Fine-Grained Generalized Zero-Shot Learning via Dense Attribute-Based Attention

Dat Huynh and Ehsan Elhamifar

Khoury College of Computer Sciences
Northeastern University
Motivation

- **Fine-Grained Recognition:**
  - Recognize visually similar objects \(\rightarrow\) **Must localize attributes**
  - Experts annotation is **costly**

- **Zero-Shot Learning:**
  - Generalize to **unseen classes** using annotations from **seen classes**
  - Existing methods **cannot localize** attributes without bbox annotations

- **Contributions:**
  - **Dense Attention:** capture attribute details
  - **Attention Embedding:** transfer attribute details
Proposed Architecture

- **Encode attributes into multiple features**

- **Dense Attention** has **multiple soft-attention channels**:
  - Output **attribute features** $h^a$
  - Guide by **attribute semantic** $\mathcal{U}_a$ (word2vec of the attribute name)

- **Attribute Embedding**:
  - Compute each **attribute score** $e^a$ from attribute feature
  - Compute **class score** $s^c$ as sum of attribute scores
Experiments

• **Outperform SOTA** on ZSL and Generalized ZSL

• **Dense Attention** localizes present attributes and finds evidence for absent attributes