Fine-Grained Recognition as HSnet Search for Informative Image Parts

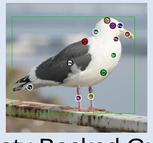
Michael Lam, Behrooz Mahasseni and Sinisa Todorovic

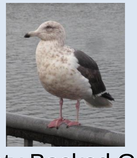
Oregon State University

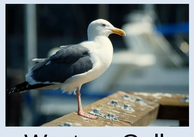
IEEE 2017 Conference on Computer Vision and Pattern Recognition
CVPR 2017 July 21-26 HONOLULU

Problem: Fine-Grained Recognition

Goal: Classify and detect parts


Slaty Backed Gull


Slaty Backed Gull


Western Gull

Challenges

- Same classes with different appearances
- Different classes with similar appearances
- **Background clutter can hurt recognition**

Key Idea: Search for Informative Parts

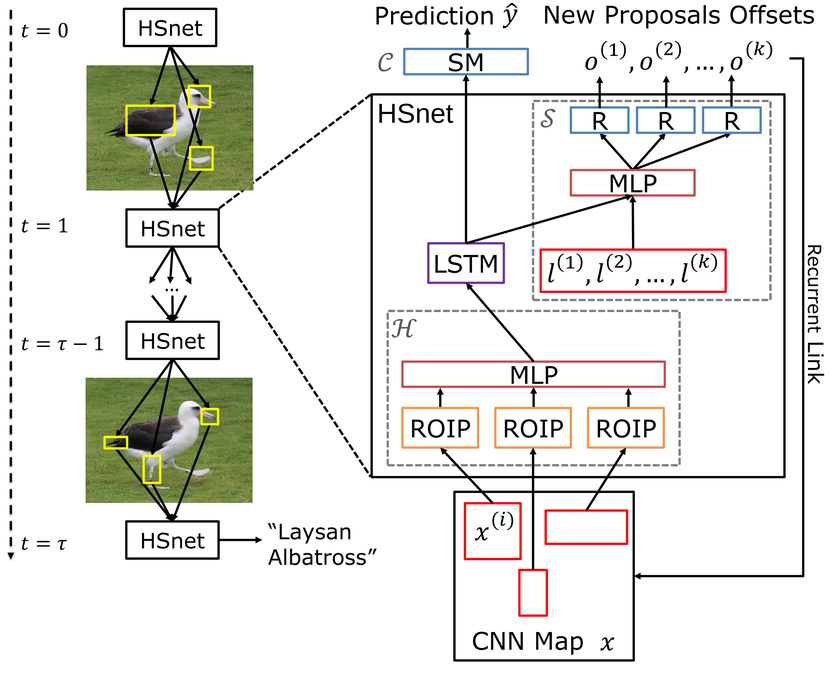
- Focus attention on informative parts
- Sequential search to find informative parts

→ **HSnet Search Framework**

Supervised vs. Weakly Supervised

- Part annotations available: train to minimize distance between predicted part locations and ground truth part locations
- Part annotations not available: use determinantal point process (DPP) to regularize bounding box positions for diverse candidates

Search Overview



HSnet Architecture

ROIP: Region of Interest Pooling
MLP: Multilayer Perceptron
SM: Softmax
R: Regression

x: CNN feature map
$x^{(i)}$: bounding box i features
$l^{(i)}$: bounding box i location
$o^{(i)}$: bounding box i offset
$\hat{y}$: class prediction

Acknowledgments: DARPA XAI, NSF RI1302700, NSF GRFP 1314109-DGE

Accuracy [%] on Birds and Cars

Caltech UCSD 2011 Birds

Method	Annotations	Accuracy
Krause et al. 2015	GT+BB	82.8
Jaderberg et al. 2015	GT	84.1
Xu et al. 2015	GT+BB+ parts+web	84.6
Lin et al. 2015	GT+BB	85.1
B1	GT	82.3
B2	GT+parts	83.1
B3	GT+parts	86.2
B4	GT+parts	85.7
Hsnet	**GT+parts**	**87.5**

Stanford Cars 196

Method	Annotations	Accuracy
Deng et al. 2013	GT+BB	63.6
Krause et al. 2013	GT+BB	67.6
Krause et al. 2014	GT+BB	73.9
Lin et al. 2015	GT	91.3
Krause et al. 2015	GT+BB	92.6
B1	GT	88.5
B4	GT	92.2
HSnet	**GT**	**93.9**

Qualitative Results

Parts Detection at Time Steps

Weak Supervision Learns Informative Parts on Cars