# Zero-shot recognition with unreliable attributes

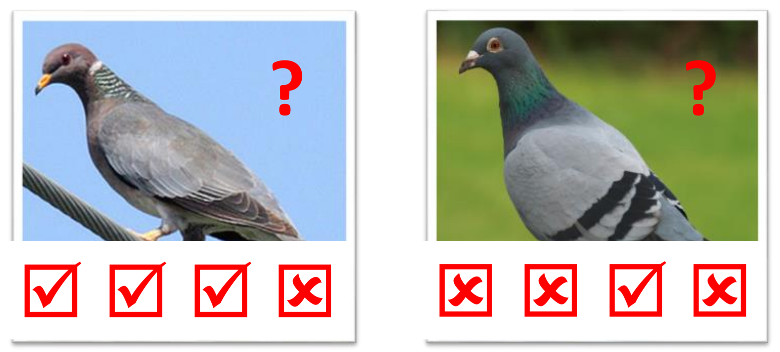
Dinesh Jayaraman and Kristen Grauman
UT Austin

## Zero-shot category recognition with attributes

**Given**: attribute classifiers, category-attribute signatures

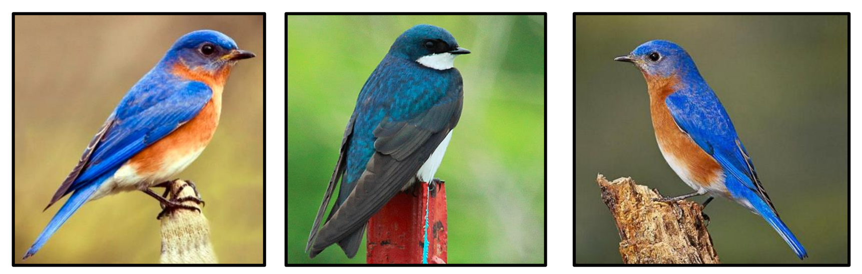
**How to identify a band-tailed pigeon:**

**Attribute signature:**
- ☑ White collar
- ☑ Yellow feet
- ☑ Yellow bill
- ☒ Red breast

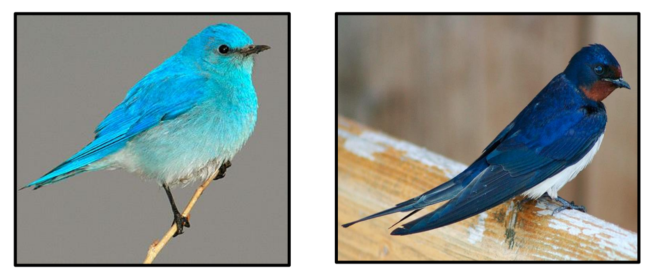



## The catch: unreliable attribute classifiers

**Training positives** ("blue back") — **Test input**



Problem 1: weak supervision



Problem 2: unseen categories

## Prior approaches: ***ignore*** unreliability

**Given:**

Ground truth attributes → F(.) (known function) → Ground truth objects

based on category-attribute signatures

**Standard framework (e.g., Direct Attribute Prediction, Lampert '09):**

Unreliable attribute predictions → F(.) (same function) → Object predictions

"soft" predictions e.g. MAP

## Our key idea: ***account*** for unreliability

Attribute error characteristics → learned function

Unreliable attribute predictions → G(.) ≈ F(.) → Object predictions

## Approach overview

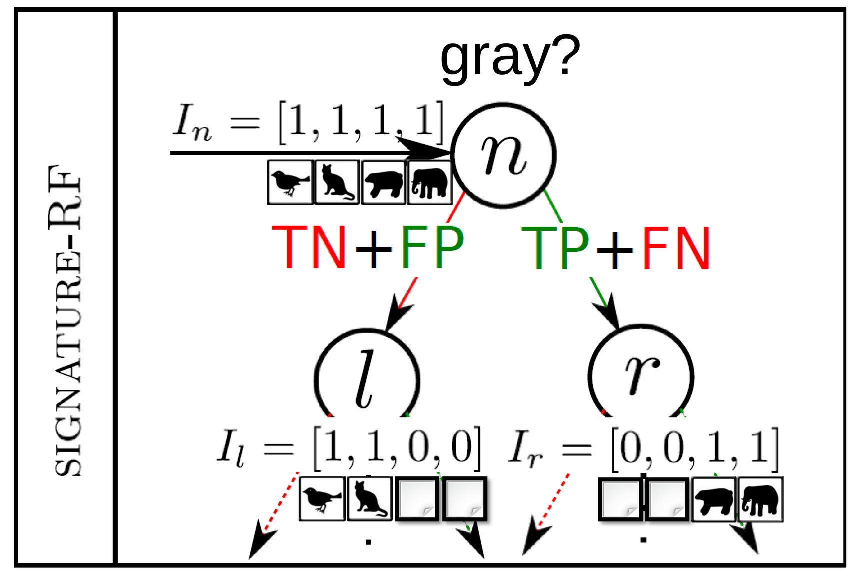
- **Random forests** trained on category-attribute signatures.
- Learning approach exploits **attribute classifier ROC curves**.
- **Fractional samples** to emulate estimated test distribution.
- Selected node splits are both **discriminative *and reliable***.

## Step 1: Train attribute classifiers

Train SVMs for M attribute classifiers on attribute-labeled data $\mathcal{D}_T$

## Step 2: Build 1-vs-rest random forest for each category k

### Signature random forest: ignore attribute unreliability




Assuming ideal classifiers

**To select at each node**:

$(\text{attribute } m, \text{threshold } t)$

**Category presence indicators:**

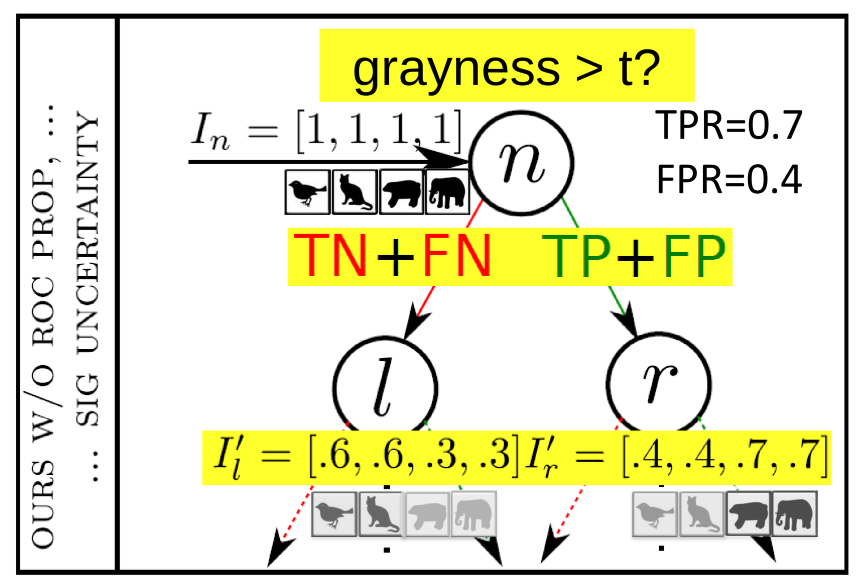
$$I_r(k) = \begin{cases} 1, & \text{if } A_k(m) > t \text{ and } I_n(k) = 1 \\ 0, & \text{otherwise} \end{cases}$$

$$I_l(k) = 1 - I_r(k)$$

**Information gain criterion:**

$$H(p_{I_n}) - \left( \frac{\|I_l\|_1}{\|I_n\|_1} H(p_{I_l}) + \frac{\|I_r\|_1}{\|I_n\|_1} H(p_{I_r}) \right)$$

### Idea #1: Attribute ROC-guided fractional samples




Modeling errors

**Set aside 20% attribute-labeled data:** $\mathcal{D}_T$ | $\mathcal{D}_V$

**Measure attribute prediction error:**

$$\text{TPR}(m,t) = \text{true positive rate on } \mathcal{D}_V$$
$$\text{FPR}(m,t) = \text{false positive rate on } \mathcal{D}_V$$

**Fractional sample propagation:**

$$I'_r(k) = \begin{cases} I'_n(k) \times \text{TPR}(m,t), & \text{if } A_k(m) = 1 \\ I'_n(k) \times \text{FPR}(m,t), & \text{if } A_k(m) = 0 \end{cases}$$

$$I'_l(k) = 1 - I'_r(k)$$

### Idea #2: Node-specific attribute error statistics

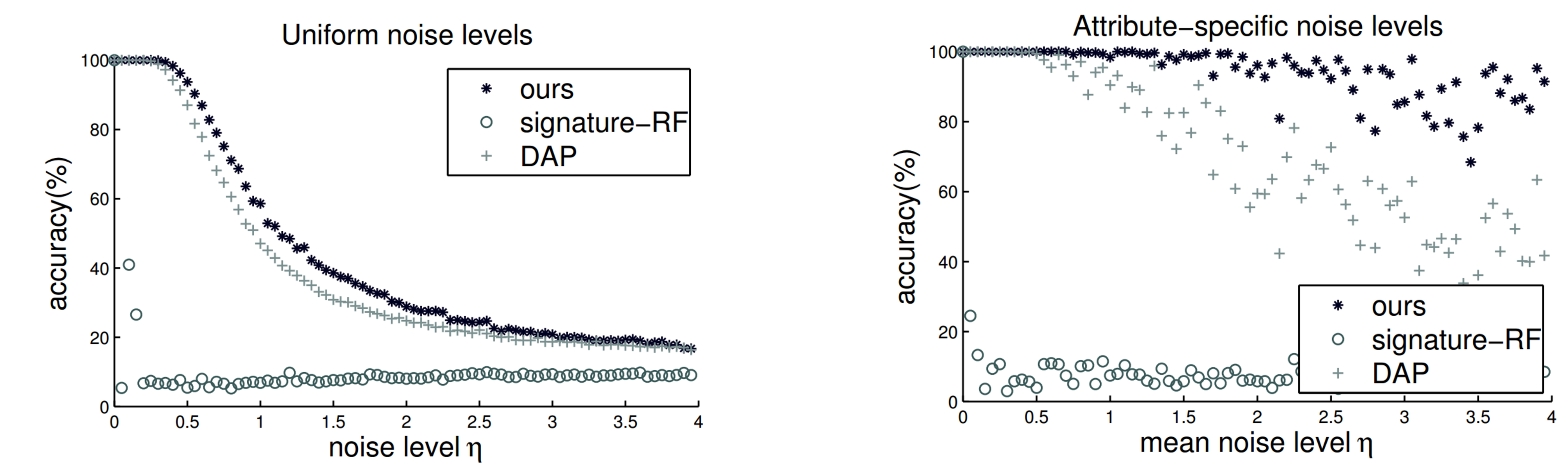
- **Validation data propagation:** Node-specific attribute validation data models test distribution better: $\mathcal{D}_V \rightarrow \mathcal{D}_V(n)$
- **Node-specific error rates**: $\text{TPR}(m,t) \rightarrow \text{TPR}(n,m,t)$ *etc.*

## Extensions

- **Few-shot learning**: Information gain criterion redefined as weighted sum of zero-shot gain and standard gain:
  $$IG_{few}(m,t) = \lambda\, IG_{zero}(m,t)\{A_1,\ldots,A_K\} + (1-\lambda)\, IG_{basic}(m,t)\{\mathcal{D}_T\}$$
- **Unreliability in category-attribute signatures:** handled with an extra probability term in child node indicator vector definition.

## Experiments

**Synthetic unreliable classifier predictions:**




Gains from (1) reliable attribute selection, (2) modeling unreliability

**Real datasets:**

AwA (animals) — aPY (objects) — SUN (scenes)

**Dataset details:**

| | AwA | aPY | SUN |
|---|---|---|---|
| # attributes | 85 | 65 | 102 |
| # unseen cls | 10 | 12 | 10 |
| # seen cls | 40 | 20 | 707 |
| # images | 30475 | 15339 | 14340 |

**Comparison to prior art (AwA):**

| Method | Accuracy |
|---|---|
| Lampert, CVPR '09 | 40.5 |
| Yu, ECCV '10 | 40.0 |
| Rohrbach, CVPR'10 | 35.7 |
| Kankuekul, CVPR '12 | 32.7 |
| Yu, CVPR '13 | 48.3 |
| OURS (named attributes) | $43.0 \pm 0.07$ |
| OURS (discovered attributes) | $\mathbf{48.7 \pm 0.09}$ |

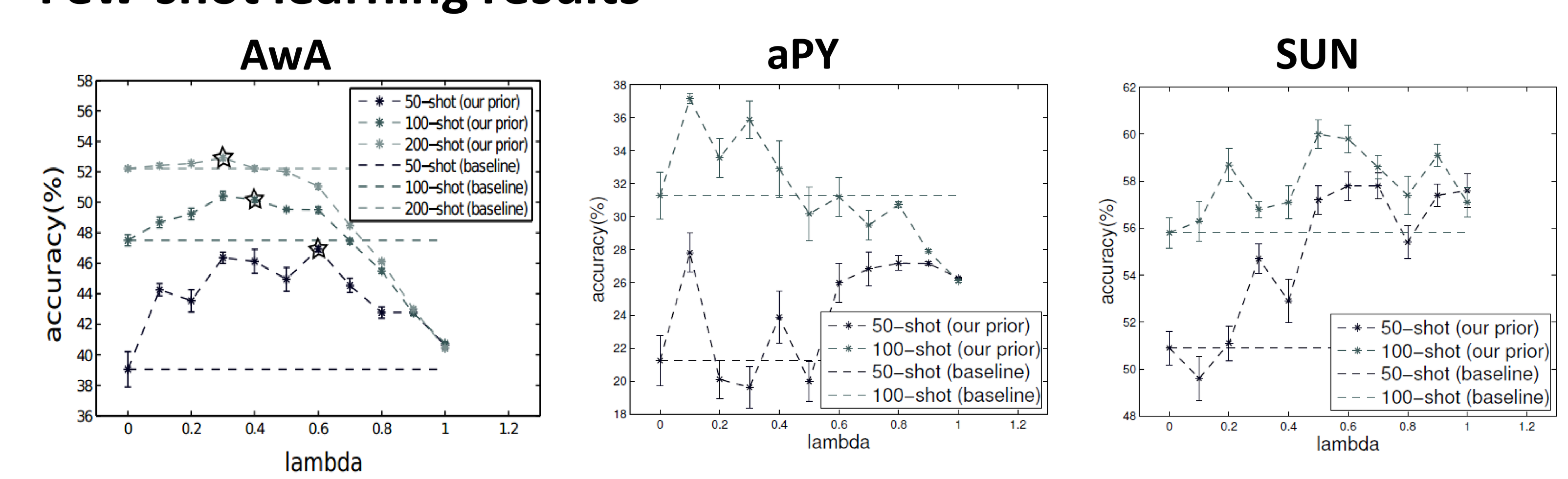
Quantifying attribute prediction unreliability *even more important* than training better attribute predictors!

**Ablation studies**

| Method/Dataset | AwA | aPY | SUN |
|---|---|---|---|
| DAP | 40.50 | 18.12 | 52.50 |
| SIGNATURE-RF | $36.65 \pm 0.16$ | $12.70 \pm 0.38$ | $13.20 \pm 0.34$ |
| OURS W/O ROC PROP, SIG UNCERTAINTY | $39.97 \pm 0.09$ | $24.25 \pm 0.18$ | $47.46 \pm 0.29$ |
| OURS W/O SIG UNCERTAINTY | $41.88 \pm 0.08$ | $24.79 \pm 0.11$ | $\mathbf{56.18 \pm 0.27}$ |
| OURS | $\mathbf{43.01 \pm 0.07}$ | $\mathbf{26.02 \pm 0.05}$ | $\mathbf{56.18 \pm 0.27}$ |
| OURS+TRUE ROC | $54.22 \pm 0.03$ | $33.54 \pm 0.07$ | $66.65 \pm 0.31$ |

Each component contributes significantly to overall gain

**Few-shot learning results**




Our method builds strong priors for knowledge transfer