UnTran: Recognizing Unseen Activities with Unlabeled data using Transfer Learning

ACM/IEEE IoTDI'18

April 18th, 2018

Md Abdullah Al Hafiz Khan, Nirmalya Roy

UMBC
AN HONORS UNIVERSITY IN MARYLAND

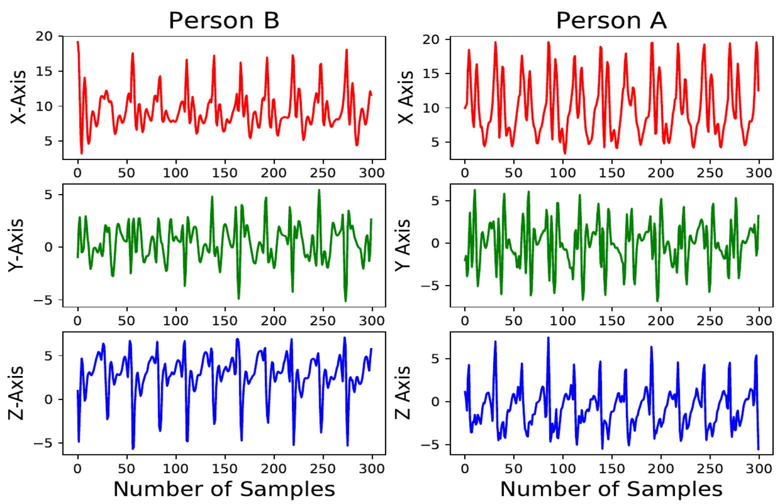
MPSC
Mobile Pervasive & Sensor Computing Lab

Challenges in Scaling Activity Recognition

- Cross User Diversity
- Device Type Diversity
- Device-instance Diversity
- Heterogeneous Environments
- Heterogeneous sensor Diversity
- Unseen Activities

Motivation

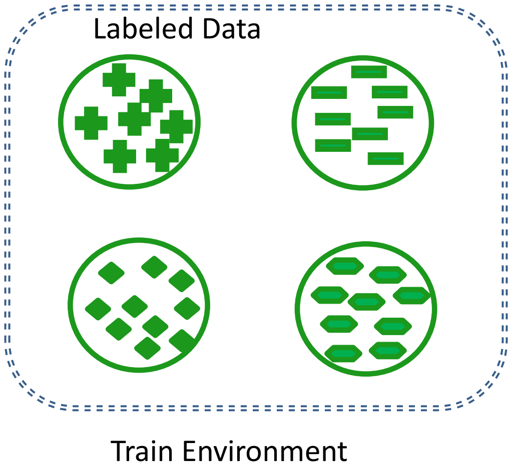
- Cross-user Diversity
  - Person A's walking pattern is different than Person B
  - One person's walking may be similar to running for another person.
  - How to cope with this diversity?

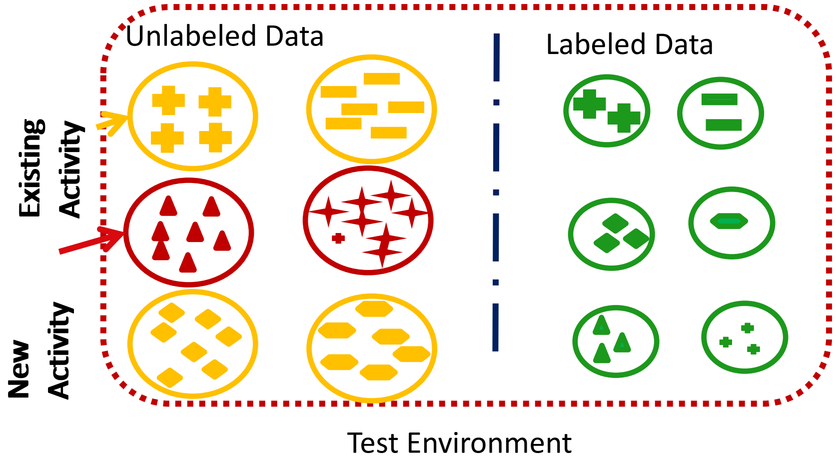


Walking

Motivation

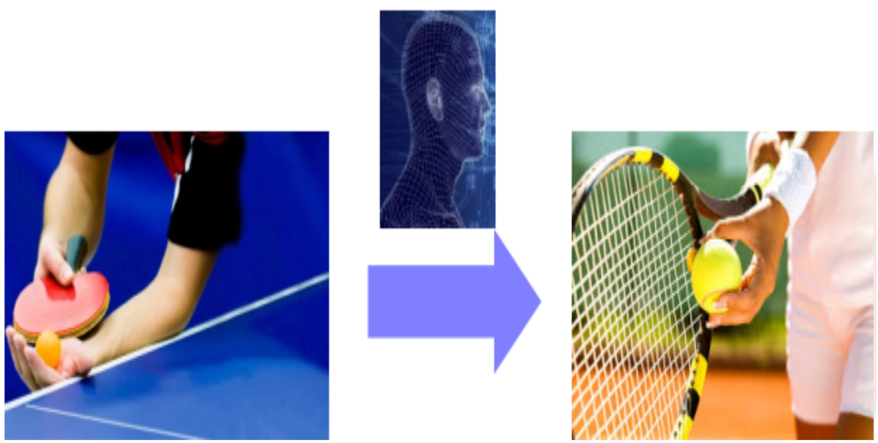
- Unseen Activities In the Target Domain
- Two Scenarios
  - Balanced Unseen Activities
  - Imbalanced Unseen Activities
- Balanced Unseen Activities
  - Both domain contains equivalent number of activities
  - New actvities
- Imbalanced Unseen Activities
  - Number of activities are larger than the training environment
  - New activities





Transfer Learning

- Psychological point of view
    - The study of dependency of human conduct, learning or performance on prior experience
        - Thorndike and Woodworth explored how individuals would transfer in one context to another context that share similar characteristics [Psychological review, 1901].



Transfer Learning

Machine learning community

- Inspired by human's transfer of learning ability
- The ability of a system to recognize and apply knowledge and skills learned in previous domains/tasks to novel tasks/domains, which share some commonality [pan et. al., TKDE 2010].

- Examples
    - Goal: to train an AR model to infer task T1 in an indoor environment E1 using machine learning techniques:
    - Sufficient training data required: sensor readings to measure the environment as well as human supervision, i.e., labels
    - A predictive model can be learned, and used in the same environment

Mobile Pervasive & Sensor Computing Lab

Our Approach

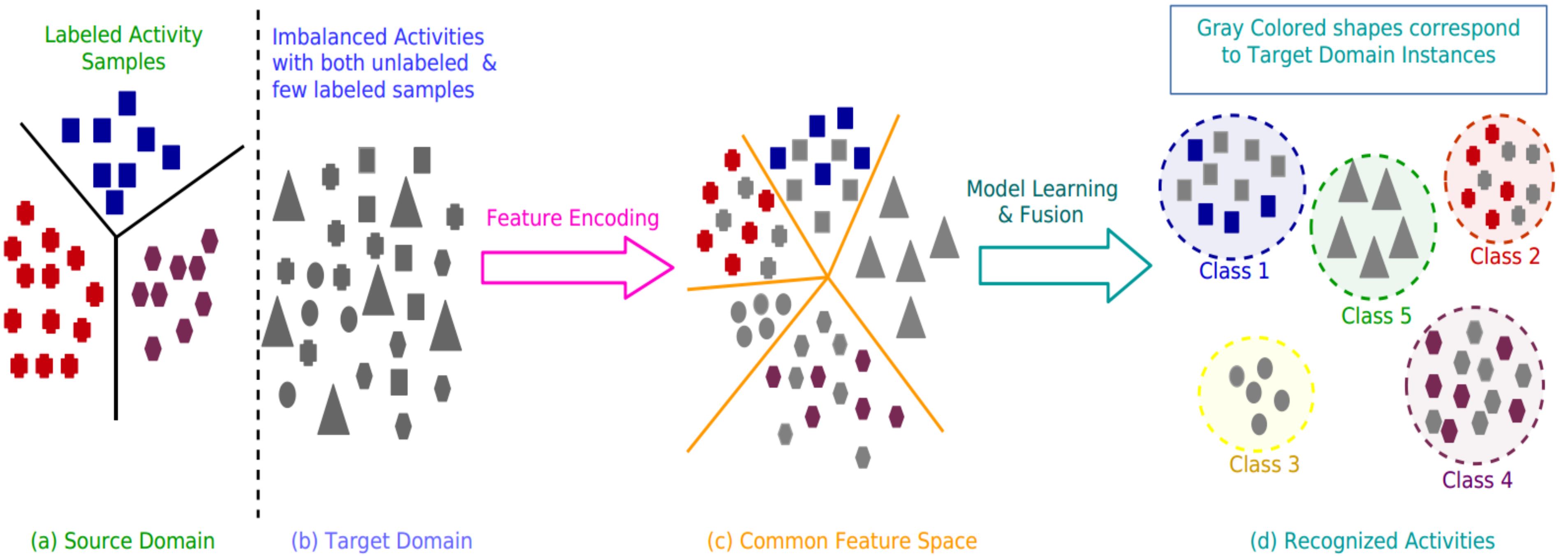
- We address the following scenarios
    - Imbalanced Activities
    - Unseen activities
    - May also contains the below challenges (inherently)
        - Cross User Diversity
        - Device-Instance Diversity
- Autoencoder
- Classifiers decision fusion

Mobile Pervasive & Sensor Computing Lab

Problem & Solution

- Collecting annotated samples are costly
- Deep models
  - Data hungry
  - Required large training time
- How to use Source trained Deep models?
  - Transfer one or more layer
  - no/small number of labels (target domain)
  - Reduce training time, reuse existing model
- Unseen (both balanced and imbalanced)
- Autoencoder + shallow classifier

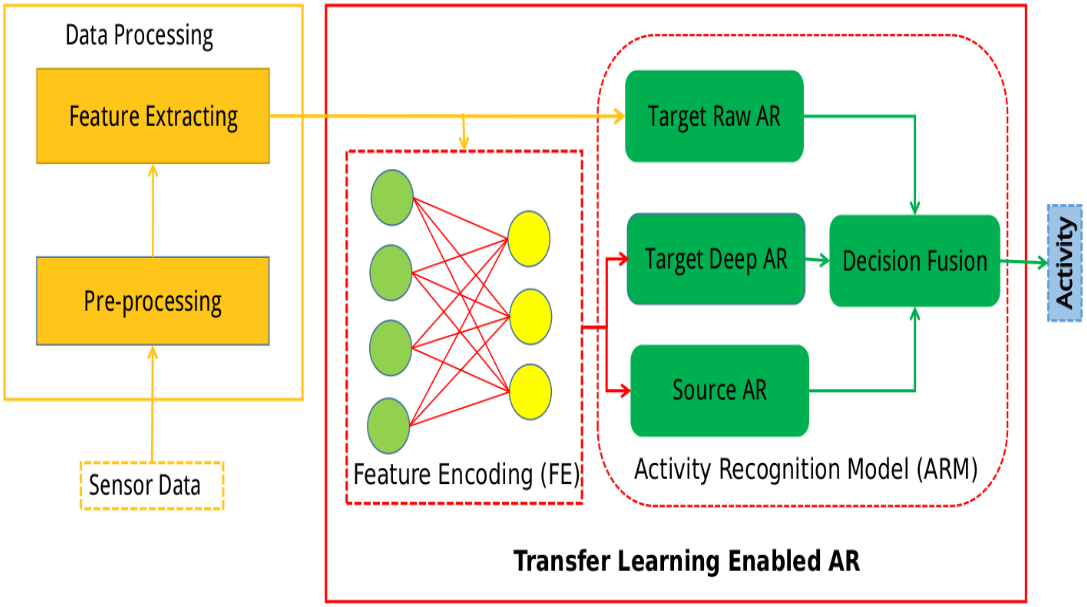
Our AR Approach



Overview of our activity recognition approach. (a) Source domain labeled activity instances, (b) Target domain contains both unlabeled and few labeled activity instances, (c) Common feature space for classification, and (d) Resulting activities after classification. Note that different shapes correspond to different activities

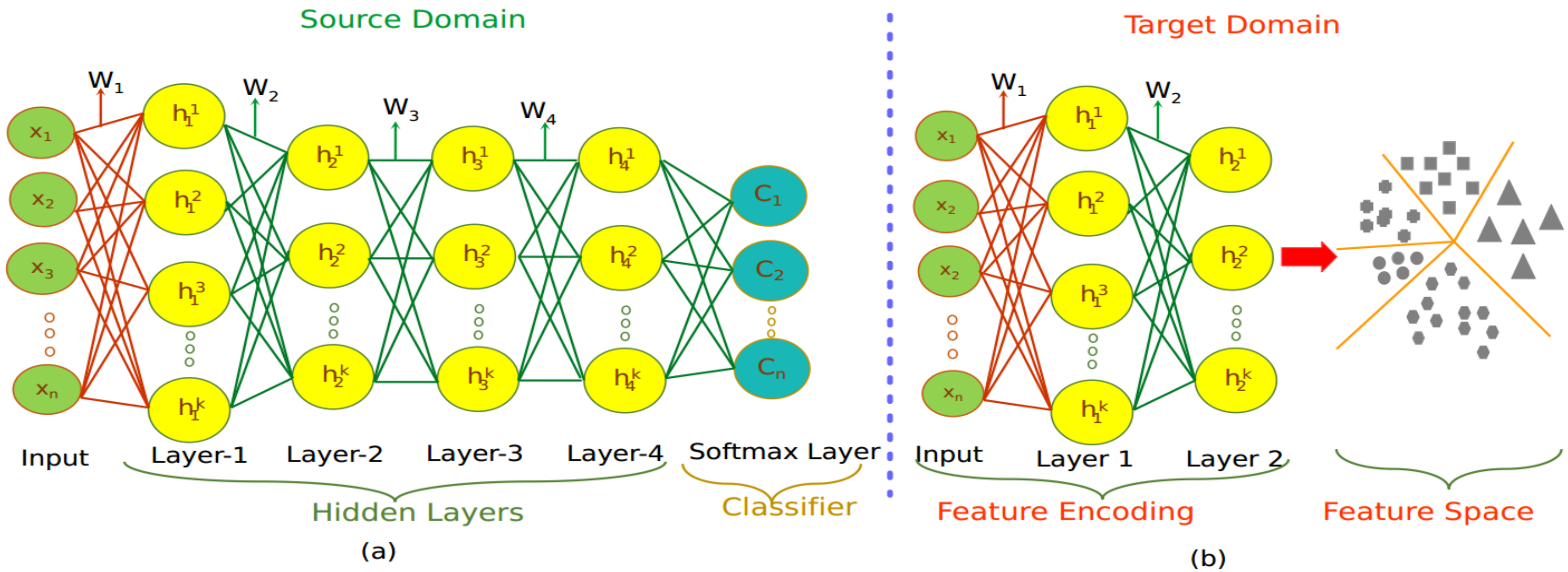
Proposed Architecture

- Three Modules
  - Data Processing
  - Feature Encoding
  - Activity Recognition
- Data Processing
  - Pre-processing
  - Feature Extracting
- Feature Encoding
  - Autoendoer
- Activity Recognition Model
  - Fusion - Traget Raw AR, Source AR, Target Deep AR



Overall Architecture

Feature Encoding



- Four layers autoencoder named it Deep Sparse Autoencoder (DSAE)
- Cost function: $J_{aen}(W,b) = \min_{\mathbf{W},\mathbf{b}} ||\mathbf{x} - \bar{\mathbf{x}}||^2 + \alpha \sum_{i=1}^{N_{L_1}} \Phi_{kl}(\rho||\hat{\rho}_i) \qquad \hat{\rho}_j = \frac{1}{m} \sum_{i=1}^{m} [a_j^{L_1} x_i]$
- Additional classifier layer (softmax layer)
- Lower layer features are more generic [6]
- Transfer two layers to implicit minimize domain distribution

Activity Recognition Model

- Fuse three classifiers
  - Based on empirical evaluation
- Source trained classifier
  - Deep feature based
- Traget classifier
  - Deep Feature based
  - Raw Feature based
- Class probability: $P(y_i|x) = \frac{1}{1+\exp(Af(x)+B)}$
- Novel class detector
  - One class SVM
  - Distinguish between seen vs unseen
- Activity Class determination

$$y^* = \arg\max_y \phi(y|x)$$

Existing activity fusion probability

$$\phi(y|x) = \begin{cases} P_s(y|x) + P_d(y|x), & \text{if } (y_d = y_r) \\ max(P_s(y|x), P_d(y|x)), & \text{else if } (y_s = y_d) \\ P_s(y|x) \times P_r(y|x) \times P_d(y|x), & \text{otherwise} \end{cases}$$

Unseen activity fusion probability

$$\phi(y|x) = \begin{cases} P_r(y|x) + P_d(y|x), & \text{if } (y_d = y_r) \\ max(P_d(y|x), P_r(y|x)), & \text{otherwise} \end{cases}$$

Mobile Pervasive & Sensor Computing Lab

Evaluation

- Three datasets -
  - Opportunistic (Opp),
  - Wisdm
  - Daily and sports (Das)

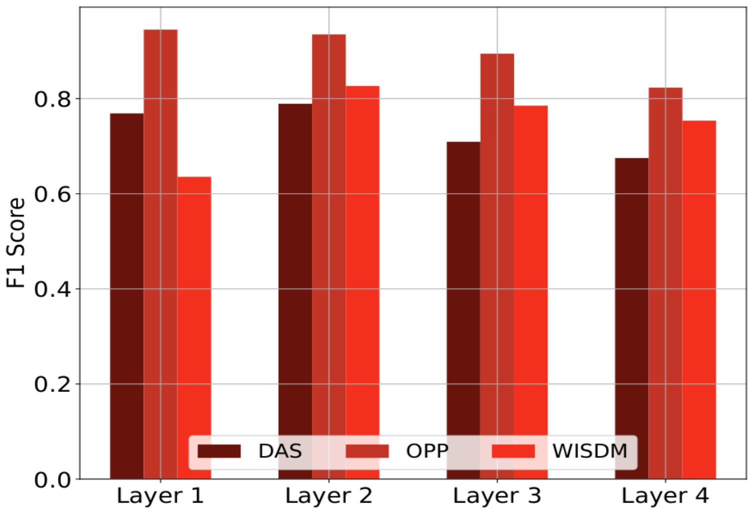
- Opportunistic
  - 17 activities (ADL), 4 participant, 64 Hz sampling frequency, accelerometer sensor

- Wisdm
  - 6 distinct activities, sampling frequency 20 Hz, 29 users, smartphone kept on pants pocket

- Daily and Sports (Das)
  - 19 activities, 8 users, sampling frequency 25 Hz, right arm data was considered

UnTran performance on different layers

- Fixed number of unseen activities in target domain
- Standard leave-two-sample-out cross-validation
- Generic features in lower layers and domain specific feature on upper layers
- 30% labeled data to train target domain classifier

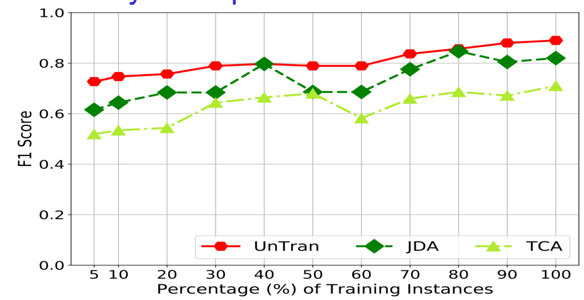


Performance on different layers

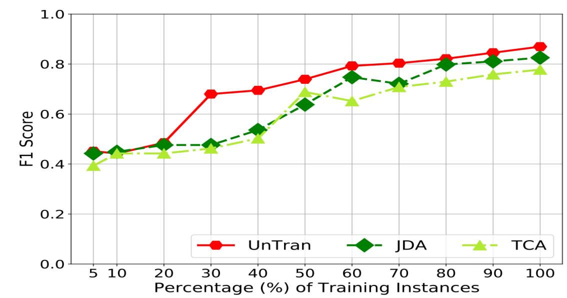
Balanced Activities: Varying Labeled Data

- Equivalent number of activities in both domain
- Standard leave-two-out cross validation
- Varying amount of labeled data of (n-2) samples randomly
- 20-30% labeled data required to get resonable performance
- Larger data distributions reduces the performance

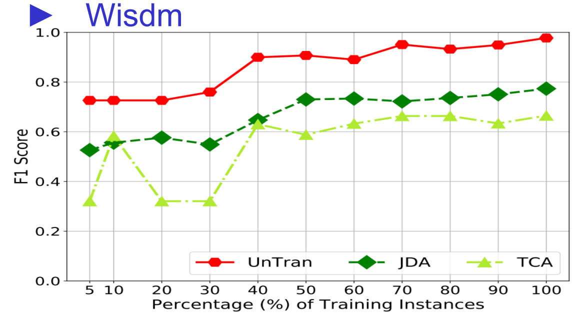
▶ Daily and sport



▶ Opportunistic



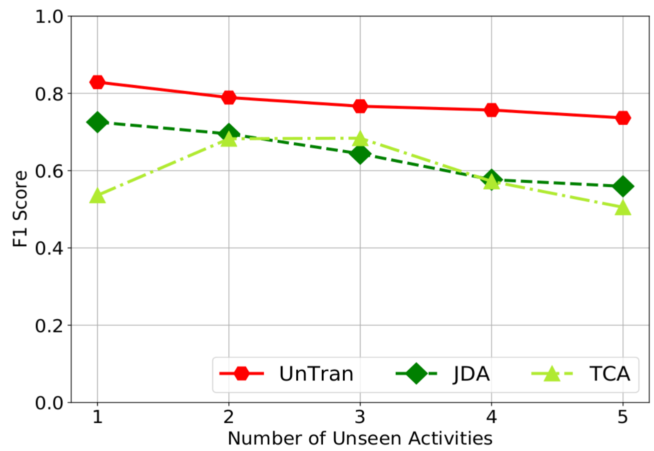
▶ Wisdm



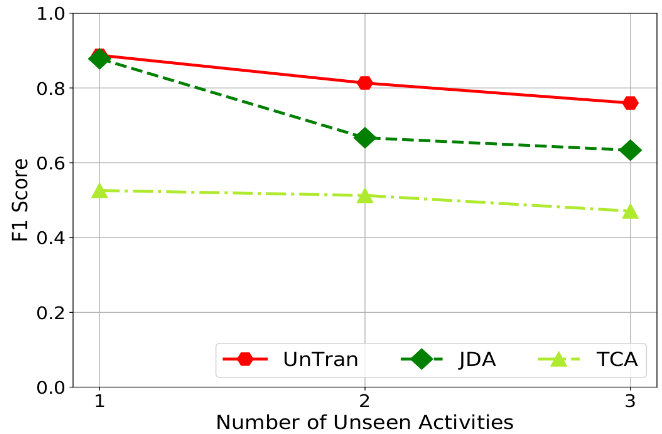
MPSC
Mobile Pervasive & Sensor Computing Lab

- Vary number of activities
- Similar leave-two-samples-out cross validation
- Model trained with 30% labeled data
- Performance drops 5-12% with increasing number of unseen activities
- Performance gain 10-13% compared to TCA and JDA

▶ Daily and sport

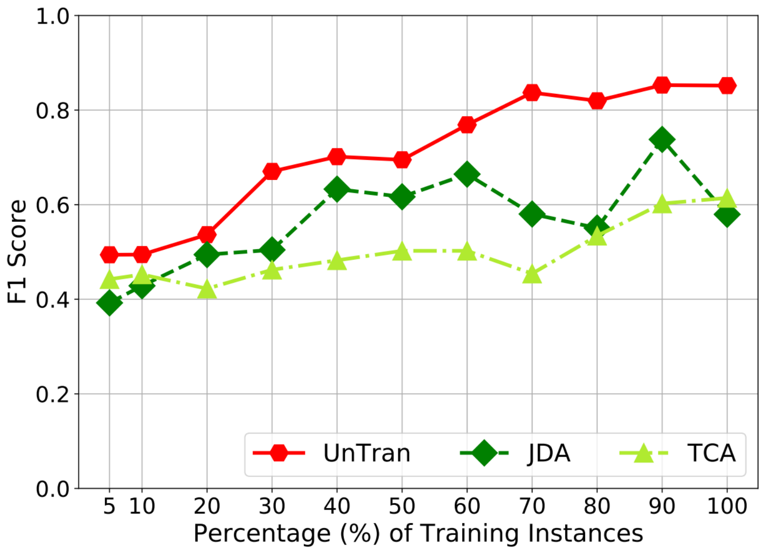


▶ Wisdm



Imbalance Activities: Varying Labeled Data

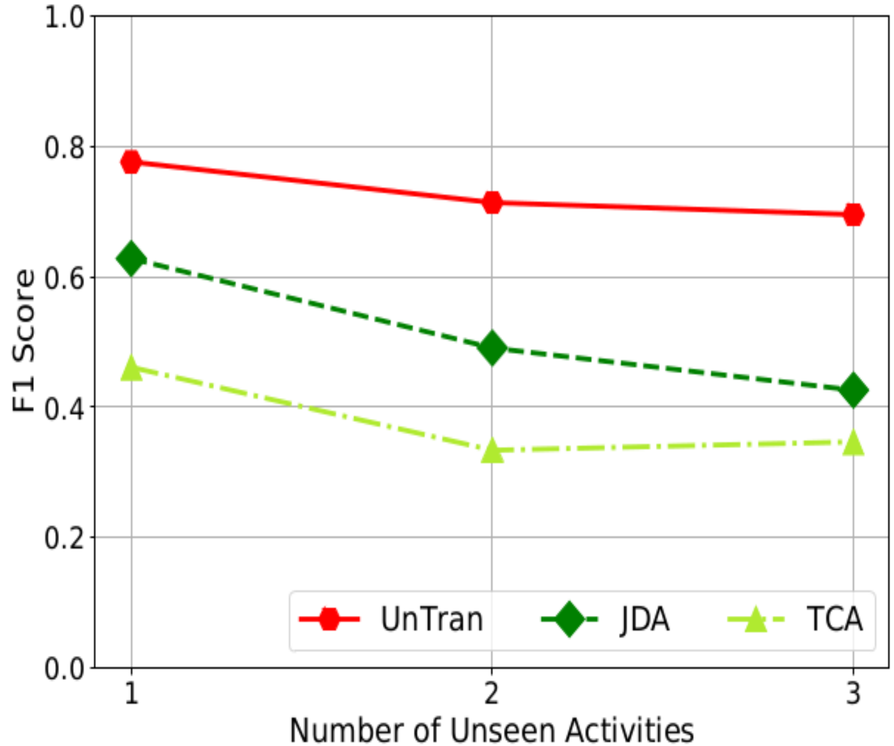
- Leave-two-class-out cross validation
- (A-2) activity classes participate in training
- Rest two class activity samples used in testing phase
- Trained with 30% annotated data
- Performance gain 10-12% compared to TCA, JDA



Opportunistic

Imbalance Activities: Varying Unseen Activities

- Leave-two-class-out cross validation
- (A-2) activity classes participate in training
- Rest two class activity samples used in testing phase
- Performance gain 15-20% compared to TCA, JDA
- Achieves F1 score about 70% on average



Opportunistic

Discussion & Conclusion

- Cross user diversity investigation are warranted
- Explicit structural pattern mapping among activities and instances are needed
- Intra- and inter-activity similarities can be exploited
- Annotation cost
  - Assumption is that the user provides few labeled data
  - One possible direction is to reduce the annotation cost
- UnTran achieves
  - Approx. 75% accuracy for coss-user differences with unlabeled data
  - Approx. 87% accuracy with 10% annotated samples in target domain

Thank you?

Md Abdullah Al Hafiz

PhD candidate
Information System
University of Maryland, Baltimore County
Email: mdkhan1@umbc.edu
https://userpages.umbc.edu/~mdkhan1

Mobile Pervasive & Sensor Computing Lab