Δ-encoder: An effective sample synthesis method for few-shot object recognition

Eli Schwartz*, Leonid Karlinsky*, Joseph Shtok, Sivan Harary, Mattias Marder, Abhishek Kumar, Rogerio Feris, Raja Giryes, Alex M. Bronstein
Who’s that dog?
Key idea – training

• The model is a variant of an auto-encoder operating in feature space
Key idea – training

• The model is a variant of an auto-encoder operating in feature space
• The network learns to encode the delta between the reference and the target image
Key idea – training

• The model is a variant of an auto-encoder operating in feature space
• The network learns to encode the delta between the reference and the target image
• This delta is used to recover the target image as a (non-linear) combination of the reference and the delta
Key idea – synthesizing
Key idea – synthesizing

• At test time we sample encoded deltas from random training image pairs
Key idea – synthesizing

- At test time we sample encoded deltas from random training image pairs
- The sampled deltas are used to create samples for new classes by combining them with the new class reference examples
- These samples are used to train a classifier for the new category
Few-shot classification experiments

- **miniImageNet**: 58.5 (previous SOA) → 59.9 (ours)
- **CIFAR-100**: 63.4 (previous SOA) → 66.7 (ours)
- **Caltech-256**: 63.8 (previous SOA) → 73.2 (ours)
- **CUB**: 69.6 (previous SOA) → 69.8 (ours)
Few-shot classification experiments

**miniImageNet**: 58.5 (previous SOA) $\rightarrow$ 59.9 (ours)
**CIFAR-100**: 63.4 (previous SOA) $\rightarrow$ 66.7 (ours)
**Caltech-256**: 63.8 (previous SOA) $\rightarrow$ 73.2 (ours)
**CUB**: 69.6 (previous SOA) $\rightarrow$ 69.8 (ours)
Real vs synthetic examples ablation study

![Graph showing accuracy vs number of generated samples for 1-shot and 5-shot accuracy, along with nearest-neighbor baseline accuracies.](image-url)
Thank you for listening!

Please meet us at our poster:

Poster Session A: 10:45 AM - 12:45 PM
@Room 210 & 230 AB #25