Transfer Learning Driven Rapid Change Detection for Disaster Management

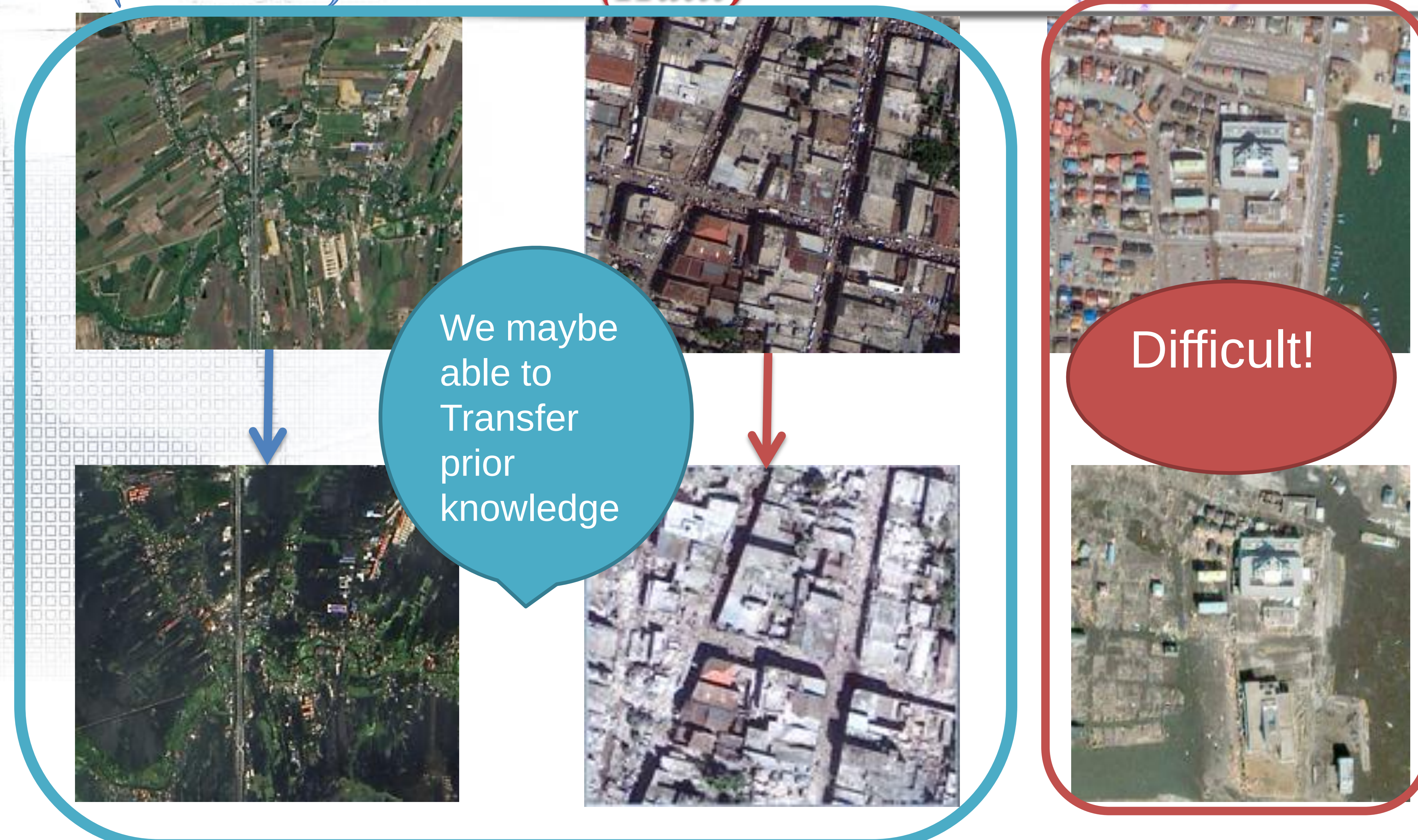
Surya S. Durbha
Associate Professor

Centre of Studies in Resources Engineering
Indian Institute of Technology Bombay

Outline

- Background
- Need for Transfer Learning
- Domain Adaptation methods
  - Approach by transfer learning
  - Case Study
- Results
- Conclusions

Floods (Thailand) **Earthquake (Haiti)** **Tsunami (Japan)**



(I). Change in Spectral Signature **(II). Change in Shape** **(III). Evolution of new class**

Contextually it is the same entity, the data distributions will be different.

Background

Post-Disaster data products

- Basic mapping or background space-map products
- Rapid damage assessment products
- Detailed damage assessment products

Techniques

- An interpretation technique applied to a single post-event data set;
- Change detection technique using a pre-post image pair with same sensor/same geometry;
- Change detection technique using pre- and post-event data from different sensors.
- Data fusion with existing pre-event GIS layers and in situ information

Damage Extent Identification

- Damage extent identification for buildings and built-up areas is relatively straightforward
  - e.g., by means of very-high-resolution (VHR) EO data
- In case of Complete collapse
  - Visible change in shape
  - Regular to less regular,
  - Minimal shadow effects
- More challenging for in complete collapse
- Level of destruction assessment is difficult to do **Quickly/Rapidly**.

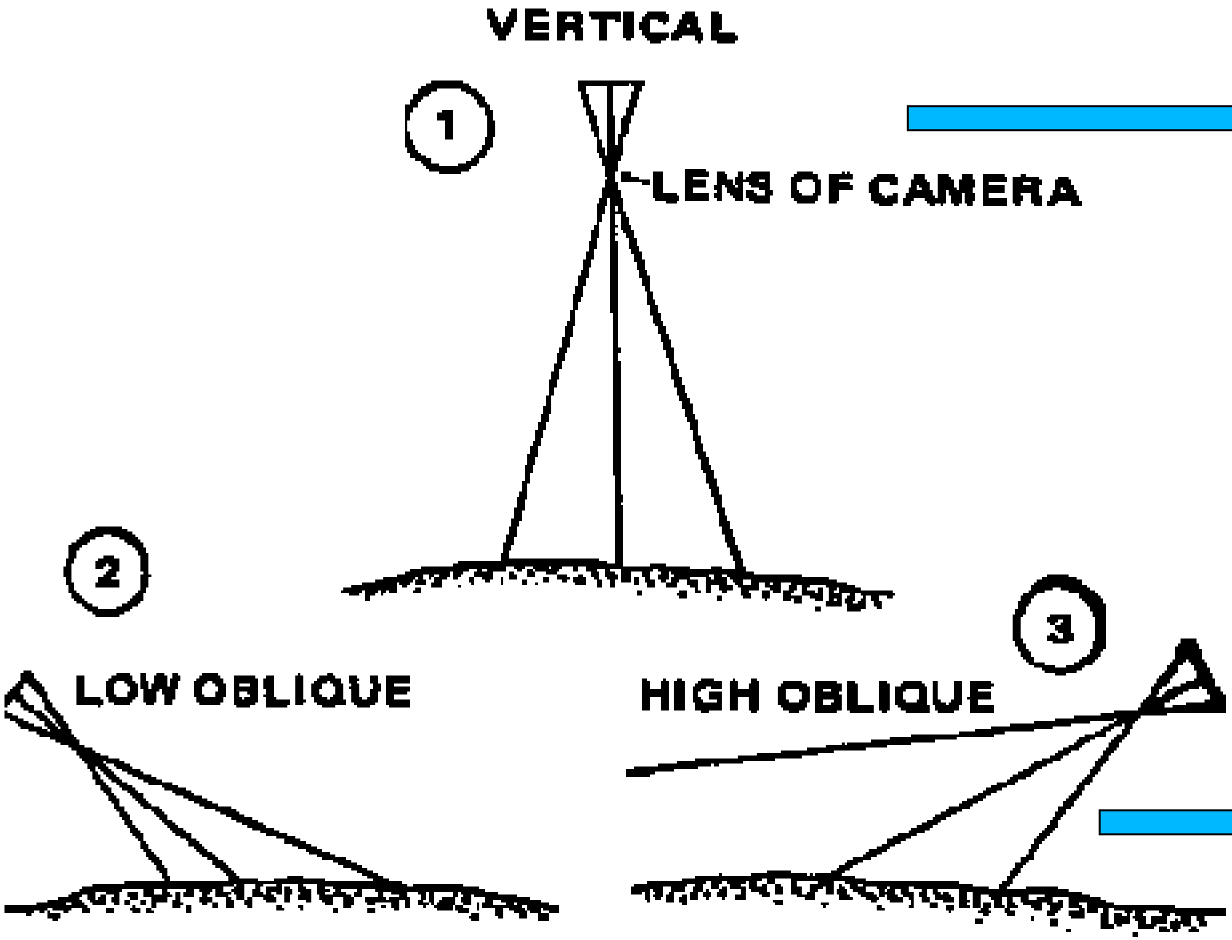
European Macroseismic Scale (EMS)

- **Grade 1:** Negligible to slight damage (no structural damage, slight non-structural damage)
- **Grade 2:** Moderate damage (slight structural damage, moderate non-structural damage)
- **Grade 3:** Substantial to heavy damage (moderate structural damage, heavy non-structural damage)
- **Grade 4:** Very heavy damage (heavy structural damage, very heavy non-structural damage)
- **Grade 5**: Destruction (very heavy structural damage).

- Studies show that it is difficult to construct an explicit one-to-one correspondence between the building damage grades from these and their appearances from remote sensing data.

- To assess different damage grades, varying resolutions of remote sensing images are analyzed to arrive at the damage levels.

- As a result several lower damage grades are aggregated as one grade for building damage classification in practice



- Mostly useful for complete collapse assessment

- Could be more useful in assessing structural building damage

http://irevolutions.org/2016/02/09/how-close/

A possible Scale

Class name	*short description*
totally collapsed	the building is totally destroyed and only debris are visible
first or mid-story collapsed	the building is collapsed, but the roof is totally or mostly untouched ("pancake" effect)
damaged roof	(most of) the walls are standing, but (most of) the roof is damaged
lightly damaged roof	the walls are standing, and a minor part of the roof is damaged
light damage	no visual change in the building roof, but debris are close to the walls

Fabio Dell'Acqua and Paolo Gamba (2012), Proceedings of IEEE,

This is what we can usually map using EO data

Problem: Paucity of Labeled samples after disaster

- Change detection studies use continuously updated information to update t… and the maps that are producec…
- Normally, after a disaster the la… classes evolve into classes whos… distribution may not be sam… disaster.
  - The two domains d… share exactly … land- cover cl…ses since
- Paucity of training samples due to
  - inaccessibility of the terrain
  - it is dangerous to venture into the disaster affected territory

Constrains the building of optimal models that can be used to classify the disaster affected areas.

http://www.satimagingcorp.com/galleryimages/digitalglobe-philippines-after.jpg

Crowdsource Mapping

- Crowdsourcing for mapping purposes requires the distribution of data to a restricted group or even to the general public, and this raises major issues with respect to timeliness and cost.
- practical results of this effort are questionable as for their accuracy, and certainly may not be worth the cost.
- Unless a commonly agreed scale is defined, and hopefully completed with many visual examples, the collaborative approach has big chances to be less conclusive than a single expert's (or expert pool's) analysis.
- However, it can give valuable Training samples for automated classification!
- For the proposed approach few accurate or pure samples are enough to run the TL-based classifier.

News | Sport | Comment | Culture | Business | Money | Life & style

News > Technology > Mapping technologies

Online volunteers map Philippines after typhoon Haiyan

Humanitarian OpenStreetMap Team coordinates mapping effort after enormous storm devastated country

Alex Hern
theguardian.com, Friday 15 November 2013 10.32 GMT
Jump to comments (5)

Tacloban City, on OpenStreetMap, after the typhoon hit. Photograph: /OpenStreetMap

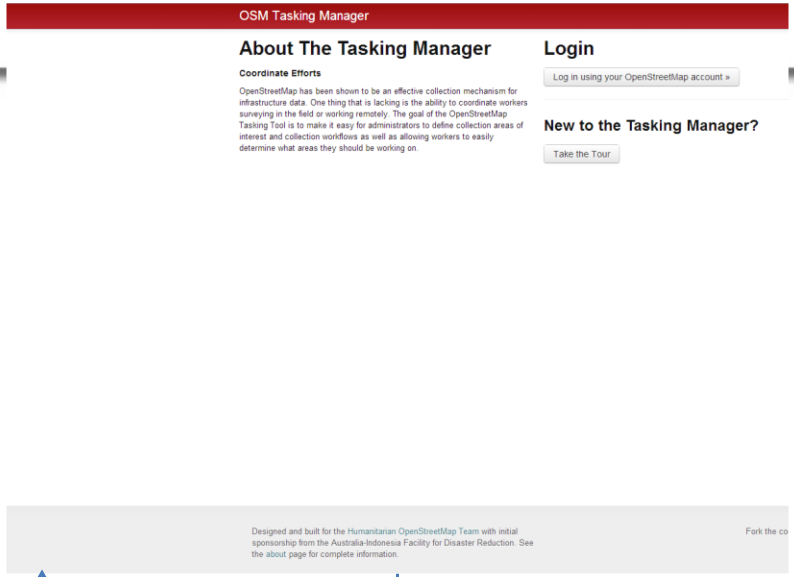
More than 700 volunteers have collaborated to provide rescue workers with high quality maps of areas in the Philippines hit by typhoon Haiyan.

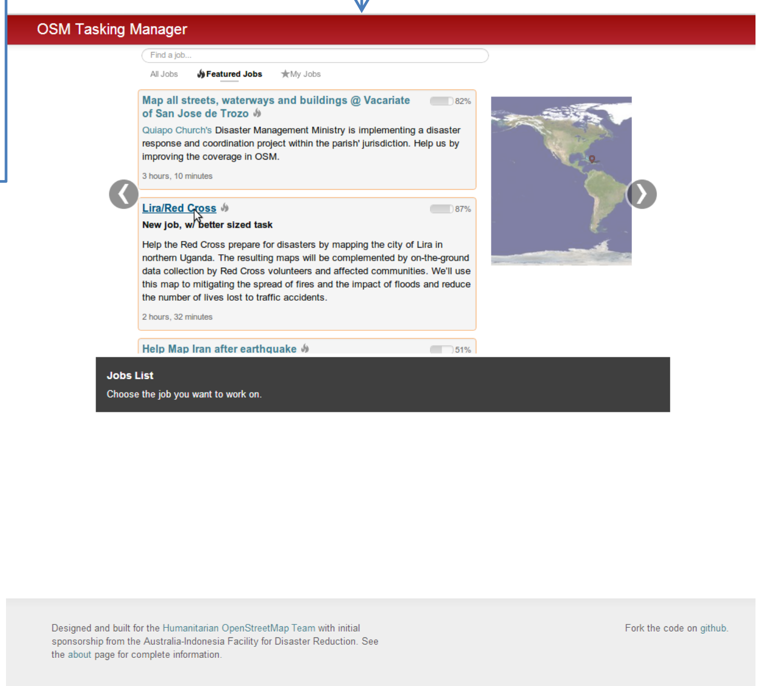
Working on OpenStreetMap, a collaboratively created map of the world – like Wikipedia, but for cartographers – the volunteers have made over 1.5m changes, providing information for humanitarian aid workers on the ground and updating maps to reflect damage from the storm.

The work is co-ordinated by the Humanitarian OpenStreetMap Team (HOT), a volunteer group which lets disaster relief workers set tasks for mapmakers at home. Users who want to help out can log-in to the tasking manager, where they are presented with a list of requests from the team.

Most of these are as simple as tracing the road network of an affected area from pre-existing satellite imagery. One asks users to trace the

http://www.theguardian.com/technology/2013/nov/15/online-volunteers-map-philippines-after-typhoon-haiyan





http://tasks.hotosm.org/tour

www.tomnod.com/nod/challenge/haiyantyphoon2013

tomnod Campaigns / Super Typhoon Haiyan

Sign In

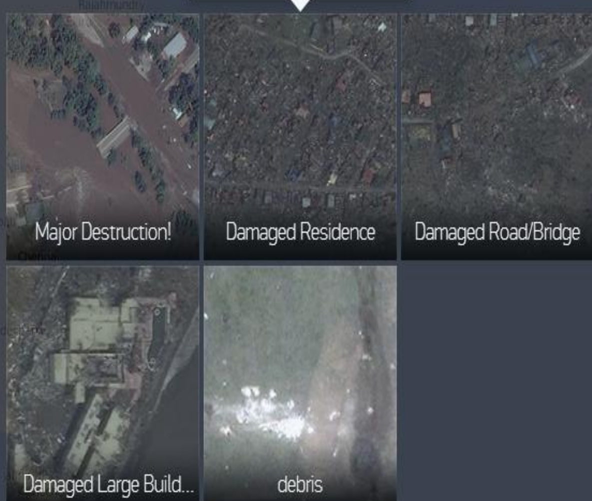
Super Typhoon Haiyan

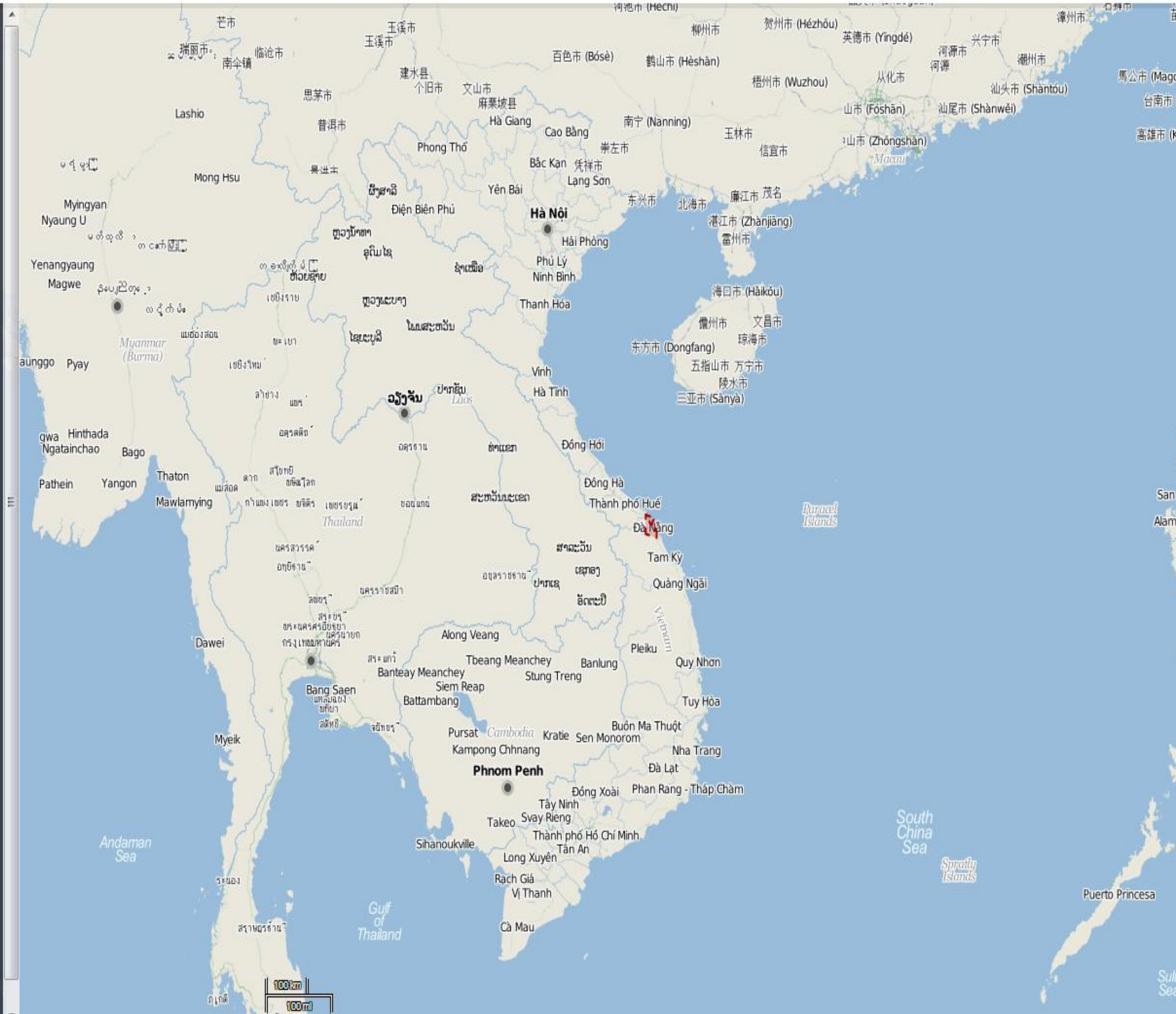
Start Tagging!

The Philippines 11/08/2013

Super Typhoon Haiyan made landfall in the Philippines late last week. The storm was one of the strongest ever recorded and devastated thousands. Help us map damage caused by the storm.

Help us tag... Destroyed Residential Area





How it Works Contact Us

Reuse what is already learned?

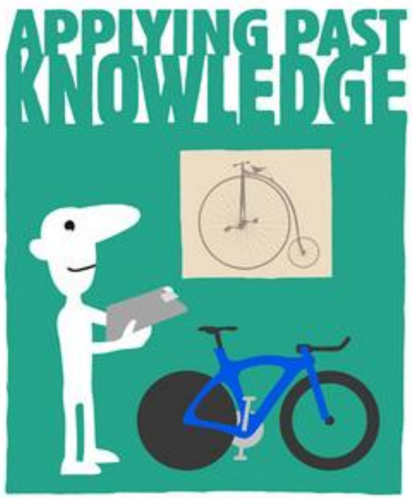
- Reduce the need and effort to collect labeled samples again
- Reuse the already available information on images acquired on the area of interest (source domain) to classify new images acquired on the area of interest (target domain).
- Transfer learning enables to develop models (using small training sample sizes) that can incorporate prior knowledge and adapt it to a new related domain or scenario thus facilitating the rapid retrieval of the affected areas.

Transfer Learning

• Humans always use past knowledge – What knowledge is relevant? – How can it be effectively leveraged?

"Transfer [learning] is a sequential process that influences the performance of learning through the reuse of structured knowledge [collected on previous tasks] and improves the behavior of the agent on new related tasks."

Pat Langley (Workshop on Structural Knowledge Transfer for Machine Learning, ICML 2006)

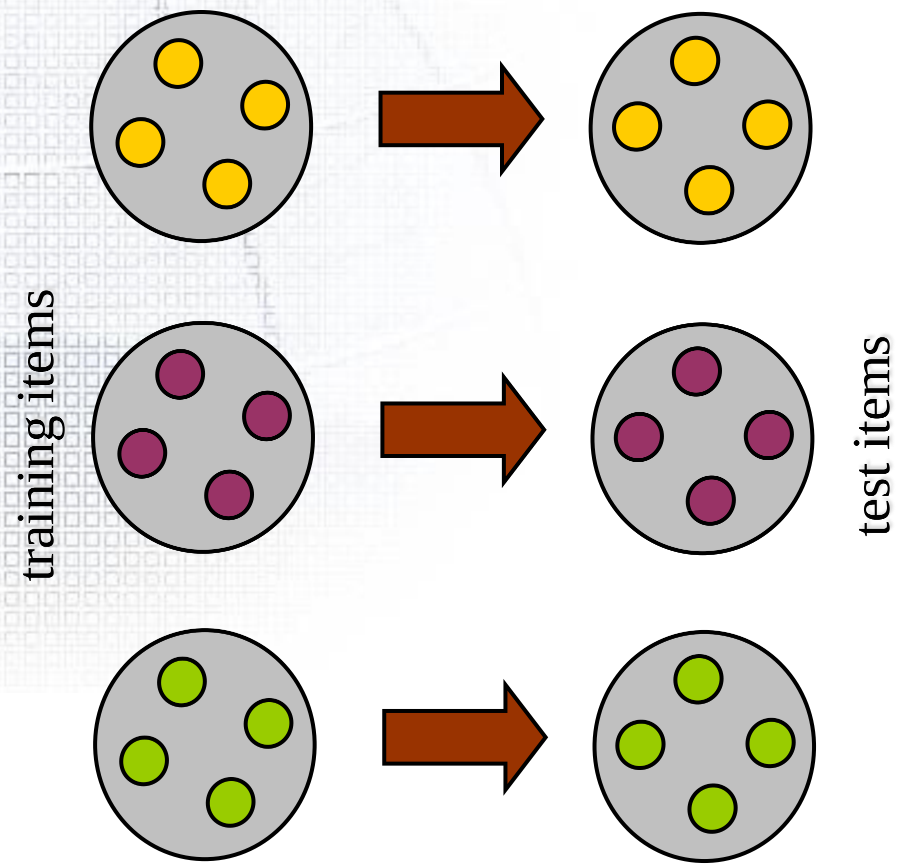


Given a source domain D_S and learning task T_S, a target domain D_T and learning task T_T, transfer learning aims to help improve the learning of the target predictive function $f_T(\cdot)$ in D_T using the knowledge in the D_S and D_T, where $D_S \neq D_T$, or $T_S \neq T_T$.

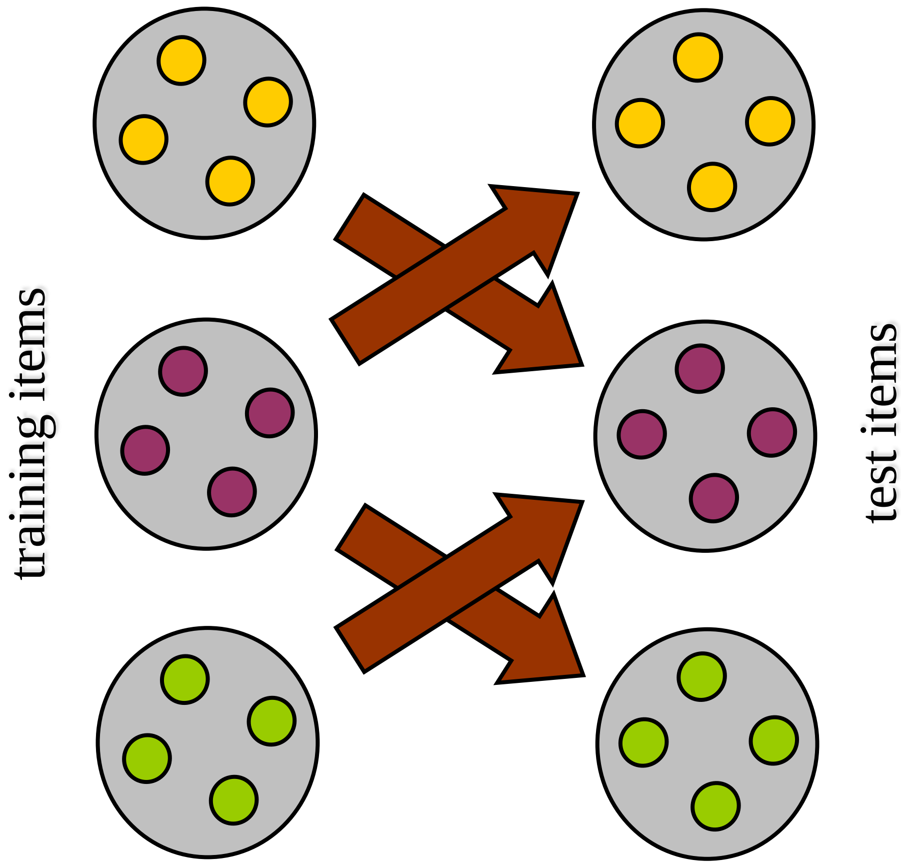
Traditional ML vs. TL

(P. Langley 06)

Traditional ML in
multiple domains



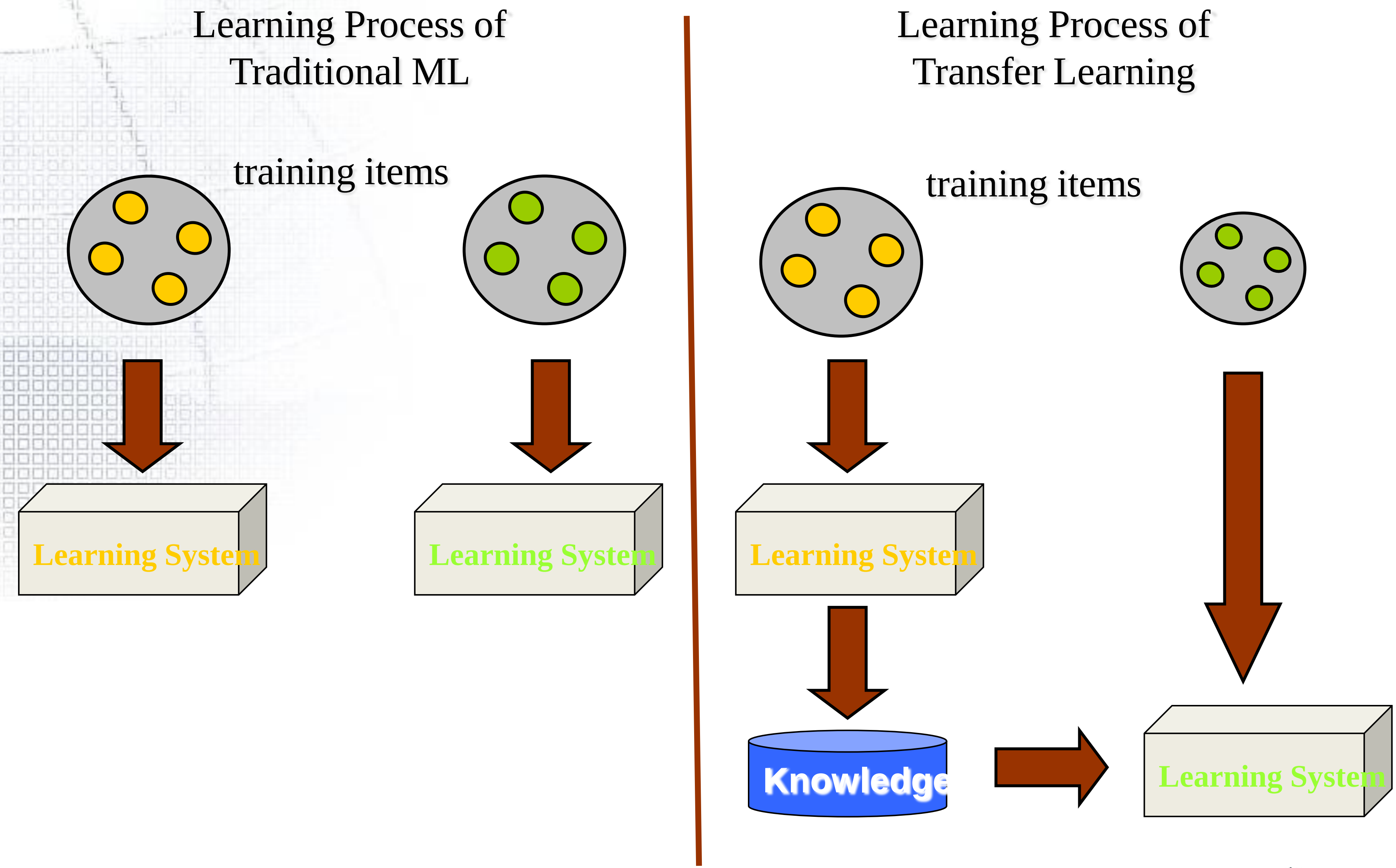
Transfer of learning
across domains



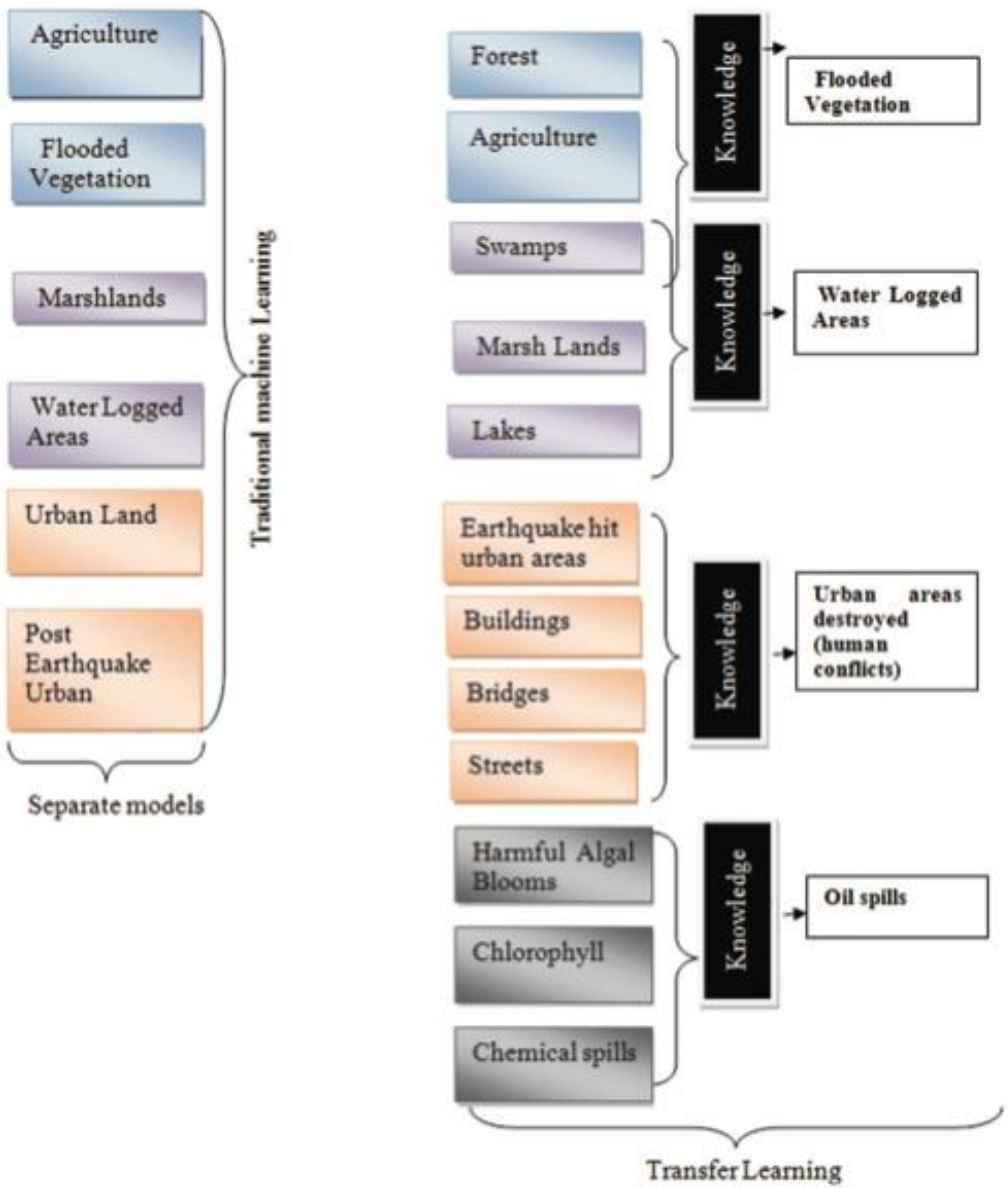
Humans can learn in many domains.

Humans can also transfer from one domain to other domains.

Traditional ML vs. TL



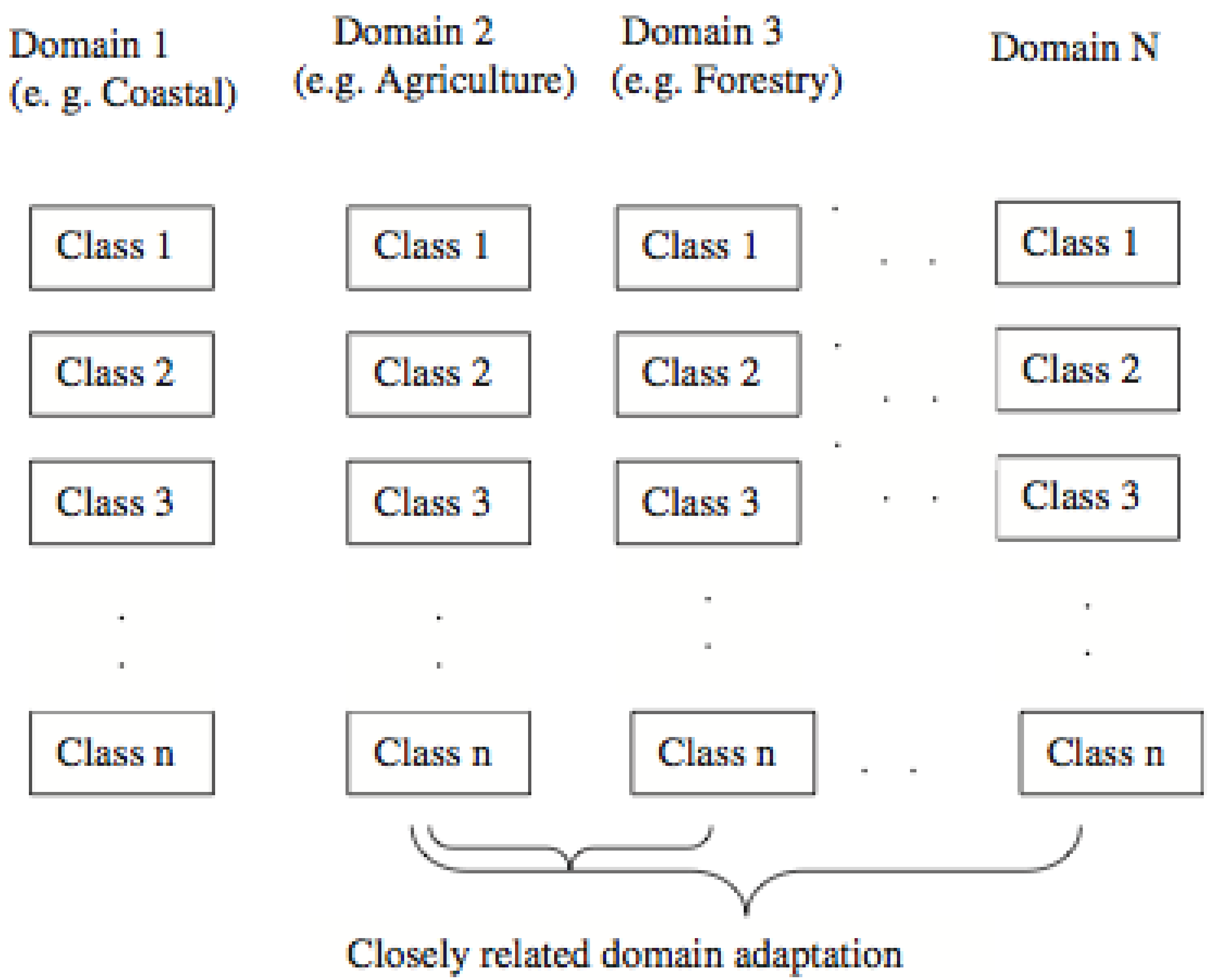
Transfer Learning in RS applications



Learning new models based on transferring prior knowledge of similar classes between closely related tasks or domains.

Water + agriculture +..+= could become a flooded agriculture class

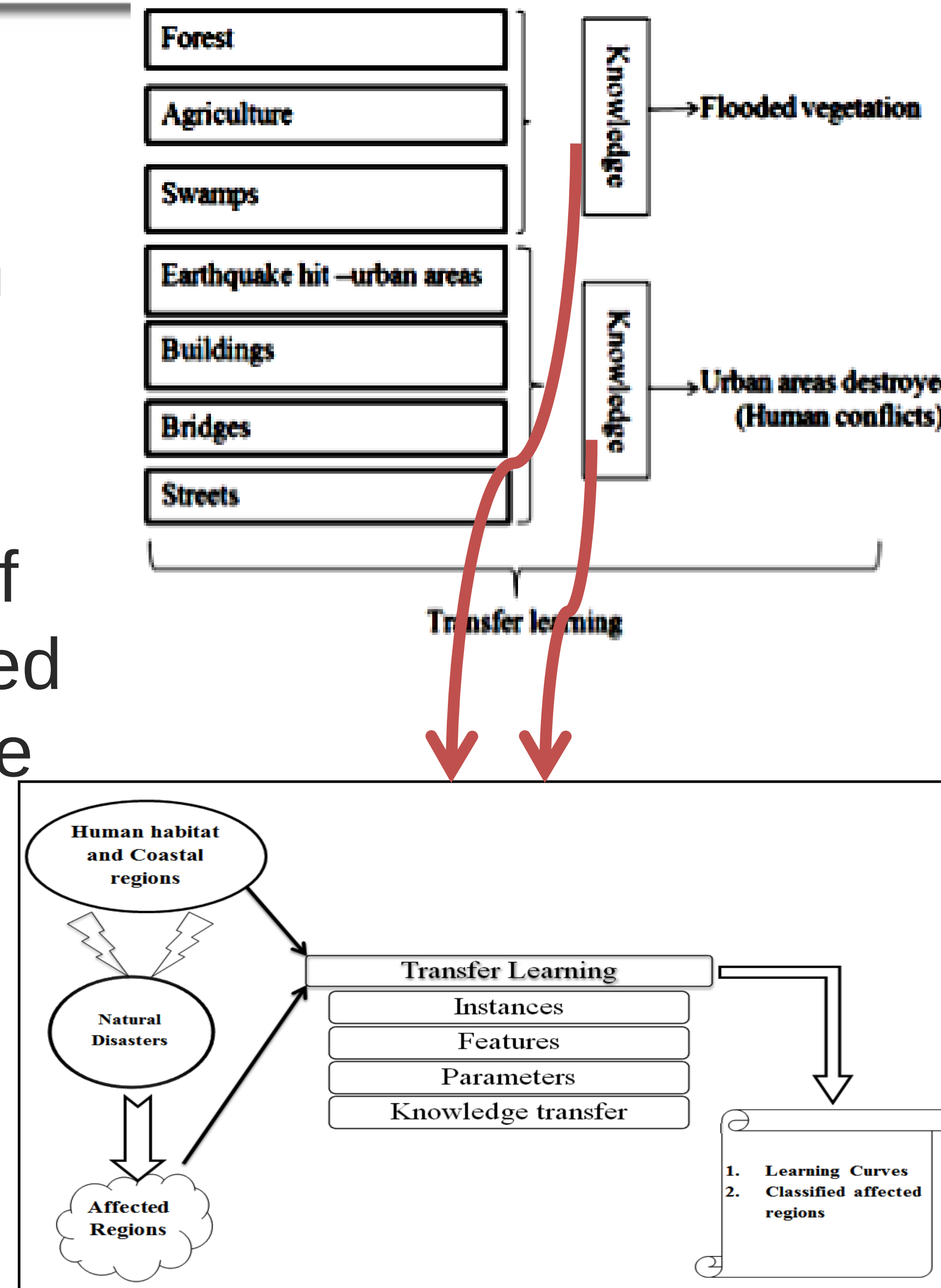
Closely related domain adaptation



Adapting and transferring knowledge from closely related source domains for enhanced classification using limited number of training samples from target domain.

What, When, How?

- **what to transfer'**
  - refers to the kind of entities that are transferred between tasks;
- **'when to transfer'**
  - depends upon the amount of prior knowledge that is shared between tasks and should be carefully selected to avoid negative transfer.
- **'how to transfer'** .
  - what part of the knowledge that is being transferred.



Instance, Features, Parameters based Transfer

- **Instance-based Transfer Learning**
  - Part of the labeled data in the source domain can be reused in the target domain after re-weighting
- **Feature-based Transfer Learning Approaches**
  - When source and target domains only have some overlapping features (lots of features only have support in either the source or the target domain)
- **Parameters-based Transfer Learning Approaches**
- A well-trained model θ_S^* has learned a lot of structure. If two tasks are related, this structure can be transferred to learn θ_T^*.

Domain Adaptation (DA) Approaches

- One of the main aspects in knowledge transfer approaches is to identify a *weighting parameter* or *similarity measure* among different domains and use this measure as a regularizer (or) parameter to facilitate labeling in the target domain.

- Because of the distributional differences among source and target domains using this *weighing parameter / similarity metric* enforces a better performance on the target classifier.

Instance based Knowledge Transfer

Given a training dataset$\{x_k , y_k\}_{k=1}^{n}$where input data $x_k \in R^n$and its corresponding output $y_k \in$ R and class labels $y_k \in \{1, -1\}$

The function to learn is

$$H(x) = \boldsymbol{w}.\phi(x) + \boldsymbol{b} \quad \textbf{(1)}$$

$\phi(x)$is used to map the input samples to a high dimensional feature space through a kernel function. To find the model parameters $\boldsymbol{w}, b$

$$min_{w,b} \frac{1}{2}||w||^2 + \frac{C}{2}\sum_{k=1}^{n}[y_i - w.\phi(x_i) - b]^2 \quad \textbf{(2)}$$

Where, 'C' is a control parameter for the size of the margin.

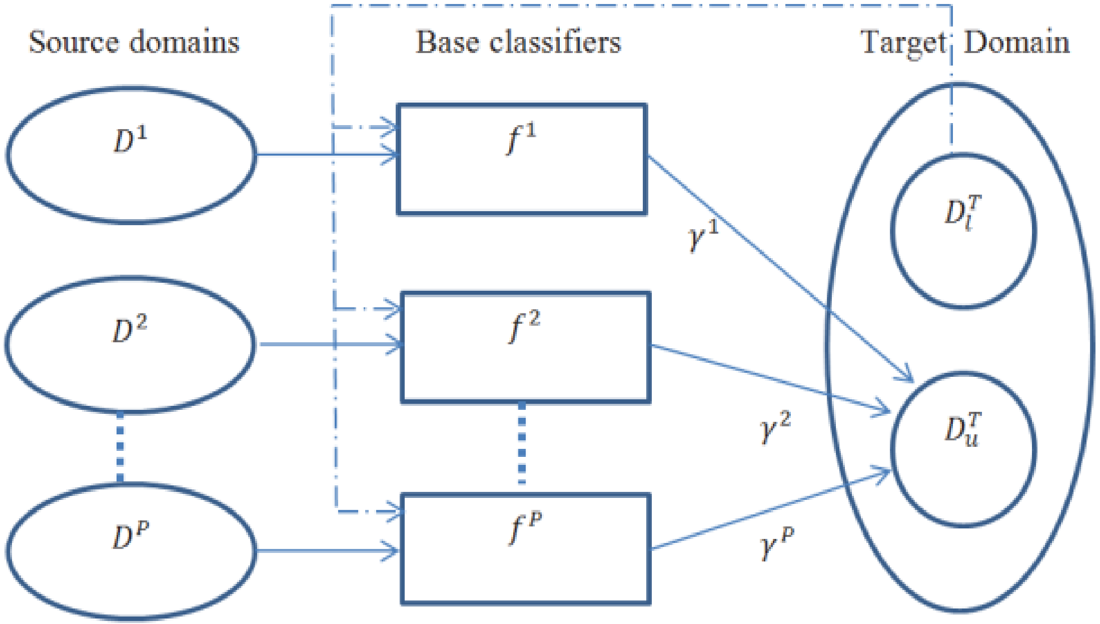
The adaptation of this formulation to constrain a *new model (e.g. Flooded urban areas)*to be close to the pre-trained models is possible by [6]

$$min_{w,b} \frac{1}{2}||w - \beta w'||^2 + \frac{C}{2}\sum_{k=1}^{n}[y_i - w.\phi(x_i) - b]^2 \quad \textbf{(3)}$$

Where $w'$is related to the old model and β is a scaling factor that controls the degree to which a new model is close to the older one and produces the best prior knowledge model and that can be used for model adaptation

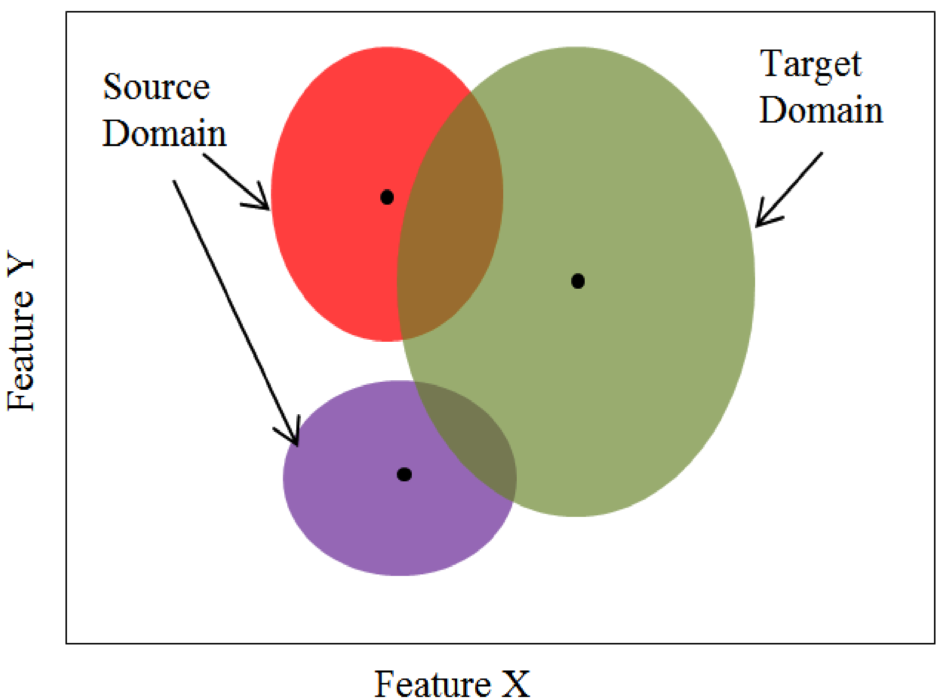
Domain Adaptation Method ...

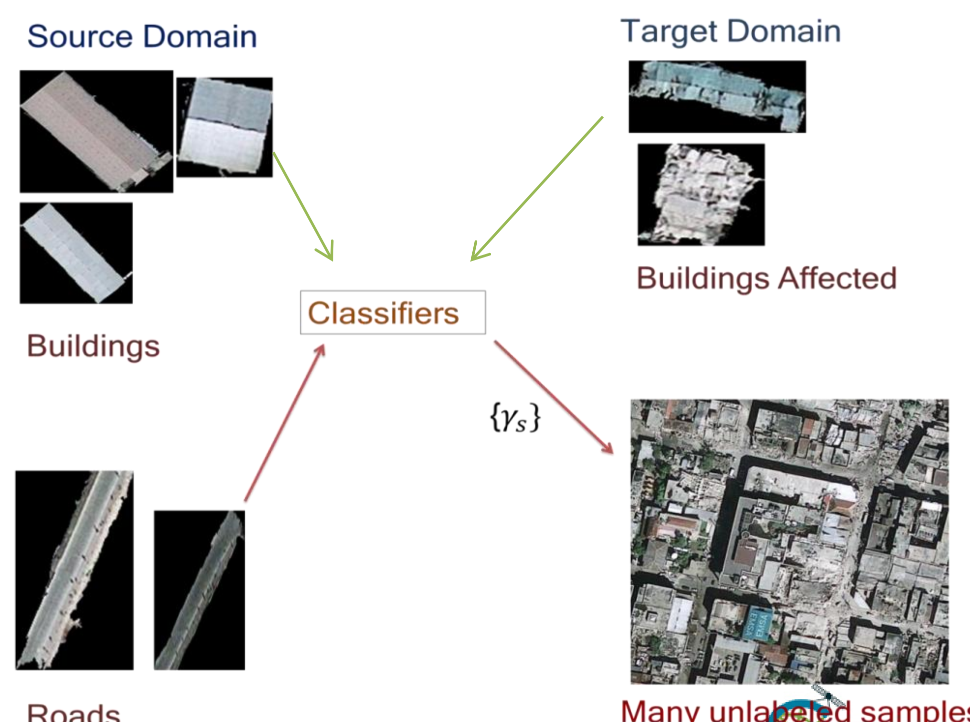
2. Find the relevance between domains



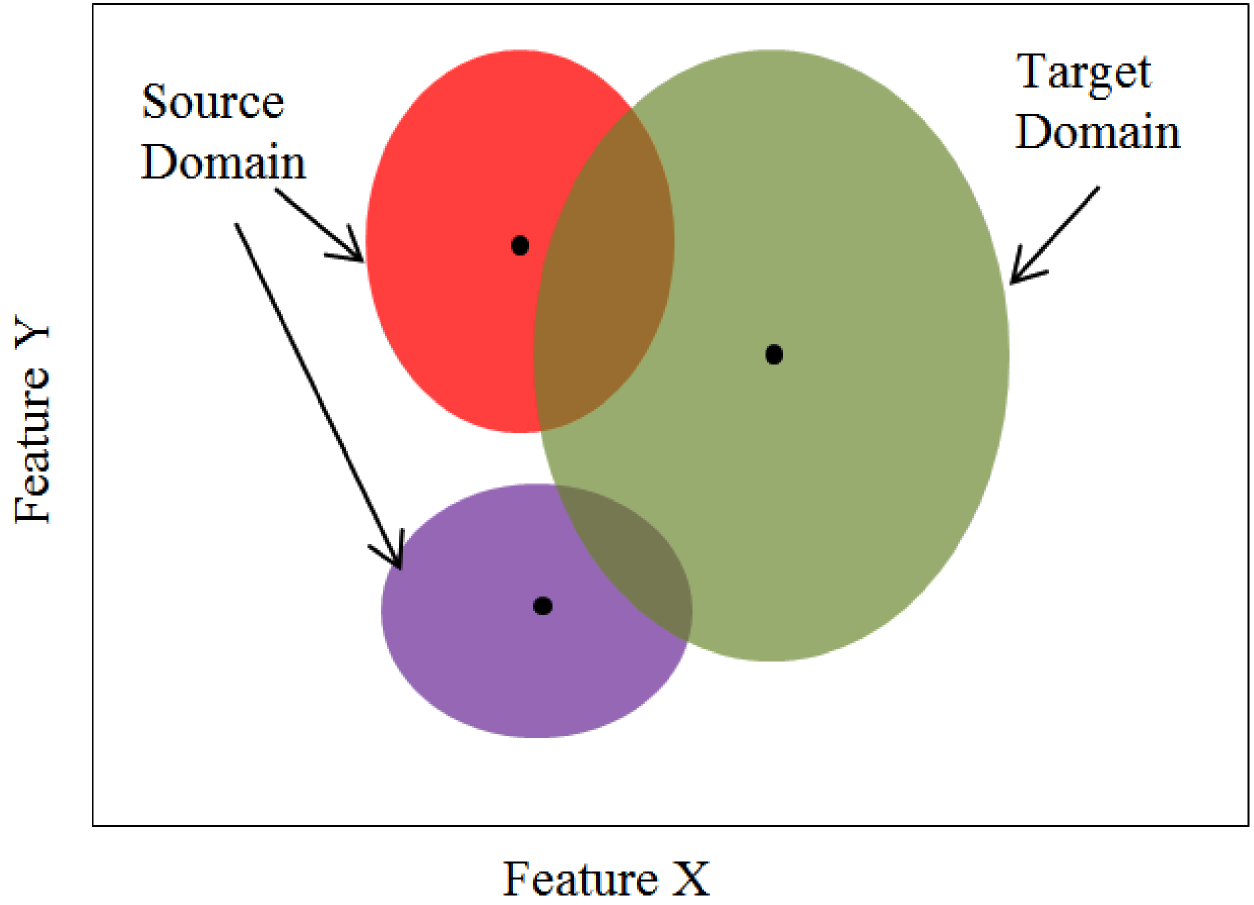
1. Build each class (Domain) classifier

3. Predict the final decision value





Domain adaptation approach, distribution difference



- ➢ Choosing a parameter(s), that identifies this relationship.
- ➢ This parameter(s) can enforce regularization during the learning process.

Urban Earthquakes

Source Domain

- *Buildings, Roads*

Target Domain

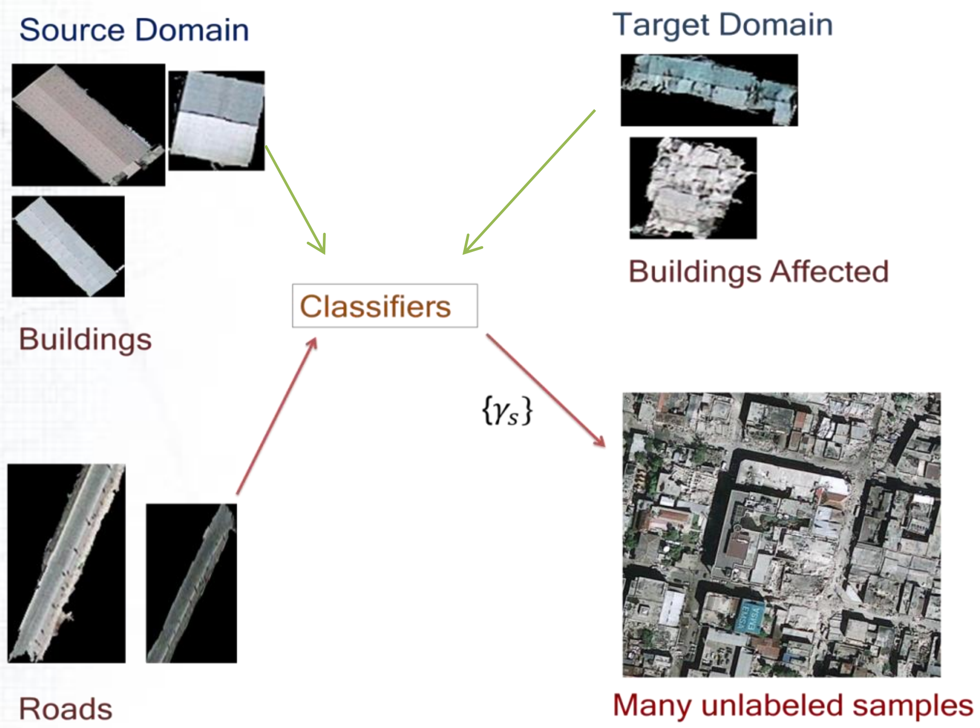
Affected Buildings

Source domains: Buildings, Roads.
Target Domain: Earthquake affected buildings.
Negative samples used: Trees, Shadows (Clutter).
Number of samples: Buildings (252),
Roads (81),
Earthquake regions(100)
Clutter (227)
Features used: Texture (GLCM and Wavelets)
Color

Urban Earthquakes: Dataset

- Haiti Earthquake on 12th January 2010
- Extent
  - (-72.363485ºW to -72.296249ºE) Longitude
  - (18.566070ºN to 18.501935ºS) Latitude
- Geo-Eye-1 (0.46 m –spatial resolution)

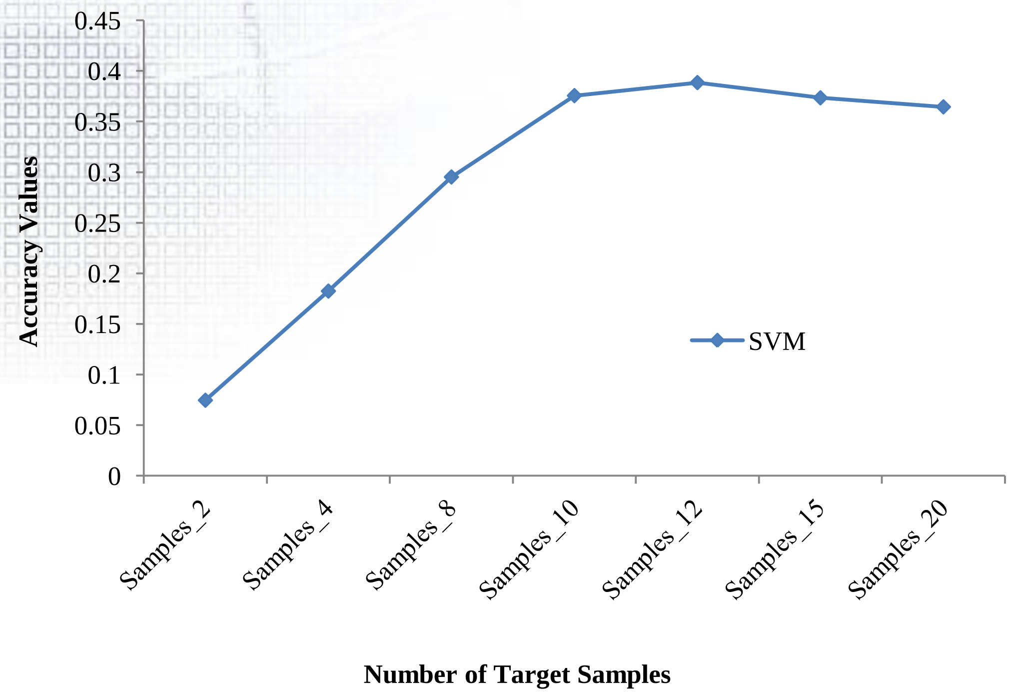
Urban Earthquakes: Results



- The accuracy results are evaluated on different combination of base classifiers and the semi-labels for the domain adaptation classifier.
- The different base classifiers are obtained when SVM is trained by
  - Only source domain samples
  - Only available target samples
  - Both source domains and the available target samples

Urban Earthquakes: Results

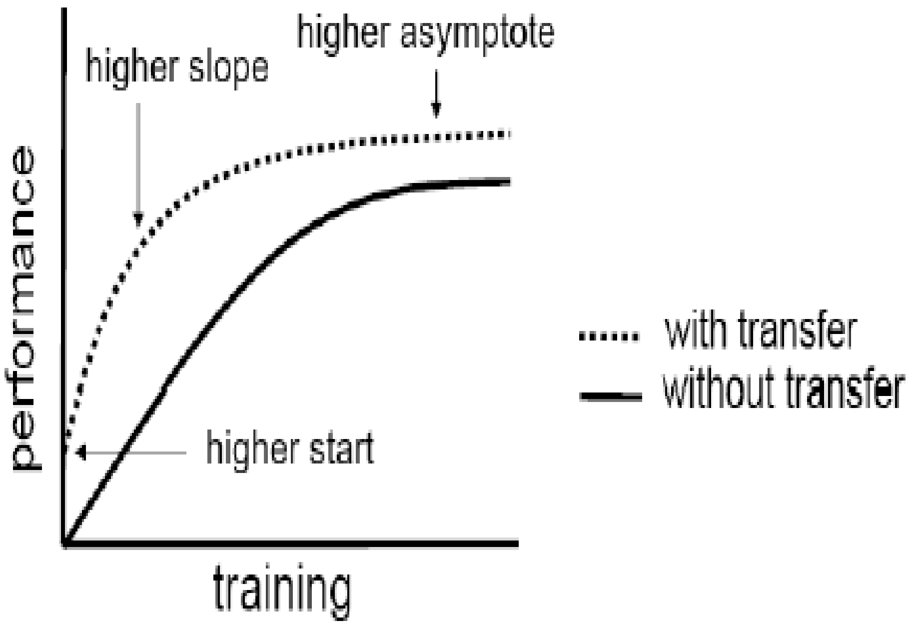
- ***Case I:*** *Estimation of unlabeled Target domain samples from the labeled target samples*
- The features used in this adaptation process are both Wavelets and GLCM.



Average accuracy values of SVM classifiers in estimating unlabeled target samples is 56 %

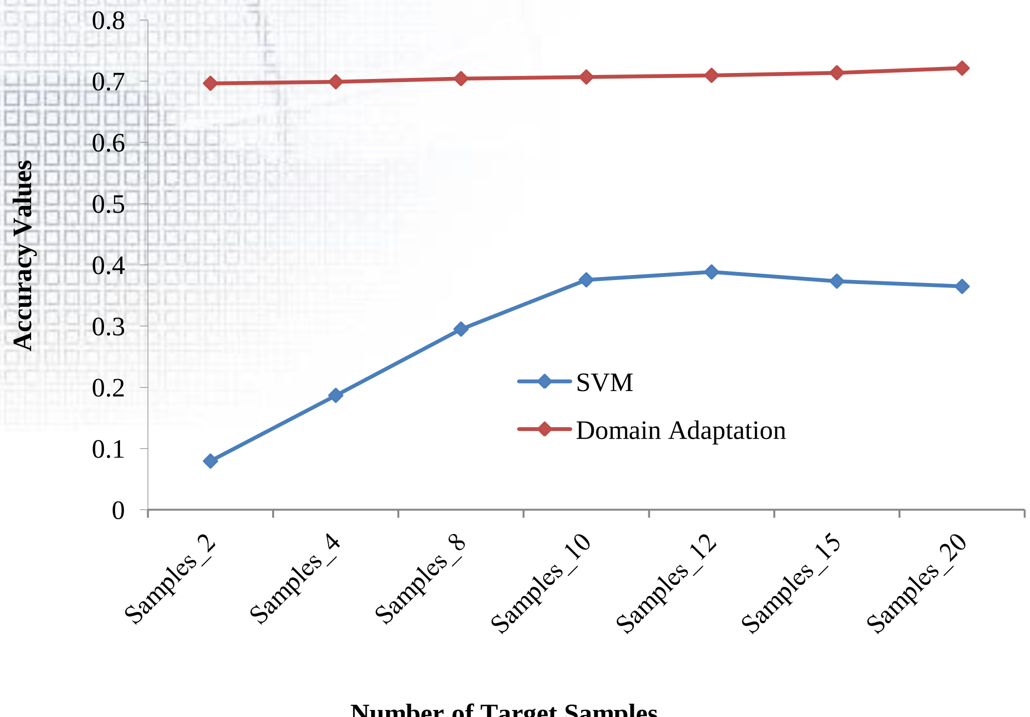
Three ways in which transfer might improve learning

- **Jumpstart:** initial performance achievable on the target task is much better
- **Higher slope:** shorter amount of time needed to fully learn the target task
- **Asymptotic Performance**: in the long run, the final performance level achievable over the target task may be higher compared to the final level without transfer



Urban Earthquakes: Results

- ***Case II:*** *Estimation of unlabeled Target domain samples from the labeled target samples vs DA approach, when base classifiers are trained by both source domain samples and the labeled target samples*
- Wavelet features are considered

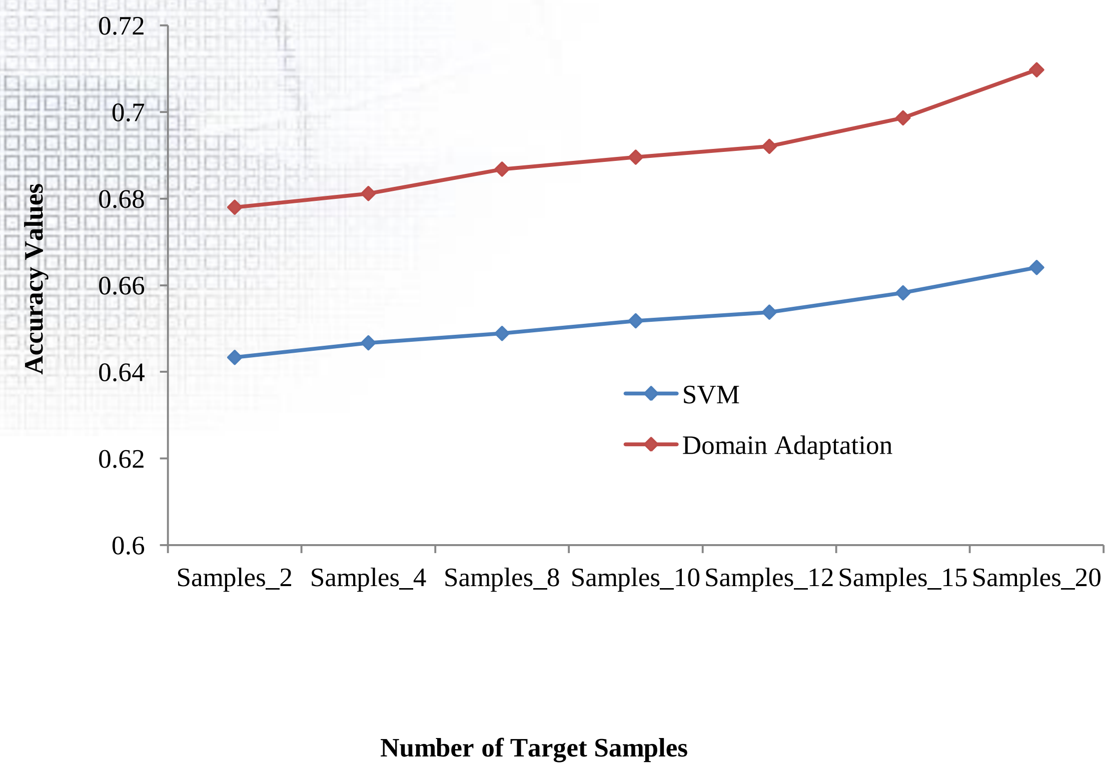


- SVM is learned using only the labeled target samples.
- DA is done by using base classifiers trained from both source domain and labeled target data

Accuracy values for increase in sample size for learning in DA approach (red line)

Urban Earthquakes: Results

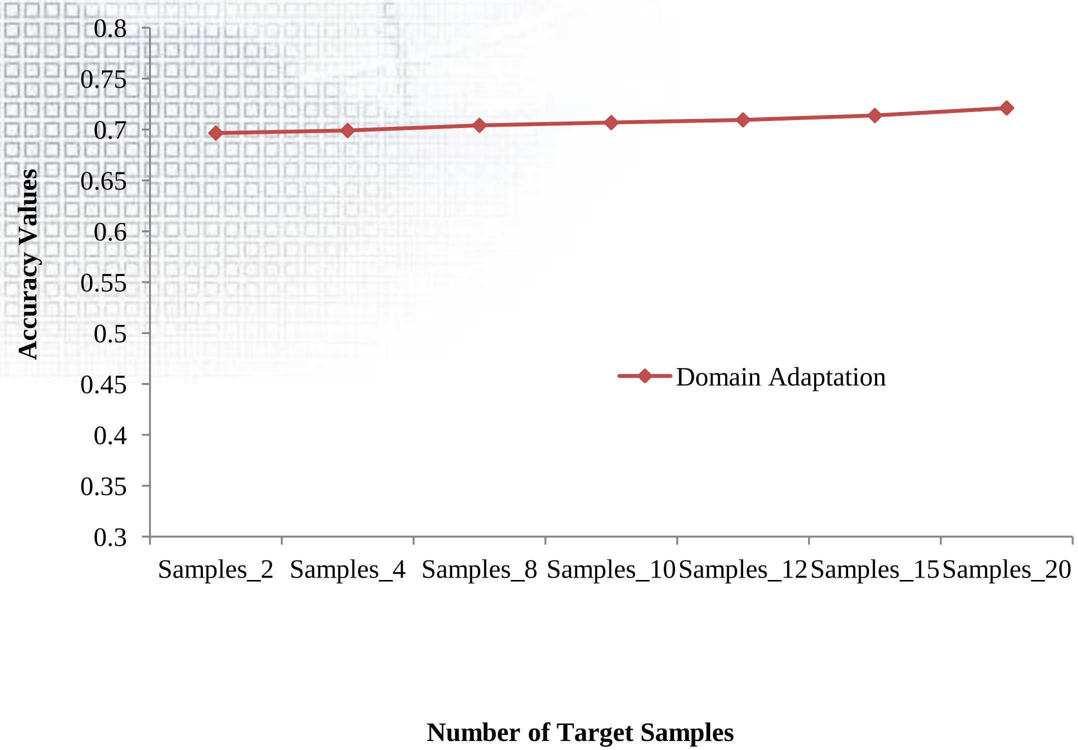
- ***Case III:*** *Estimation of unlabeled Target domain samples from source domain samples and the labeled target samples vs DA approach, when base classifiers are trained by source domain samples and the labeled target samples*
- GLCM features are considered



Accuracy values for increase in sample size for learning in DA approach. SVM learned by source domain and labeled target samples and DA by using base classifiers trained from both source domain and labeled target

Urban Earthquakes: Results

- ***Case IV:*** *Estimation of unlabeled Target domain samples from source domain samples vs DA approach, when base classifiers are trained by source domain samples and the labeled target samples*
- Features used here are ***color*** and wavelet features



Accuracies of DA approach when color and wavelet features are considered and the base classifiers are trained from both source domain and labeled target samples.

Urban Earthquakes: Results

International Journal of Image and Data Fusion
Vol. 3, No. 3, September 2012, 203–219

Taylor & Francis
Taylor & Francis Group

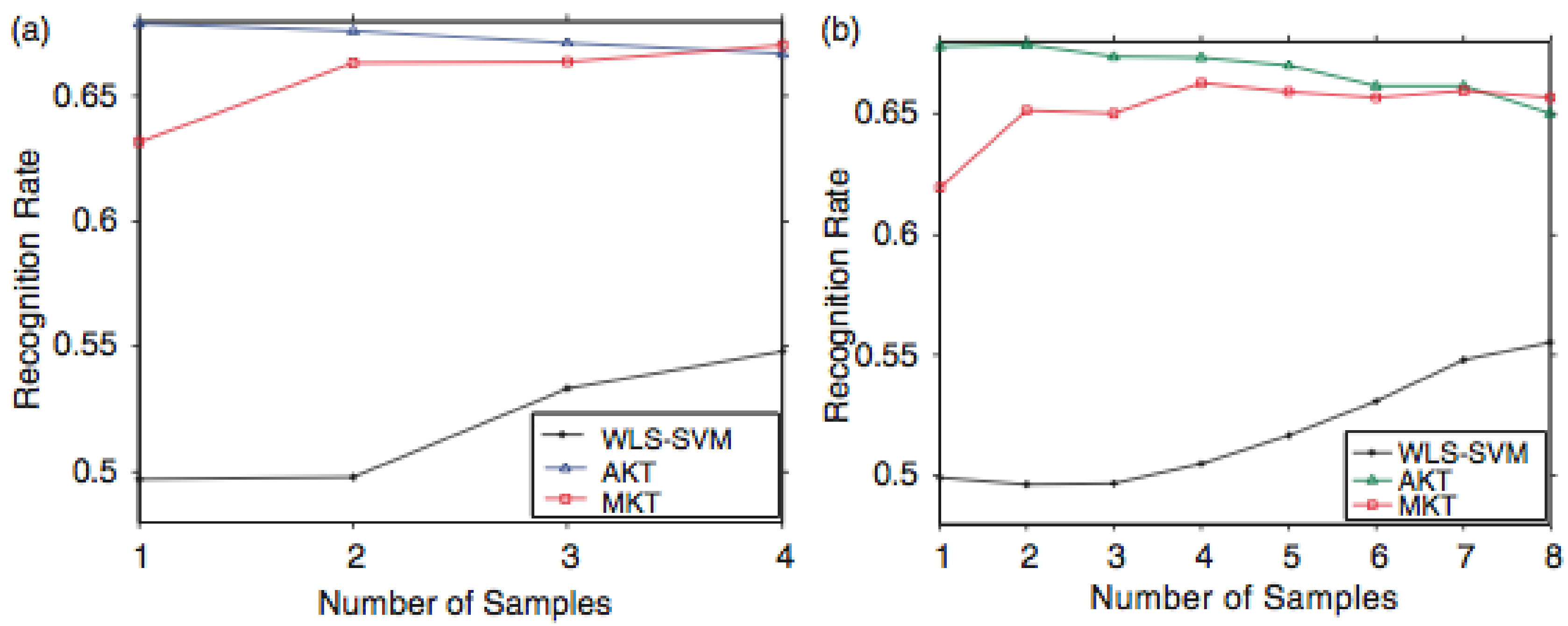
Floods

Transfer learning for image information mining applications

Surya Durbha[a], Roger King[b*], Prakash Andugula[a] and Nicolas H. Younan[b]

[a]Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai, Maharashtra, India; [b]Department of Electrical & Computer Engineering and Center for Advanced Vehicular Systems, Mississippi State University, PO 5405, Mississippi State, MS 39762, USA

(Received 15 November 2011; final version received 27 May 2012)



Recognition rate for six classes (agriculture, fallow, flooded agriculture and developed, flooded forest, forest and water) showing the number of samples used for the target class (4 and 8, respectively).

Floods

International Journal of Image and Data Fusion
Vol. 3, No. 3, September 2012, 203–219

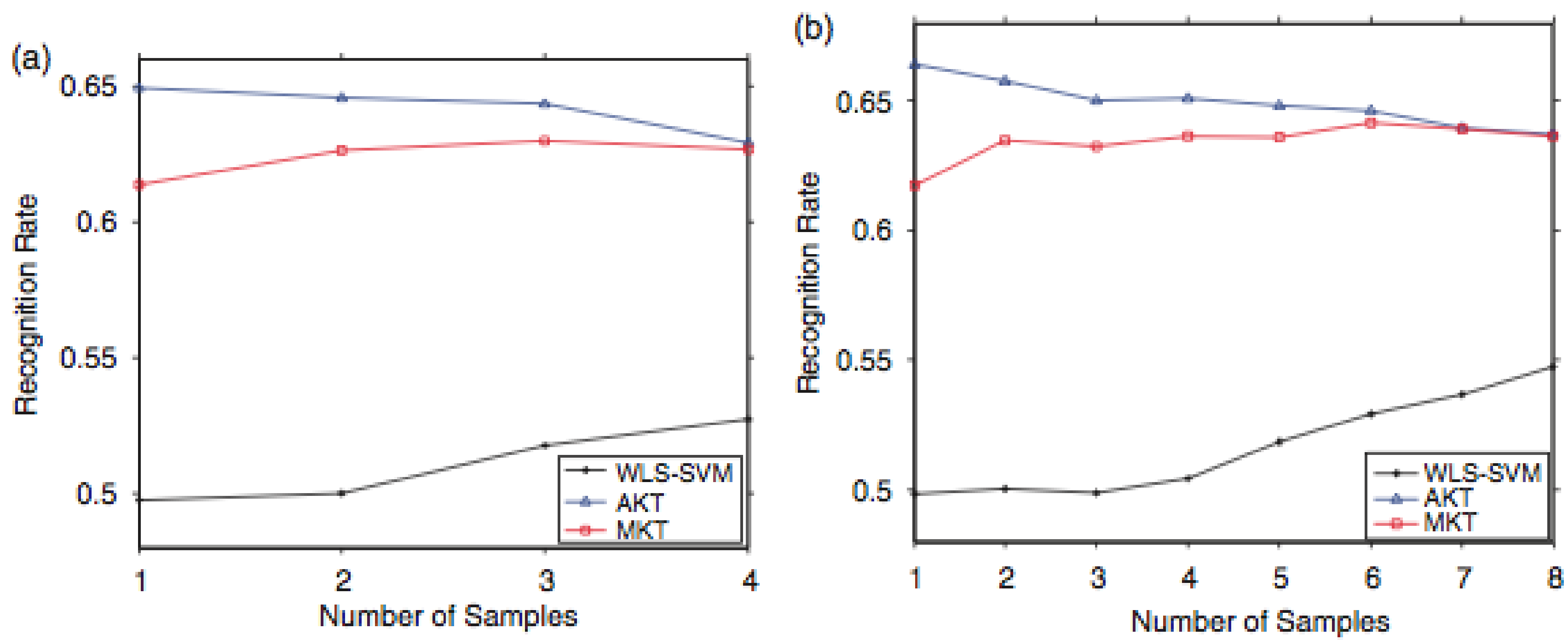
Taylor & Francis
Taylor & Francis Group

Transfer learning for image information mining applications

Surya Durbha[a], Roger King[b]*, Prakash Andugula[a] and Nicolas H. Younan[b]

[a]*Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai, Maharashtra, India;* [b]*Department of Electrical & Computer Engineering and Center for Advanced Vehicular Systems, Mississippi State University, PO 5405, Mississippi State, MS 39762, USA*

(Received 15 November 2011; final version received 27 May 2012)



Recognition rate for four classes (water, forest, agriculture, and flooded agriculture developed) showing the number of samples used for the target class (4 and 8, respectively).

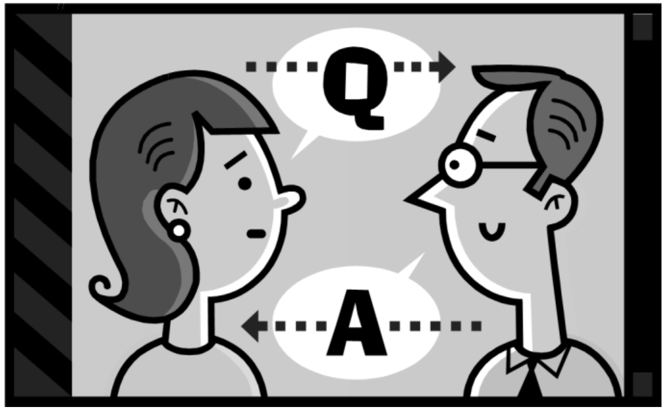
Conclusions

- Natural Disasters are a regular phenomenon in the India
  - several efforts are underway to use remote sensing technology for DSS
  - This work directly augments these activities through the development of a novel, state of the art methodology for rapid disaster assessment activities.
- The results of the work can be readily incorporated into operational disaster management activities that would give time sensitive information of the affected areas through classified maps.

Conclusions (contd.)

- The proposed approach can be used to update dynamic GIS databases more quickly than normally possible using traditional change detection techniques.
- The proposed approach can be scaled to various change detection problems also such as:
  - urban sprawl analysis, Harmful Algal blooms detection, chemical /oil spills, Translating thematic information from one classification system to another, etc.

Thank you !



Notation

Domain:

It consists of two components: A feature space $\mathcal{X}$, a marginal distribution $\mathcal{P}(X)$, where $X = \{x_1, x_2, ..., x_n\} \in \mathcal{X}$

In general, if two domains are different, then they may have different feature spaces or different marginal distributions.

Task:

Given a specific domain and label space $\mathcal{Y}$, for each x_i in the domain, to predict its corresponding label y_i, where $y_i \in \mathcal{Y}$

In general, if two tasks are different, then they may have different label spaces or different conditional distributions $\mathcal{P}(Y|X)$, where $Y = \{y_1, ..., y_n\}$ and $y_i \in \mathcal{Y}$

Notation

For simplicity, we only consider at most two domains and two tasks.

Source domain:

$$\mathcal{P}(X_S), \text{ where } X_S = \{x_{S_1}, x_{S_2}, \ldots, x_{S_{n_S}}\} \in \mathcal{X}_S$$

Task in the source domain:

$$\mathcal{P}(Y_S|X_S), \text{ where } Y_S = \{y_{S_1}, y_{S_2}, \ldots, y_{S_{n_S}}\} \text{ and } y_{S_i} \in \mathcal{Y}_S$$

Target domain:

$$\mathcal{P}(X_T), \text{ where } X_T = \{x_{T_1}, x_{T_2}, \ldots, x_{T_{n_T}}\} \in \mathcal{X}_T$$

Task in the target domain

$$\mathcal{P}(Y_T|X_T), \text{ where } Y_T = \{y_{T_1}, y_{T_2}, \ldots, y_{T_{n_T}}\} \text{ and } y_{T_i} \in \mathcal{Y}_T$$

Key References

1. http://unosat-maps.web.cern.ch/unosat-maps/HT/EQ20100114HTI/PDNA_HTI_EQ2010_Building DamagePosterA0_v1_HR.pdf

2. U.S. Geological Survey, http://earthexplorer.usgs.gov/ accessed on December 24th 2011, Sensors Advanced Land Imager (ALI), http:// eo1.usgs.gov/sensors/ali, Data source for Haiti Earth Quake, ftp://selene.tacc.utexas.edu/Magic/Hazard/, accessed as on October, 2012

3. Durbha, S., King, R.,Prakash, A., and Younan, N,H., Transfer learning for image information mining applications, International Journal of Image and Data Fusion, pp.203-219,2012.

4. Sinno Jialin Pan and Qiang Yang., A Survey on Transfer Learning. IEEE Transactions on knowledge and data engineering, pp.1345- 1359, 2010.

5. Prakash, A., Durbha,S,S., King, R,L., Younan, N,H., Multi-Task Learning For Disaster Image Mining, Proc. of ESA-EUSC-JRC 8th Conference on Image Information Mining, DOI: 10.2788/49465, Oct. 2012, pp. 78-81

6. Lixin, D., Dong, Xu., and Tsang, I,W., Domain Adaptation From Multiple Sources: A Domain-Dependent Regularization Approach, Neural Networks and Learning Systems, IEEE Transactions, pp.504-518, 2012.

7. Dell'Acqua, Fabio, and Paolo Gamba. "Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives." Proceedings of the IEEE 100.10 (2012): 2876-2890.