Zero-Shot Learning

Venkatesh Saligrama
Boston University
Joint Work with Greg Castanon, Joe Wang, Yuting Chen

BOSTON UNIVERSITY

Outline

- Conventional (Supervised) Machine Learning:
  - Large amount of training data required to train high accuracy classifiers.
- Challenge
  - Diverse range of objects, object attributes (size, materials, chemistry, composition).
  - Very few (or negligible) positive examples for many scenarios. Data collection for all these scenarios is clearly infeasible or impractical.
- Approach: Zero-Shot Learning
  - How to learn classifiers for new classes for which you have no (training) data?
- Relevance to TSA:
  - Luggage inspection: homemade explosives
    - New classes of threats for which we don't have parametric models/samples
    - Variations: chemical formula, concentration, processes
    - Discovery of new explosive classes and how to relate to what seen before
  - Video forensics: suspicious activity detection…
- How does it work? Identify latent structural thematic properties of known classes
  - Predict classifiers for new classes based on how threats manifest in latent space

Supervised (conventional) Learning

- Conventional Learning
  - Training Data
    - Images → Class-Labels
    - Xray images → Threat/non-threat
    - Video → what activity
  - Learning Problem
    - Train classifier with training data
    - Accurate prediction of class-labels for new images during test-time

class: horse

class: elephant

New Sample ≈ Old Sample

x → Classifier f(x) → y

??

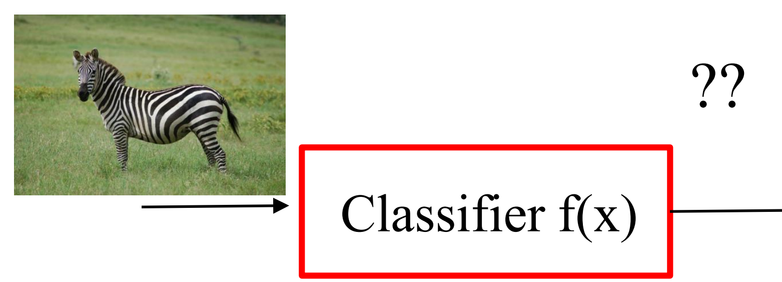
Zero-Shot Learning

- Zero-Shot Learning
  - Training Data (x,y)
    - Labeled images of Horses, elephants
    - Existing Explosive/Non-Explosive data
    - Video: Existing Activity Classes
  - Learning Problem:
    - Learn a classifier for new classes that not seen in training data.
    - Zebra class, New Explosives, New suspicious activity…
  - Traditional concept makes no sense

horse

elephant

New Sample $\not\approx$ Old Sample



Zebra is not seen before: How to minimize error for things not seen before

Airport Security Context

- Millions of types of homemade threats:
  - Fine grained classification

- Myriad Scanner Outputs

Key Idea: Leverage structure in descriptions

Source domain

Target domain

Horse	
Kingdom:	Animalia
Phylum:	Chordata
Class:	Mammalia
Order:	Perissodactyla
Family:	Equidae
Genus:	*Equus*
Species:	*E. ferus*
Subspecies:	***E. f. caballus***

Elephants	
Kingdom:	Animalia
Phylum:	Chordata
Subphylum:	Vertebrata
Class:	Mammalia
Superorder:	Afrotheria
Order:	Proboscidea
Family:	Elephantidae
	Gray, 1821

Seen classes

Zebra	
Kingdom:	Animalia
Phylum:	Chordata
Class:	Mammalia
Order:	Perissodactyla
Family:	Equidae
Genus:	*Equus*
Subgenus:	***Hippotigris*** and ***Dolichohippus***

Domestic dog	
Kingdom:	Animalia
Phylum:	Chordata
Class:	Mammalia
Order:	Carnivora
Family:	Canidae
Genus:	*Canis*
Species:	*C. lupus*
Subspecies:	***C. l. familiaris***

Unseen classes

?

What if we are given thematic information during training?
Can we recognize new class from thematic information?

Key Idea: Reduction to Standard Binary Classification

- View attributes/themes (d) and image (x) as two pieces of puzzle
  - Predict whether or not they are associated

A zebra is an animal that looks like a horse. It has stripes like a tiger does. It has black and white stripes on its body.

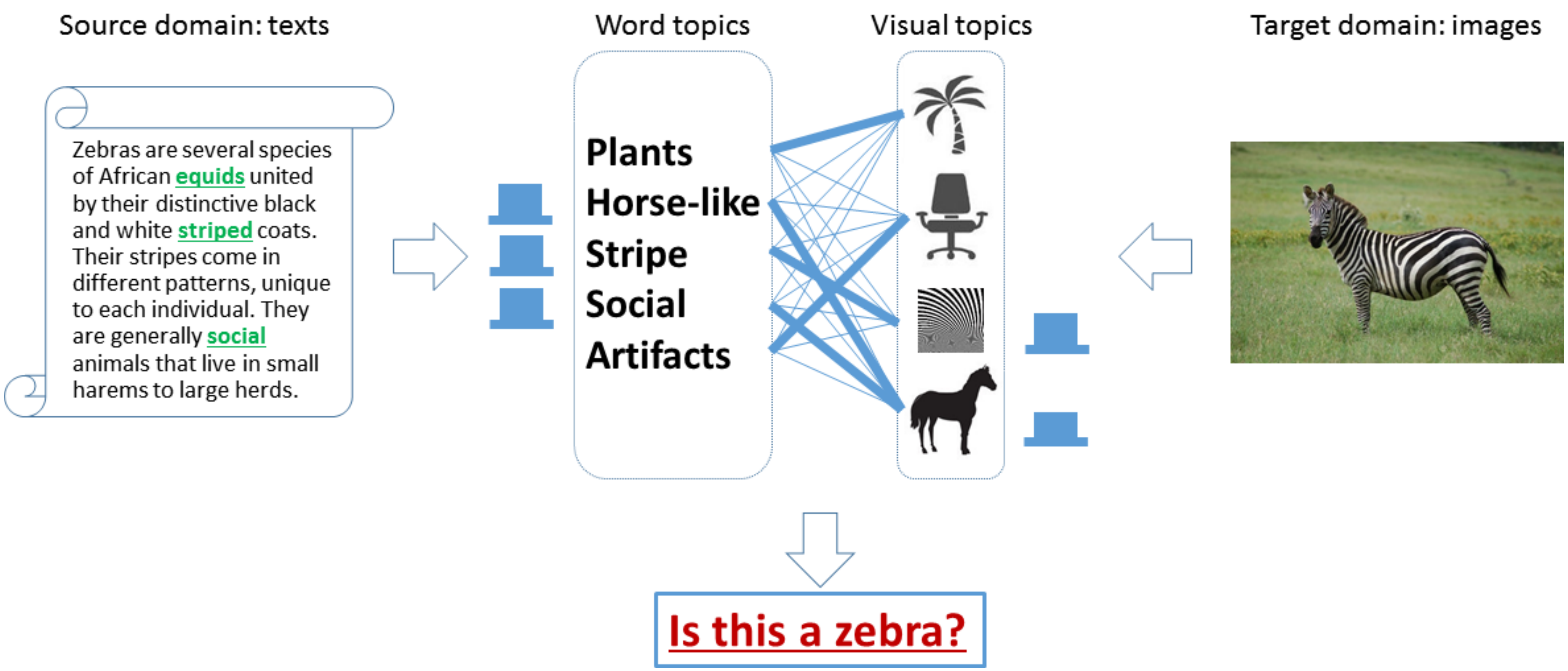
$\boldsymbol{x}_i$ → Classifiers f(x,d) ← $\boldsymbol{d}_j$

No match to image

Yes, description matches image

With thematic info we can pose it as conventional learning with unconventional outputs for classifiers.

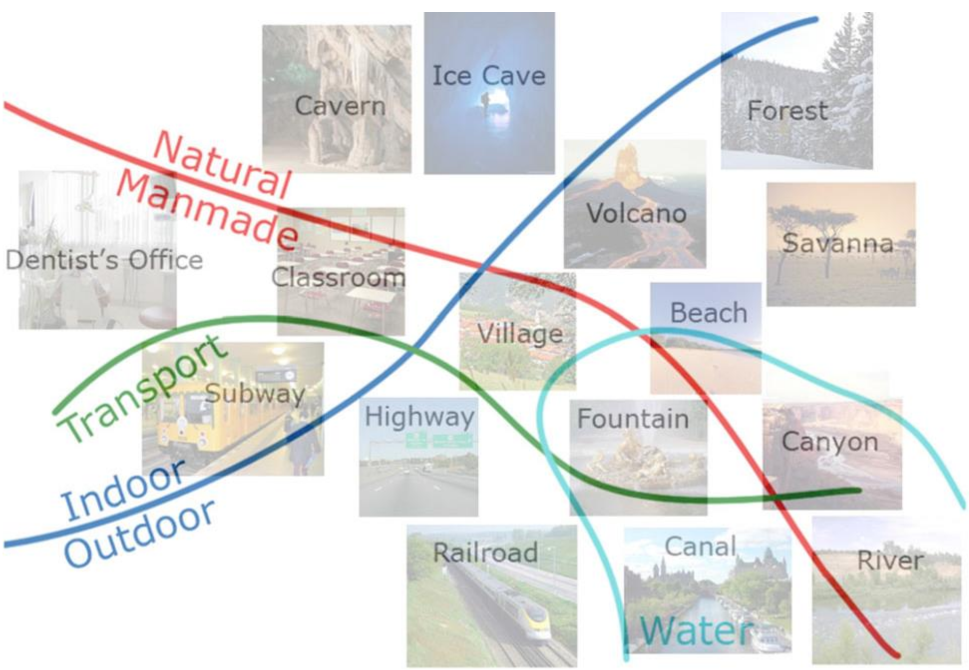
Key Idea 2: Latent Topic Model



What if themes/attributes are unknown?
Can we infer these themes from generic information about other classes?

Experiments: Benchmark datasets

Dataset	# instances	# attributes	# seen/unseen classes
aP&Y	15,339	64 (continuous)	20 / 12
AwA	30,475	85 (continuous)	40 / 10
CUB-200-2011	11,788	312 (binary)	150 / 50
SUN Attribute	14,340	102 (binary)	707 / 10



Performance Comparison

Method	**aP&Y**	**AwA**	**CUB-200-2011**	**SUN Attribute**	**Average**
Akata et al. **CVPR’15**	-	61.9	40.3	-	-
Lampert et al. **PAMI’14**	38.16	57.23	-	72.00	-
R.-Paredes and Torr **ICML’15**	24.22±2.89	75.32±2.28	-	82.10±0.32	
SSE, **ICCV’15**	46.23±0.53	76.33±0.83	30.41±0.20	82.50±1.32	58.87
SDL, **arXiv’15**	***50.35±2.97***	***79.12±0.53***	***41.78±0.52***	***83.83±0.29***	***63.77***

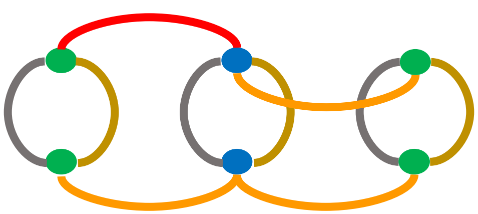
Zero Shot Inference

User @RealUser · 10h

Going to give Tom his backpack

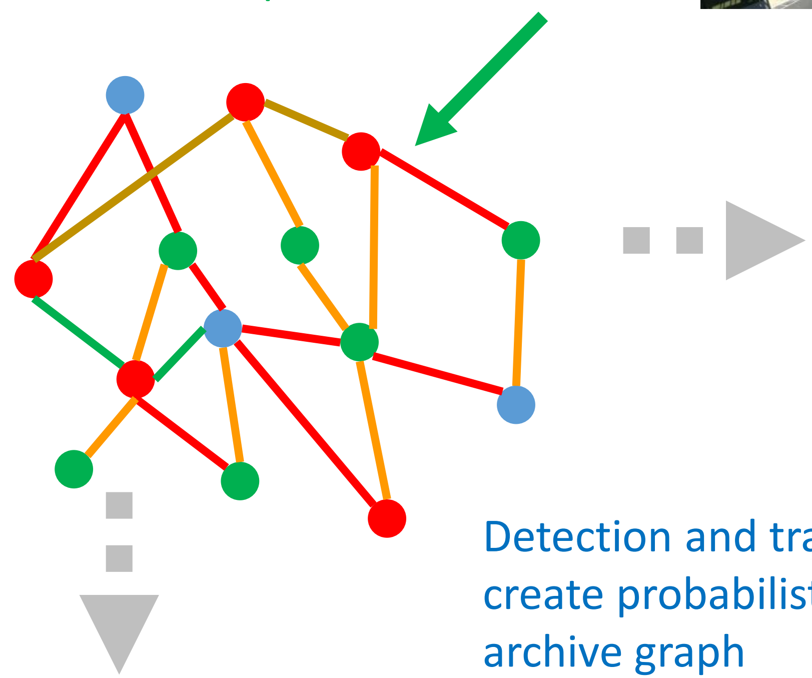
Semantic Gap

Represent archive



Semantic Query Graph

Problem reduced to subgraph matching



Detection and tracking create probabilistic archive graph

Outline

- Conventional (Supervised) Machine Learning:
  - Large amount of training data required to train high accuracy classifiers.
- Challenge
  - Diverse range of objects, object attributes (size, materials, chemistry, composition).
  - Very few (or negligible) positive examples for many scenarios. Data collection for all these scenarios is clearly infeasible or impractical.

- Approach: Zero-Shot Learning
  - How to learn classifiers for new classes for which you have no (training) data?
- Intuition:
  - Leverage known classes to identify latent structural thematic properties of threats/non-threats. Match/Identify thematic properties of new classes.
- Relevance to TSA:
  - Luggage inspection: homemade explosives
    - New classes of threats for which we don’t have parametric models/samples
    - Variations: chemical formula, concentration, processes
    - Discovery of new explosive classes and how to relate to what seen before
  - Video forensics: suspicious activity detection…