# ICPR 2020
# Graph-based Interpolation of Feature Vectors for Accurate Few-Shot Classification

By: **Yuqing Hu,** Vincent Gripon, Stéphane Pateux

Jan. 12th, 2020

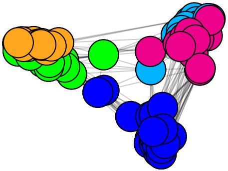
# Motivation

## Few-Shot Image Classification

- **Image Classification with few labelled data**
  
  ⟶ Model likely to overfit.

- **Transfer Learning**
  
  ⟶ A model developed for a task is reused as the starting point for a model on a second task.

- **Graph Neural Network**
  
  ⟶ A type of Neural Network which directly operates on the Graph structure, gathering transductive information between nodes.



## Problem settings

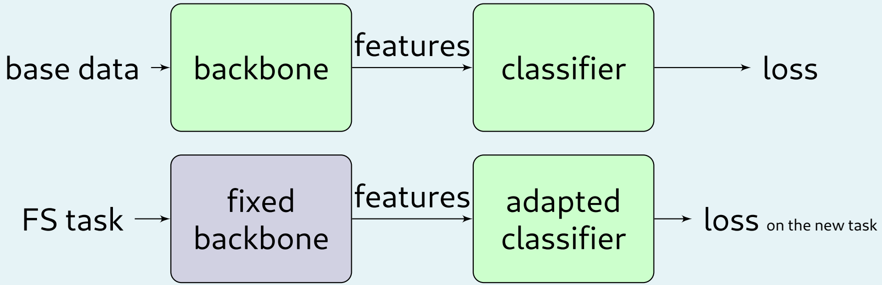
- **$s$-shot, $K_n$-way transductive classification**
  - $\longrightarrow$ Given a task containing $K_n$ classes, $s$ labelled images (support set) and $Q/K_n$ unlabelled images per class, the goal is to predict the class of the $Q$ unlabelled images (query set).


- **Dataset**
  - $\longrightarrow$ **miniImageNet** (64 base classes, 16 validation classes, 20 novel classes).
- **Benchmark settings**
  - $\longrightarrow$ Typically, $K_n = 5$, $s = 1$ or $5$, $Q = 75$ (number of unlabelled inputs $q = Q/K_n = 15$ per class). Accuracy averaged on $10,000$ randomly drawn tasks from novel classes.

# Methodology: first step (transfer learning)

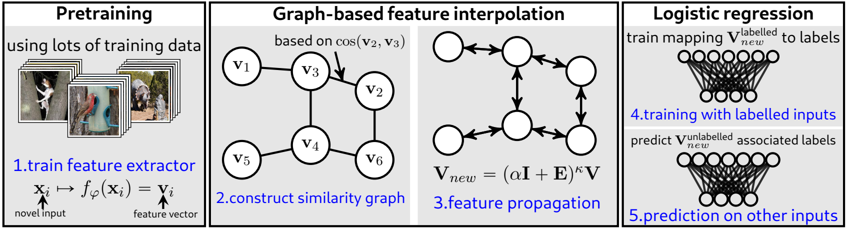
**Transferring knowledge from the large training database (64 classes with 600 samples/class)**

Exploit the **features** of **few data samples** after going through a **pretrained backbone** for **classifying new classes**.




*Snell, Jake, Kevin Swersky, and Richard Zemel. "Prototypical networks for few-shot learning." 2017.*
*Wang, Yan, et al. "SimpleShot: Revisiting Nearest-Neighbor Classification for Few-Shot Learning." 2019.*

# Methodology: transductive few-shot using graphs




- **Pretraining** Train a backbone deep neural network able to discriminate inputs from the base class dataset.
- **Graph-based feature interpolation** Firstly, construct a graph $G_T(\mathbf{V}, \mathbf{E})$ where vertices matrix $\mathbf{V}$ contains the stacked features of labelled and unlabelled inputs. Build the adjacency matrix $\mathbf{E}$ based on a $k$ nearest neighbor graph using cosine similarity function $\cos(\cdot, \cdot)$. Secondly, apply feature propagation to obtain new features $\mathbf{V}_{new}$ for each vertex.
- **Logistic regression** Learn a linear classifier on the new labelled features $\mathbf{V}_{new}^{\text{labelled}}$ and use the classifier to predict the class of query set $\mathbf{V}_{new}^{\text{unlabelled}}$.

## Comparison with state-of-the-art methods

| Method | Backbone | miniImageNet 1-shot | miniImageNet 5-shot |
|---|---|---|---|
| MAML | ResNet18 | $49.61 \pm 0.92\%$ | $65.72 \pm 0.77\%$ |
| Baseline++ | ResNet18 | $51.87 \pm 0.77\%$ | $75.68 \pm 0.63\%$ |
| Matching Networks | ResNet18 | $52.91 \pm 0.88\%$ | $68.88 \pm 0.69\%$ |
| ProtoNet | ResNet18 | $54.16 \pm 0.82\%$ | $73.68 \pm 0.65\%$ |
| SimpleShot | ResNet18 | $63.10 \pm 0.20\%$ | $79.92 \pm 0.14\%$ |
| S2M2_R | ResNet18 | $64.06 \pm 0.18\%$ | $80.58 \pm 0.12\%$ |
| LaplacianShot | ResNet18 | $72.11 \pm 0.19\%$ | $82.31 \pm 0.14\%$ |
| Transfer+Graph Interpolation (ours) | ResNet18 | $\mathbf{72.40 \pm 0.24}\%$ | $\mathbf{82.89 \pm 0.14}\%$ |
| ProtoNet | WRN | $62.60 \pm 0.20\%$ | $79.97 \pm 0.14\%$ |
| Matching Networks | WRN | $64.03 \pm 0.20\%$ | $76.32 \pm 0.16\%$ |
| S2M2_R | WRN | $64.93 \pm 0.18\%$ | $83.18 \pm 0.11\%$ |
| SimpleShot | WRN | $65.87 \pm 0.20\%$ | $82.09 \pm 0.14\%$ |
| SIB | WRN | $70.00 \pm 0.60\%$ | $79.20 \pm 0.40\%$ |
| BD-CSPN | WRN | $70.31 \pm 0.93\%$ | $81.89 \pm 0.60\%$ |
| LaplacianShot | WRN | $74.86 \pm 0.19\%$ | $84.13 \pm 0.14\%$ |
| Transfer+Graph Interpolation (ours) | WRN | $\mathbf{76.50 \pm 0.23}\%$ | $\mathbf{85.23 \pm 0.13}\%$ |

| Method | Backbone | CUB 1-shot | CUB 5-shot |
|---|---|---|---|
| S2M2_R | ResNet18 | $71.43 \pm 0.28\%$ | $85.55 \pm 0.52\%$ |
| ProtoNet | ResNet18 | $72.99 \pm 0.88\%$ | $86.64 \pm 0.51\%$ |
| Matching Networks | ResNet18 | $73.49 \pm 0.89\%$ | $84.45 \pm 0.58\%$ |
| LaplacianShot | ResNet18 | $80.96\%$ | $88.68\%$ |
| Transfer+Graph Interpolation (ours) | ResNet18 | $\mathbf{86.05 \pm 0.20}\%$ | $\mathbf{90.87 \pm 0.10}\%$ |
| S2M2_R | WRN | $80.68 \pm 0.81\%$ | $90.85 \pm 0.44\%$ |
| Transfer+Graph Interpolation (ours) | WRN | $\mathbf{88.35 \pm 0.19}\%$ | $\mathbf{92.14 \pm 0.10}\%$ |

| Method | Backbone | CIFAR-FS 1-shot | CIFAR-FS 5-shot |
|---|---|---|---|
| BD-CSPN | WRN | $72.13 \pm 1.01\%$ | $82.28 \pm 0.69\%$ |
| S2M2_R | WRN | $74.81 \pm 0.19\%$ | $87.47 \pm 0.13\%$ |
| SIB | WRN | $80.00 \pm 0.60\%$ | $85.30 \pm 0.40\%$ |
| Transfer+Graph Interpolation (ours) | WRN | $\mathbf{83.90 \pm 0.22}\%$ | $\mathbf{88.76 \pm 0.15}\%$ |

# Conclusion

- Introduction of a novel method to solve the few-shot transductive classification problem. It consists in combining three steps: a pretrained backbone, a graph-based interpolation technique and a logistic regression.
- Proposed method obtained state-of-the-art results, with the most important gains in the case of 1-shot classification.
- Proposed method requires to tune few hyperparameters, and these have a little impact on accuracy. We thus believe that it is an applicable solution to many practical problems.
- There are still open questions to be addressed, such as the case of imbalanced classes, or settings where prediction must be performed on streaming data, one input at a time.

**See the paper for more details: arxiv.org/abs/2001.09849**