Semantics-guided Data Hallucination for Few-shot Visual Classification

Chia-Ching Lin\textsuperscript{1,2} Yu-Chiang Frank Wang\textsuperscript{1} Chin-Laung Lei\textsuperscript{1} Kuan-Ta Chen\textsuperscript{2}

\textsuperscript{1}Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan
\textsuperscript{2}Institute of Information Science, Academia Sinica, Taipei, Taiwan

INTRODUCTION

- Few-shot learning (FSL): only few samples would be available for selected object categories during learning
  - Base classes: each has a sufficient amount of training samples
  - Novel classes: each has a limited amount of training samples
- Data hallucination: generate additional training samples for novel classes based on intra-class variation learned from base classes

PROPOSED MODEL

- Training
  - Stage 1 (representation learning on $D_{\text{base}}$): train the feature extractor $\phi$ and the hallucinator $H$ by meta-learning
  - Stage 2 (few-shot learning on $D_{\text{base}} \cup D_{\text{novel}}$): use the trained $H$ to augment additional training samples for $C_{\text{novel}}$ and use the augmented dataset to train the final classifier $V$

EXPERIMENTS

- Top-5 accuracy of $V$
  - CIFAR-100: $|C_{\text{fin base}}| = 30, |C_{\text{fin novel}}| = 20$

NOTATIONS

- $D_{\text{base}}$: dataset of base classes $C_{\text{base}}$
- $D_{\text{novel}}$: dataset of novel classes $C_{\text{novel}}$
- Both datasets consist of tuples $((x_i, y_i, R_i))$
  - $x_i$: the $i$-th image
  - $y_i$: the one-hot label vector
  - $R_i$: the semantic information associated with $y_i$
- CIFAR-100: a word embedding vector of the label name
- Animals with Attributes: an attribute vector

CONCLUSION

- We incorporate semantic information into the data hallucination process to generate additional training data that exhibit semantics-oriented modes of variation for improved FSL performances