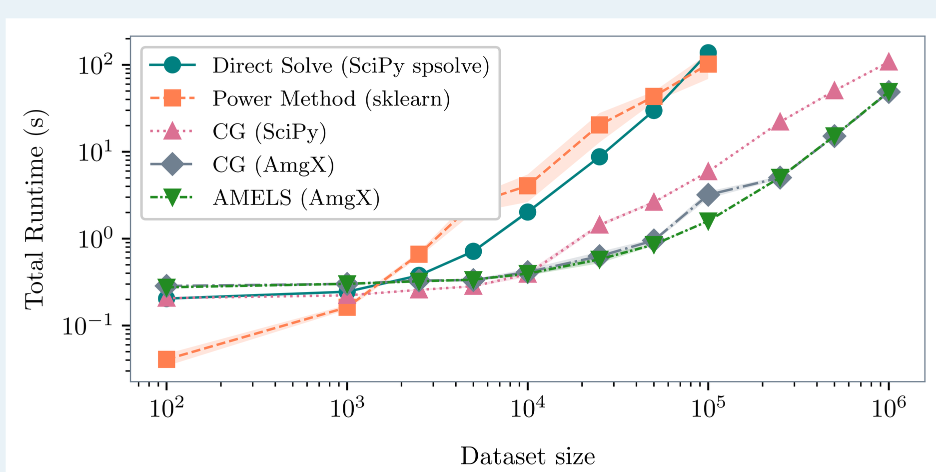
Solution is found faster and more reliably

- Other methods converge significantly slower
- We find an accurate solution faster and with little variance in results



Scalability to large data

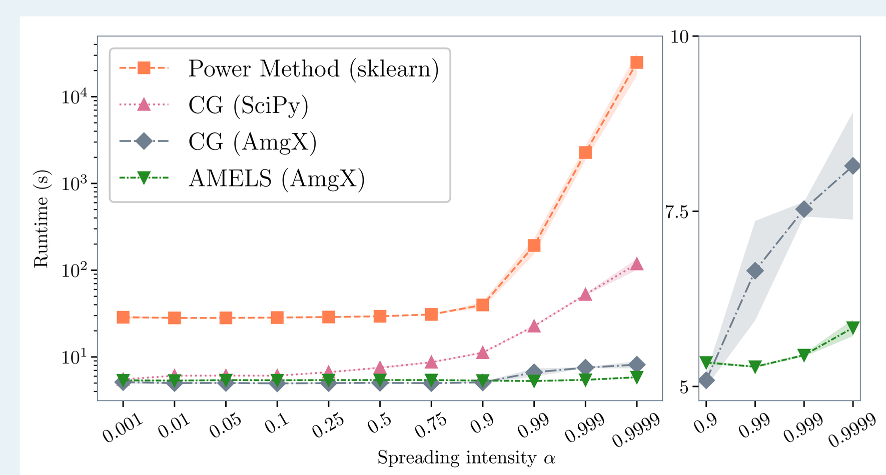
- We increase the dataset size and observe runtime
- We improve scalability, enabling large-scale applications of label spreading



Robustness Results

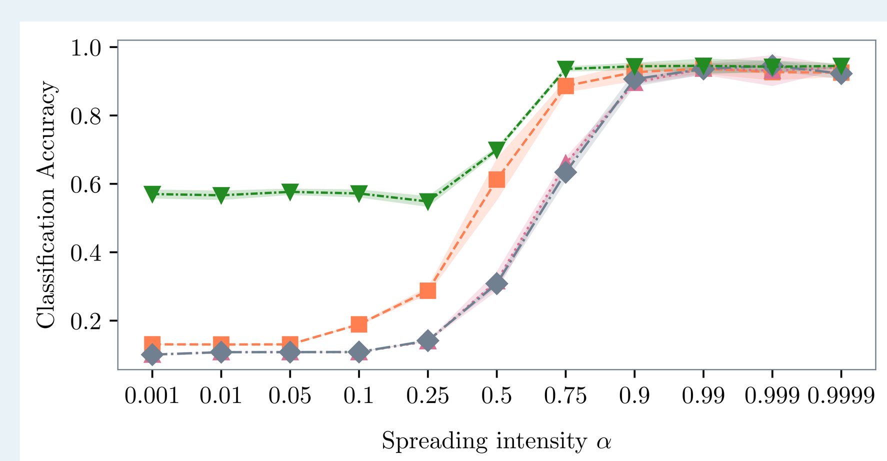
Ill-conditioned systems

- Large spreading intensity α $\Rightarrow$ system ill-conditioned
- Our method can handle this better in terms of runtime



Hyperparameter sensitivity

- Label spreading performance depends on α
- For small α, our multilevel approach still propagates label information



Takeaways

- We leverage multilevel AMG solvers for efficient label spreading
- We improve the spectral properties of the system matrix and approximate direct and long range connections at once
- Our implementation
  - significantly reduces runtime of label spreading
  - makes it applicable to large image datasets
  - is more robust to hyperparameters (runtime and performance)

Contact

{antonia.vanbetteray,jonathan.klees,
mathias.rottmann}@uni-osnabrueck.de
m.schaefers@rub.de

GitHub
github.com/JonathanKlees/efficient_label_spreading

Acknowledgment: Supported by the German Federal Ministry of Research, Technology and Space (BMFTR) within the project RELiABEL (grant no. 16IS24019B).

Gefördert durch: Bundesministerium für Forschung, Technologie und Raumfahrt