ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Augmented Bi-path Network for Few-shot Learning

Baoming Yan[1,2*], Chen Zhou[1*], Bo Zhao[3], Kan Guo[2], Jiang Yang[2], Xiaobo Li[2], Ming Zhang[1] and Yizhou Wang[1]

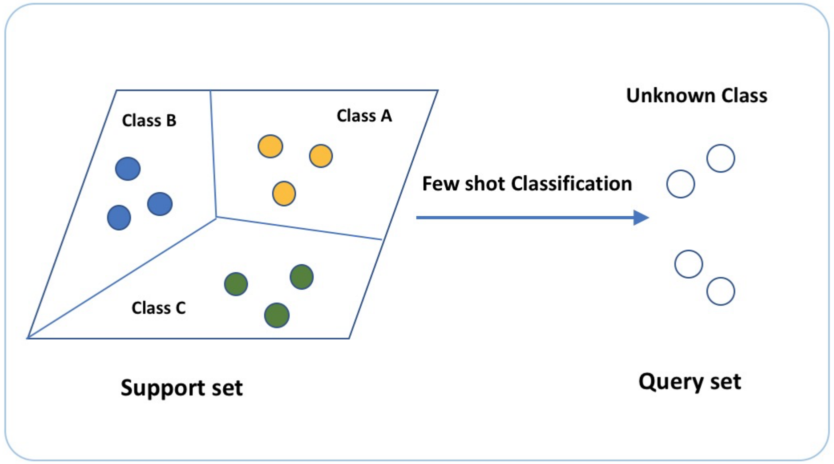
[1] Peking University, [2] Alibaba Group, [3] The University of Edinburgh

(* Indicates equal contribution)

Introduction

Few-shot learning problem definition:

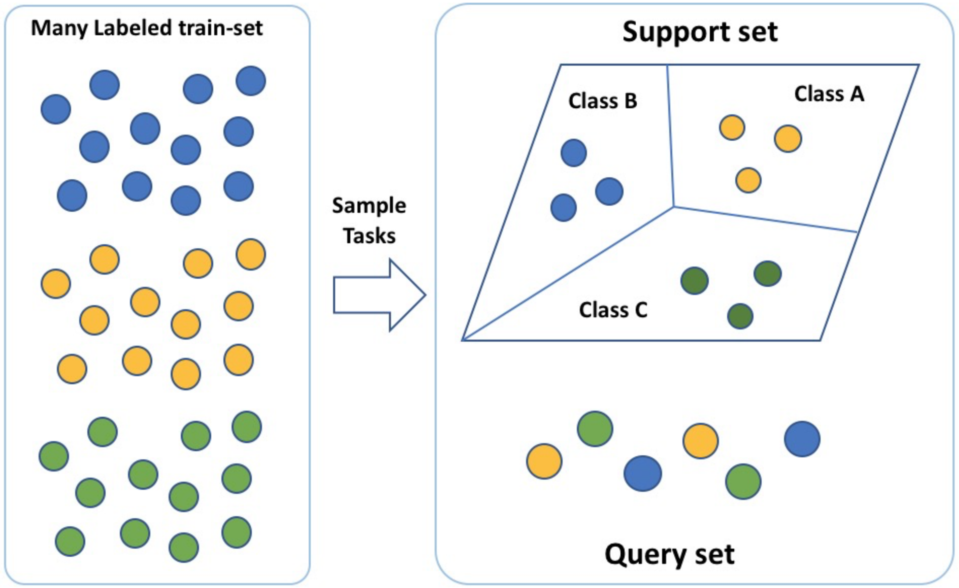
Aims to learn from few labeled training data, train a good classification model with few labeled data. E.g. fine-grained categories or medical data



N-way-K-shot: N classes with K labeled samples

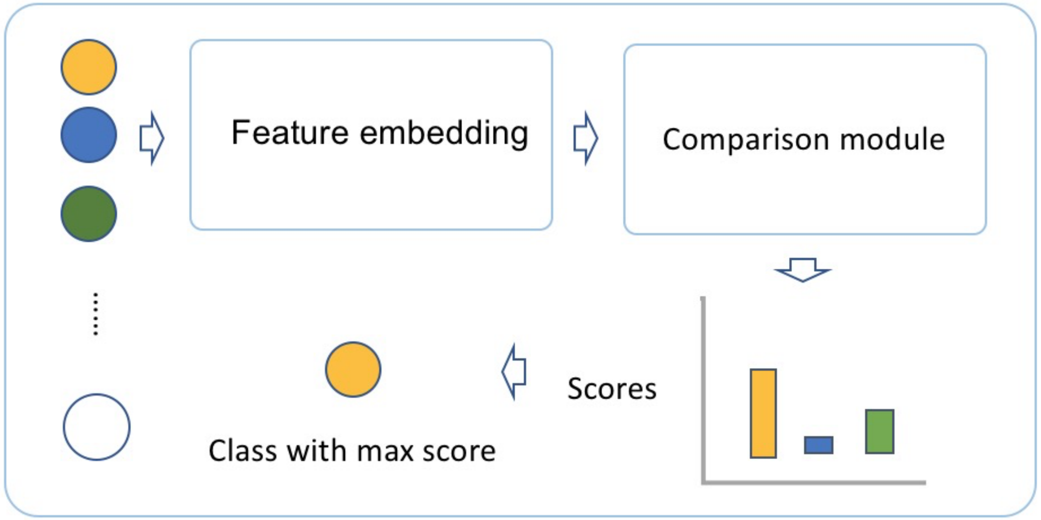
Typical solution: meta-learning based methods

Learn meta-knowledge from sampled FSL tasks



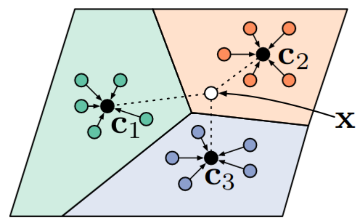
Introduction

Differ in the design of the classifier for basic tasks

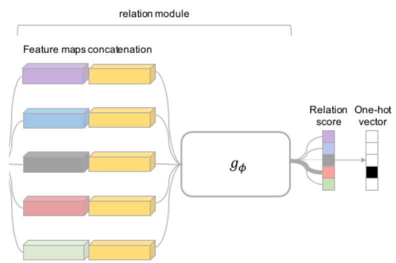


Previous methods:

Porotype Network

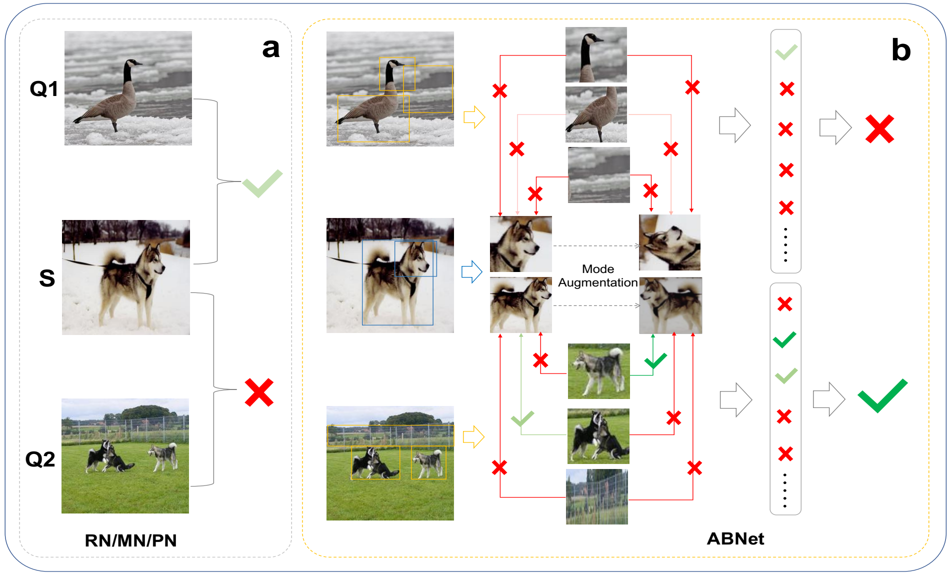


Relation Network



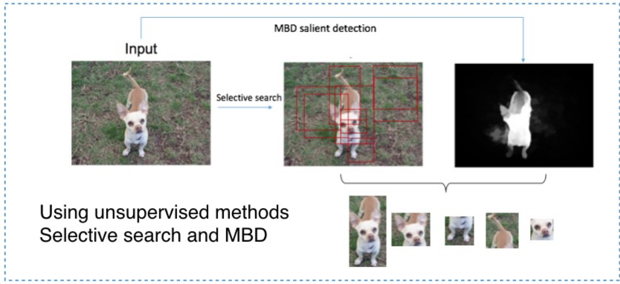
The neural network has difficulty in learning or comparing the **local features** of two images

Motivation

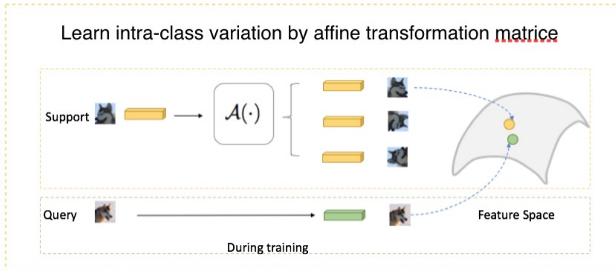


a. Image-level comparison may cause serious mis-classification
b. Introduction of salient patch makes the comparison precise.

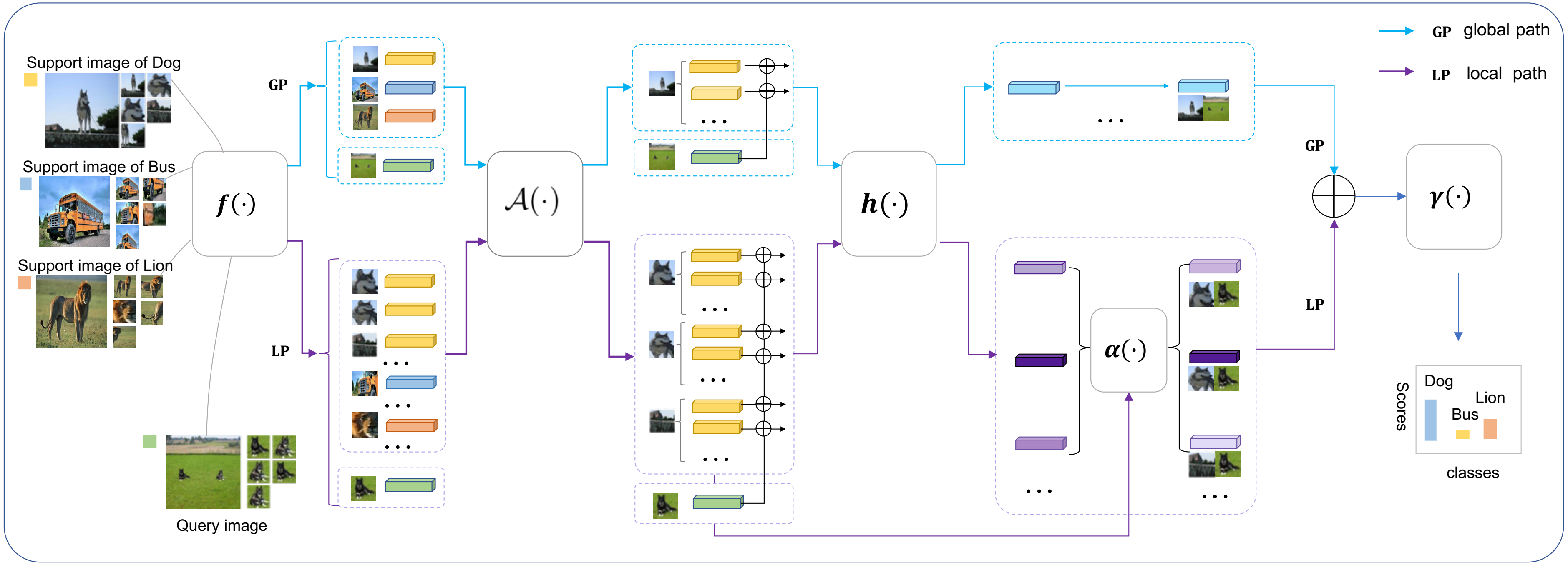
- Introduce salient patches



- Learn to augment



Methods



Salient Patch extraction

$$S(r_i) = \Pi(r_i) \cdot \frac{1}{K} \sum_{j=0}^{K} v(p_j)$$

$$v(p_j) = \min_{\pi \in \mathbb{S}} [\max_{t=0}^{T} I(\pi(t)) - \min_{t=0}^{T} I(\pi(t))]$$

Feature Embedding

Learn to augment

$$\begin{bmatrix} \mathcal{A}_{11}, & \mathcal{A}_{12}, & \mathcal{A}_{13}, & \mathcal{A}_{14} \\ \mathcal{A}_{21}, & \mathcal{A}_{22}, & \mathcal{A}_{23}, & \mathcal{A}_{24} \\ \mathcal{A}_{31}, & \mathcal{A}_{32}, & \mathcal{A}_{33}, & \mathcal{A}_{34} \\ 0, & 0, & 0, & 1 \end{bmatrix}$$

Learn to compare

a) Generating Similarity Maps
b) Learning to Re-weight:
c) Learning to Merge

Methods

Loss Function:

Classification Loss:

$$\mathcal{L}_{cls} = \frac{1}{B \times C} \sum_{i=1}^{B} \sum_{j=1}^{C} (P(i,j) - \mathbb{I}(y_i == y_j))^2$$

Attention Loss:

$$\mathcal{L}_{att} = \frac{S_{att}}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} |\alpha(i,j)|$$

Augmentation Loss:

$$\mathcal{L}_{aug} = \frac{S_{aug}}{B \times C} \sum_{i=1}^{B} \sum_{j=1}^{C} \sum_{k=1}^{K} (f(I_q) - f_k(I_s))^2 \mathbb{I}(y_i == y_j)$$

Total Loss:

$$\mathcal{L} = \mathcal{L}_{cls} + \lambda_{att} \cdot \mathcal{L}_{att} + \lambda_{aug} \cdot \mathcal{L}_{aug}$$

Experiments

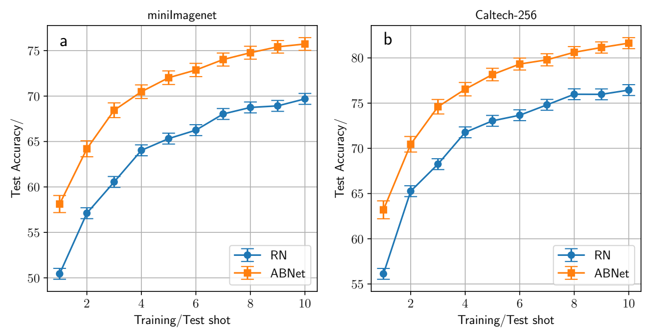
Compare with SOTA:

Method		miniImageNet		Caltech-256		tieredImageNet	
		5way1shot	5way5shot	5way1shot	5way5shot	5way1shot	5way5shot
MatchingNet [16]	NIPS'16	43.56±0.84	55.31±0.73	45.59±0.77	54.61±0.73	54.02	70.11
MetaLSTM [21]	ICLR'17	43.44±0.77	60.60±0.71	-	-	-	-
MAML [20]	ICML'17	48.70±0.84	55.31±0.73	48.09±0.83	57.45±0.84	51.67±1.81	70.30±0.08
MetaNet [17]	ICML'17	49.21±0.96	-	-	-	-	-
ProtoNet [12]	NIPS'17	49.42±0.87	68.20±0.70	-	-	54.28±0.67	71.42±0.61
RelationNet [13]	CVPR'18	50.44±0.82	65.32±0.77	56.12±0.94	73.04±0.72	54.48±0.93	71.32±0.78
CTM [42]	CVPR'19	41.62	58.77	-	-	-	-
Spot&Learn [37]	CVPR'19	51.03±0.78	67.96±0.71	-	-	-	-
MetaOptNet [43]	CVPR'19	52.87±0.57	68.76±0.48	-	-	54.71±0.67	71.79±0.59
ABNet		**58.12±0.94**	**72.02±0.75**	**63.20±0.99**	**78.42±0.69**	**62.10±0.96**	**75.11±0.78**

Resnet Backbone:

Method		Backbone	miniImageNet	
			5way1shot	5way5shot
RelationNet [13]	CVPR'18	ResNet-18	58.21	74.29
CTM [42]	CVPR'19	ResNet-18	62.05±0.55	78.63±0.06
MetaOptNet [43]	CVPR'19	ResNet-12	62.64±0.61	78.63±0.46
ABNet		ResNet-18	**63.15±0.63**	**78.85±0.56**

Results with different shots:



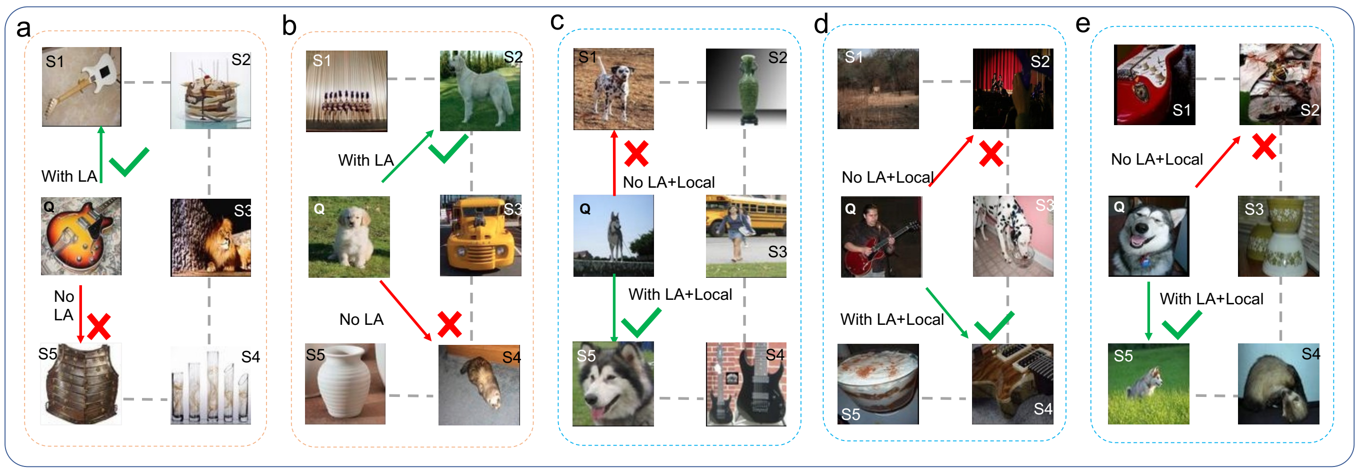
Ablation Study

Quantitative comparison of four variants of our methods:

- Baseline: fixed global features
- Baseline + LA (Learn to Augment)
- Baseline + LA + SP (Salient Patch)
- Baseline + LA + SP + LR (Learn to Re-weight)

Model	miniImageNet 5way1shot	miniImageNet 5way5shot
Baseline	52.44±0.91	66.50±0.77
Baseline+LA	54.27±0.91	68.02±0.77
Baseline+LA+SP	55.76±0.89	69.70±0.72
Baseline+LA+SP+LR	**58.12±0.94**	**72.02±0.75**

Visualization Comparison:



25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
ICPR 2020
Milan, Italy 10 | 15 January 2021

Augmented Bi-path Network for Few-shot Learning

Baoming Yan[1,2*], Chen Zhou[1*], Bo Zhao[3], Kan Guo[2], Jiang Yang[2], Xiaobo Li[2], Ming Zhang[1] and Yizhou Wang[1]

[1] Peking University, [2] Alibaba Group, [3] The University of Edinburgh

(* Indicates equal contribution)

Poster #2318 @PS T3.10