
Supplementary Material of Spatio-Temporal Domain Adaptation for Gait Based User Recognition from Radar Data

Anonymous Author(s)

Affiliation

Address

email

1 Contents

2	1 Dataset	1
3	2 Model	2
4	3 Experiment Embeddings	2
5	4 Experiment Shift Metrics	6
6	5 Metric Learning Based Shift and Malignancy Detection Algorithm	7
7	6 GRL, GAN and Student-Teacher Methods for Unsupervised Domain Adaptation	8
8	1 Dataset	

9 Our dataset was collected using a TI IWR1642 mmWave sensor [1]. The radar chirp configuration
10 used in [2] was utilized to collect data from 4 different locations. The primary location, which was
11 used as the source domain, was a research lab with cubicles. Each of the ten subjects was asked to
12 walk back and forth in the room with a radar pointed at them. The subject's path was pre-determined
13 and was always perpendicular to the radar signal's propagation direction. Moreover, each day of a
14 user's data contains 100 data samples. 4-seconds of walking data was considered as an individual
15 data sample.

16 Apart from the source domain, three other locations were used to collect data. These locations were
17 used as the target domains. The locations consisted of a server, conference, and office room. Of
18 the three locations, the office location was reserved for the test set, and none of the experiments
19 utilized the office data for training. The radar device was deployed at all three target locations, and
20 the walking path setup is similar to the source domain. Furthermore, unlike the source domain, the
21 users generated only 50 data samples on each day at each target location. In addition, the data from
22 all four locations collected each day was not necessarily from the same day. The source location's
23 day one for a user could be today, while the data from the same user at each target location could be
24 collected on three different days. Similarly, all ten users were not required to generate data on the
25 same day for a given location. This allowed us to simulate a real-world setup, such as an office with
26 new employs generating data, upon being hired or given access to a restricted space.

27 **2 Model**

28 All our experiments used an 18-layer Resnet as defined in [3]. The last fully-connected layer was
29 replaced with a Multi-Layered Perceptron (MLP) with 128 hidden units and a final layer with hidden
30 units equal to the number of classes (10). The 128 dimensional fully connected layer's output was
31 used as an embedding layer. Moreover, a Constrictive Regularizer [4] was utilized as the activity
32 regularizer. Moreover, the last fully connected layer also used a Constrictive Regularizer for the
33 kernel weights instead of the activations. Additionally, given the definition of Additive Margin Loss
34 [5], we do not have any bias variables in the last layer. Finally, the logits were activated following the
35 definition in [5].

36 In all our supervised experiments, data from each available domain was shuffled together and sampled
37 in batches. This was done irrespective of the domain each sample belonged to. For our unsupervised
38 experiments, data from each available domain was sampled separately, which allowed us to generate
39 embeddings for data from each domain. We generate domain predictions with an MLP with 128
40 hidden units followed by another layer with hidden units equal to the number of domains being used
41 for training. The domain classifier MLP was either prepended by a gradient reversal layer (GRL) [6]
42 or treated as a discriminator in a Generative Adversarial Network (GAN) [7].

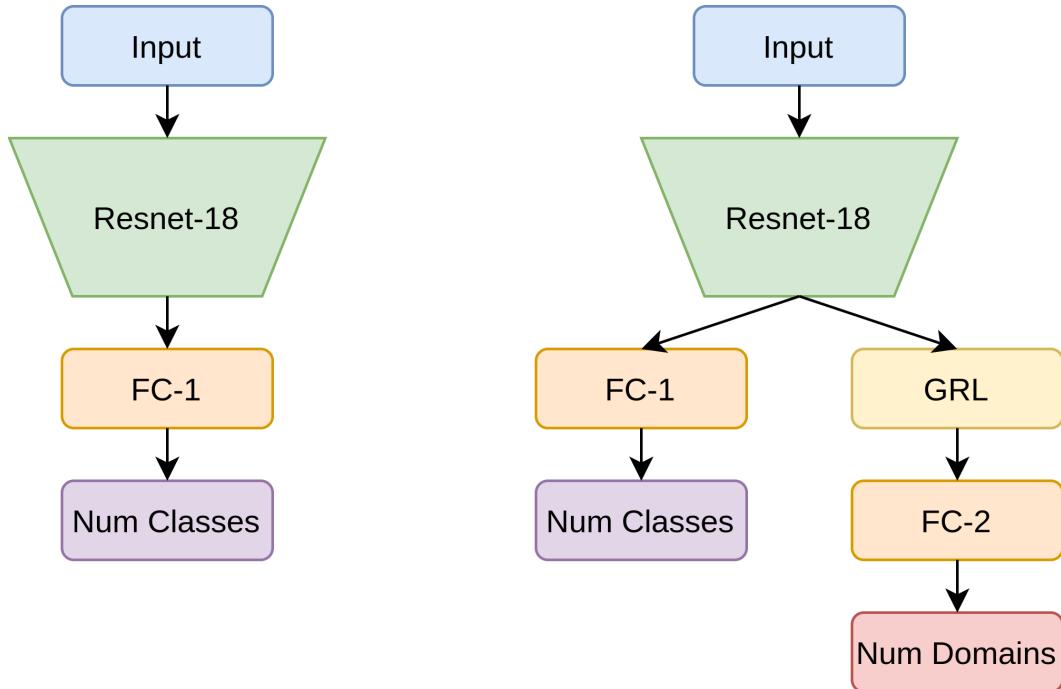


Figure 1: (Left) Supervised model architecture. (Right) Unsupervised model architecture.

43 **3 Experiment Embeddings**

44 We present the t-SNE embeddings for all our experiments in the section.

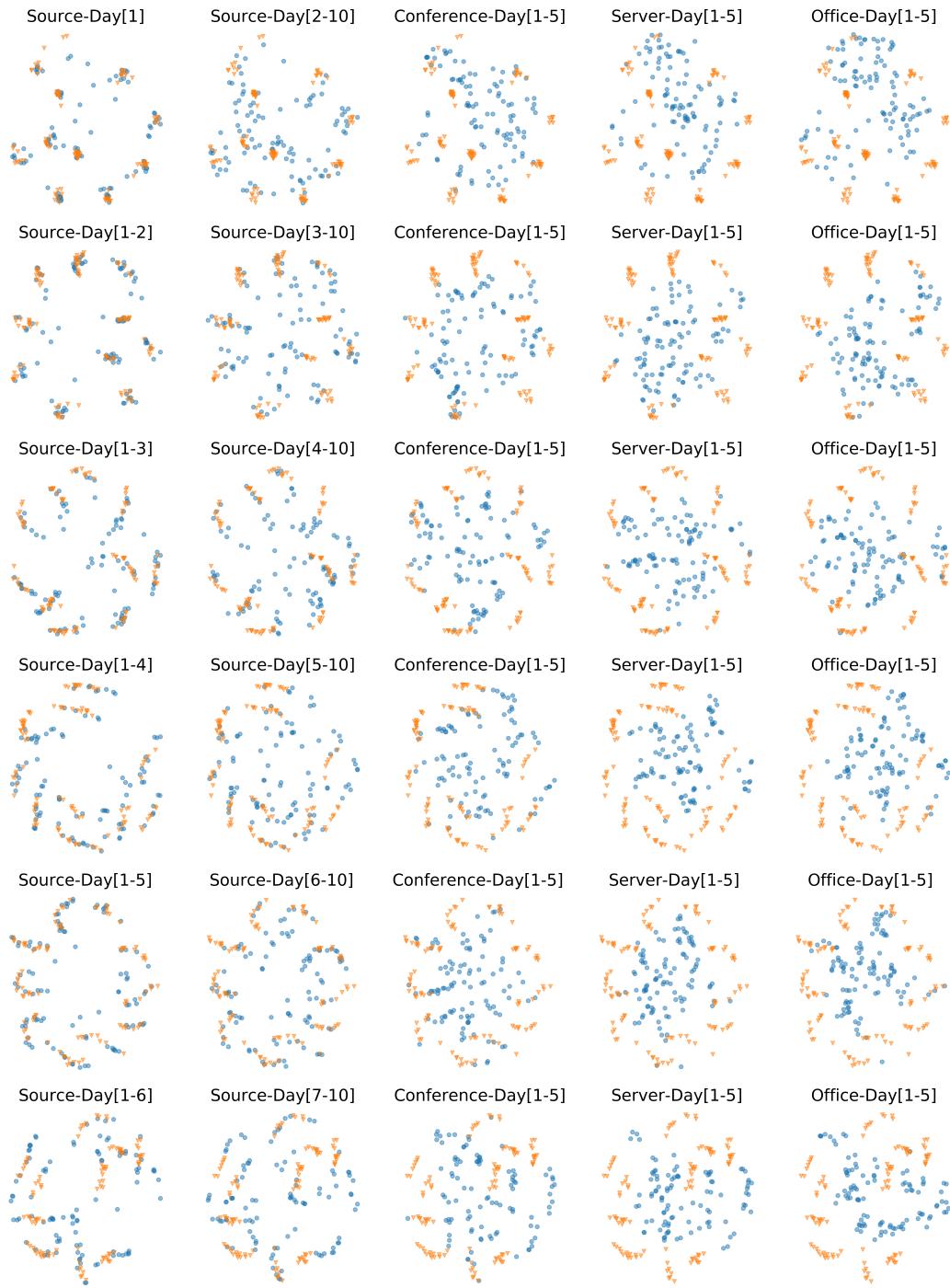


Figure 2: Embeddings from the model trained using 1 to 6 days of source-only data. The number in brackets indicates the day indices. For example, [1 - 6] means the data from day 1 to day 6. Orange triangle clusters represent the training data (Source Location's Day 1), where each cluster represents a distinct human subject. Blue circles represent each domain's testing data, where the data points were sampled from all the classes of a given domain.

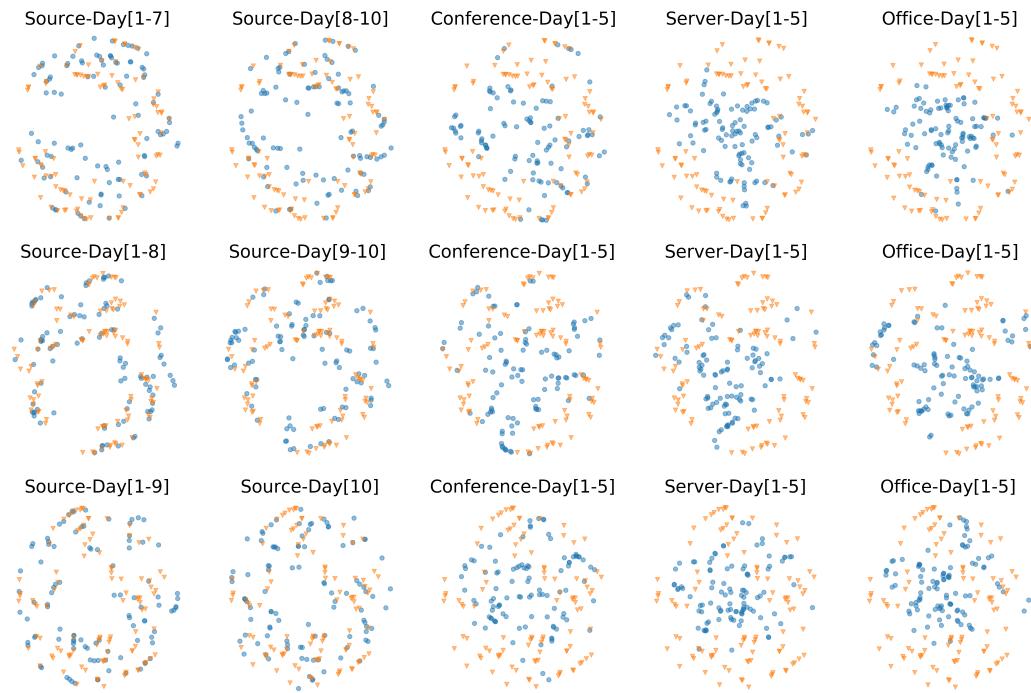


Figure 3: Embeddings from the model trained using 7 to 9 days of source-only data

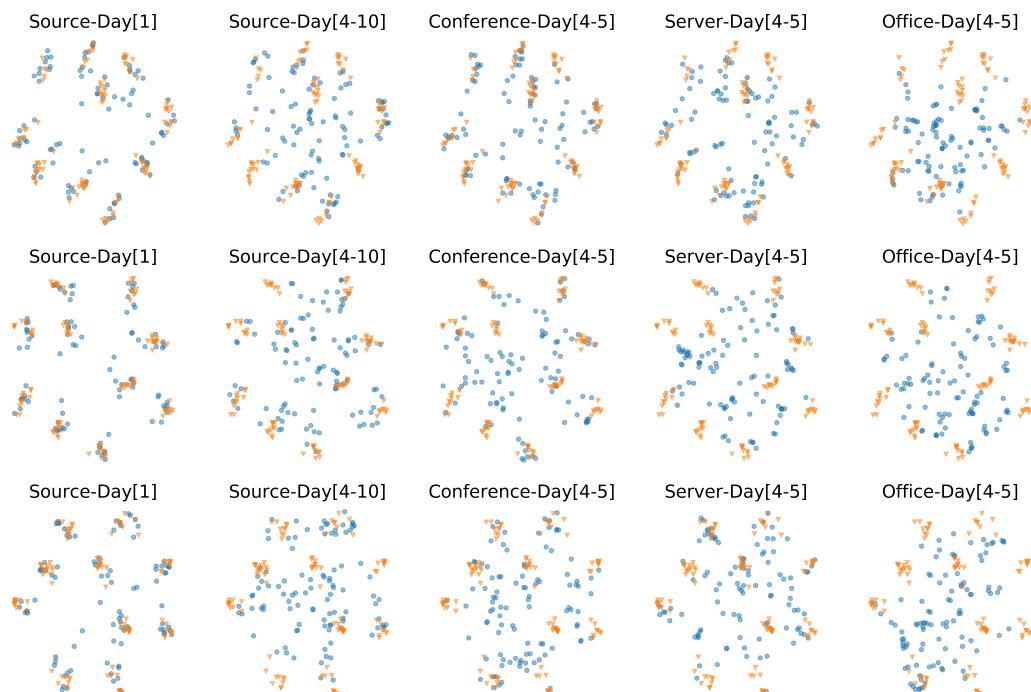


Figure 4: Embeddings from model trained on source and target domain data. (Top) model trained one day of source, server and conference data, (Middle) model trained via one day of source and conference data, (Bottom) model trained via one day of source and server data

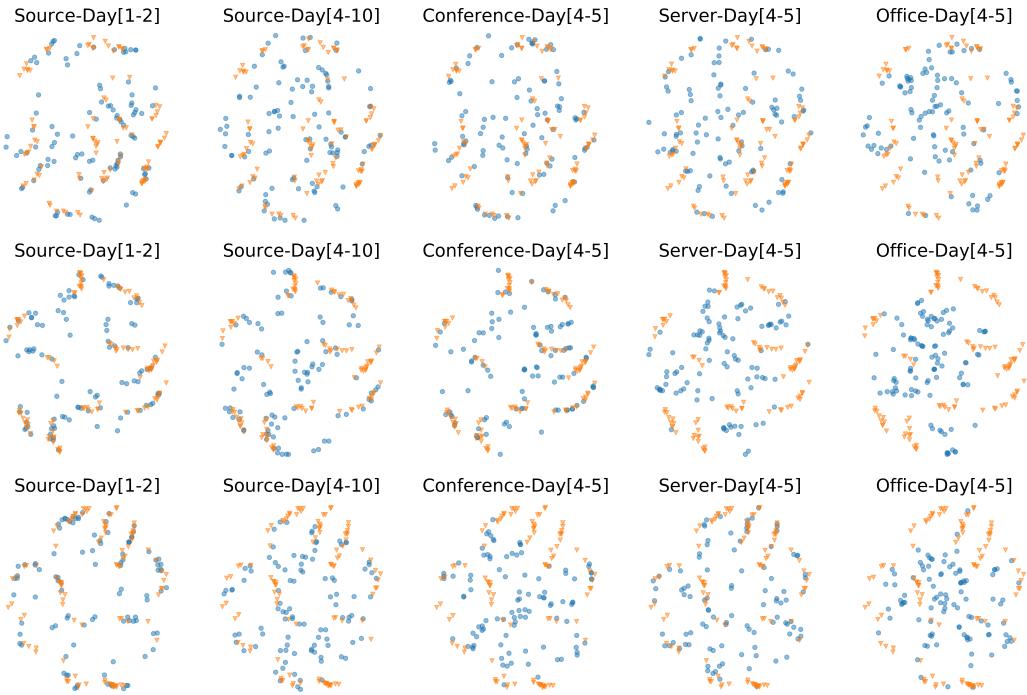


Figure 5: Embeddings from model trained on source and target domain data. (Top) model trained via two days of source, server and conference data, (Middle) model trained via two days of source and conference data, (Bottom) model trained via two days of source and server data

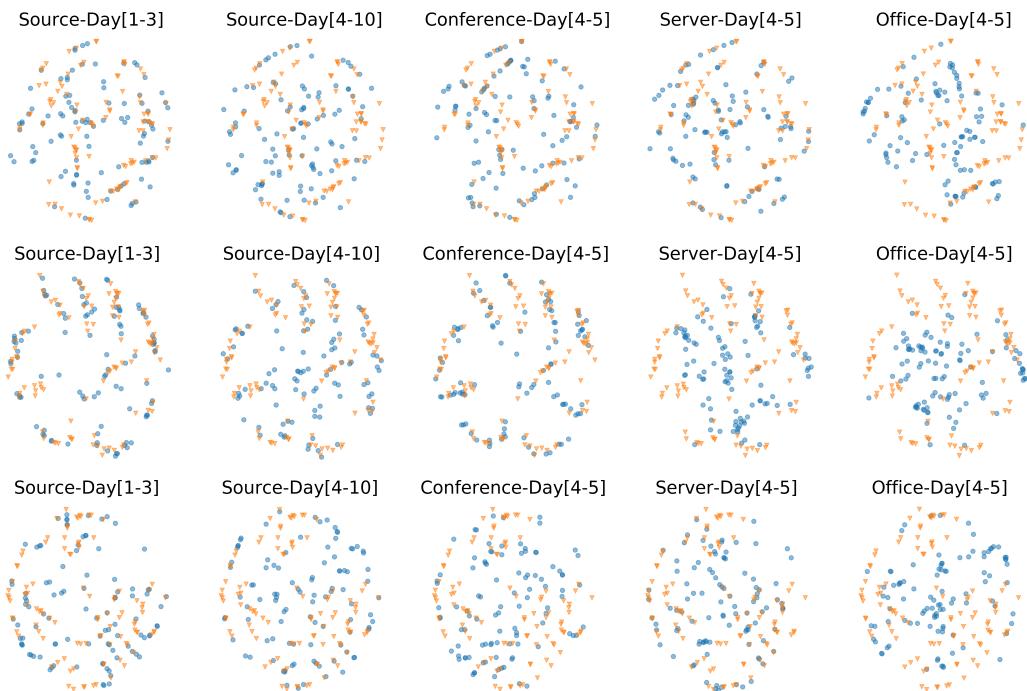


Figure 6: Embeddings from model trained on source and target domain data. (Top) model trained via three days of source, server and conference data, (Middle) model trained via three days of source and conference data, (Bottom) model trained via three days of source and server data

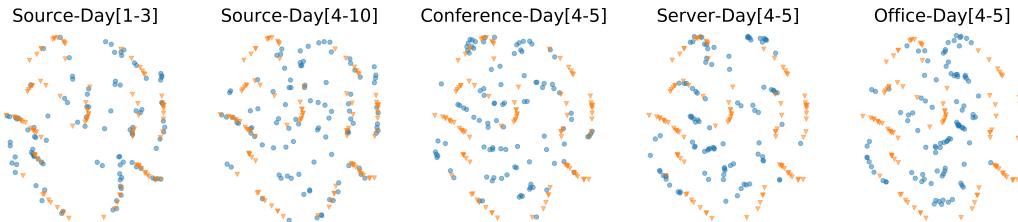


Figure 7: Embeddings from model trained on three days of labeled source and three days of unlabeled server domain data. The unsupervised data was trained using a GAN based approach.

45 4 Experiment Shift Metrics

46 Here we present the accuracies obtained from all our experiments alongside the corresponding
 47 metric learning based distance metrics. We find that not only does the distance metrics agree with
 48 the accuracy based findings, it also makes it more evident. Indeed we see an apparent temporal
 49 degradation in Fig. 8 in the source domain. Moreover, introducing more training data from the source
 50 domain alleviates the SDS to a certain extent, but it also biases the model more towards the source
 51 domain in some cases.

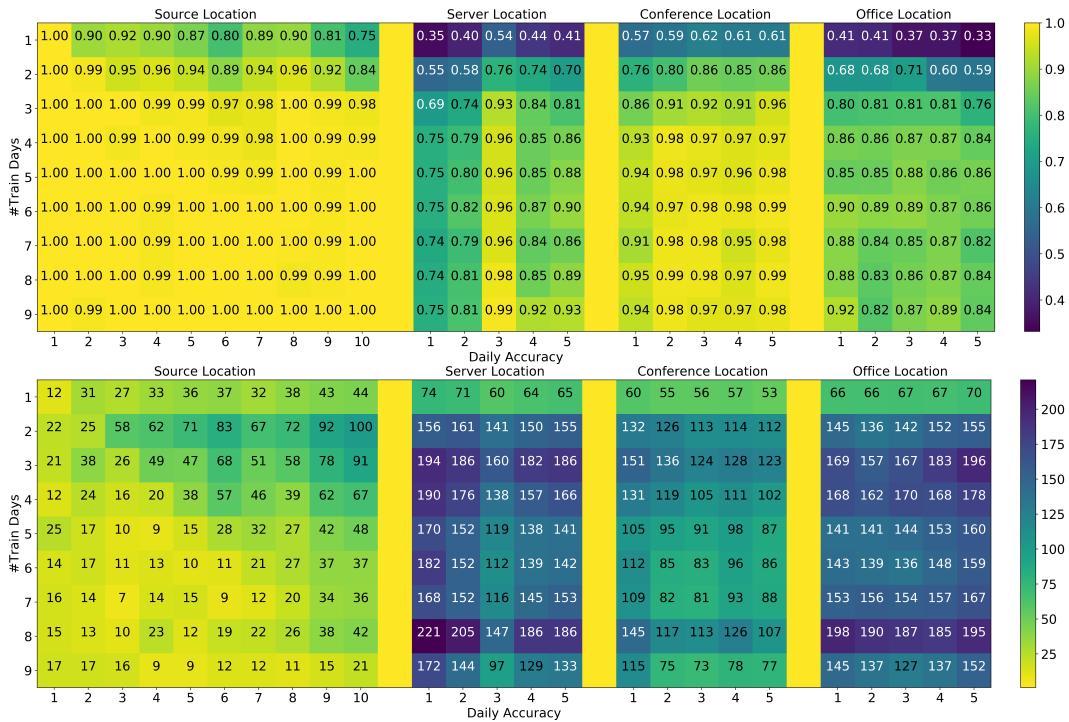


Figure 8: Results obtained from training only on source location data for varying number of days.

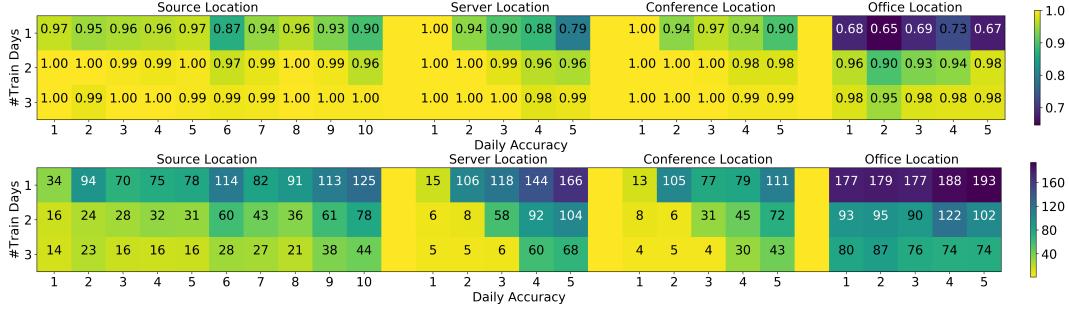


Figure 9: Results obtained from training on source, server, and conference location data for varying number of days. The data from office location is not used for training and only for testing

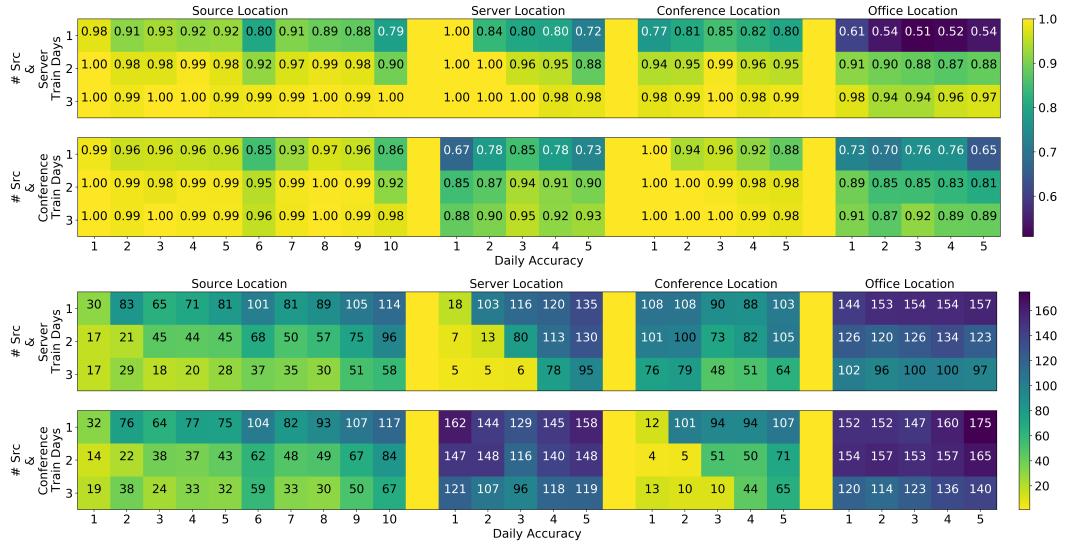


Figure 10: (Top) Results obtained from training on the source and server data for a varying number of days. (Bottom) Results obtained from training on the source and conference data for a varying number of days. Data from office location is for testing only

5 Metric Learning Based Shift and Malignancy Detection Algorithm

We utilize Algorithm 1 to compute thresholds used for detecting outliers. In our experiments, L2 distance was used along with a margin m of 10. Upon computing class thresholds $Threshold$ and trained model M , any new data samples that were observed, were subject to class-wise thresholding $DistanceMetric$ between the known class centers C^{Train} and the generated embedding for the new data point.

Algorithm 1 Algorithm to Compute Class-Wise Thresholds

Input: Training data X_{Train} , validation data X_{Val} , number of classes n , $DistanceMetric$ (L2 or cosine distance), hyperparameter safety margin m .

Output: Class-wise threshold $\{Threshold_j | j = 1, 2, \dots, n\}, M, C^{Train}$

- 1: Train model M on X_{Train} with metric learning
- 2: Generate class-wise embeddings $\{C_j^{Train} | j = 1, 2, \dots, n\}$ from X_{Train} and M
- 3: Generate class-wise embeddings $\{C_j^{Val} | j = 1, 2, \dots, n\}$ from X_{Val} and M
- 4: Compute $\{Threshold_j | j = 1, 2, \dots, n\} = DistanceMetric(C_j^{Val} - C_j^{Train} | j = 1, 2, \dots, n) + m$

58 **6 GRL, GAN and Student-Teacher Methods for Unsupervised Domain
59 Adaptation**

60 We addressed unsupervised domain adaptation using a GAN based approach. However, we also
61 tried a few other methods. The most notable of which were, based on GRL and Student-Teacher(ST)
62 based methods. We report the accuracies obtained from each method in Table 1. These models
63 were trained using three days of labeled source location data and three days of unlabeled server
64 location data. Apart from the three methods mentioned above, we also tried introducing several other
65 methods. Specifically, in combination with the model with GRL and AMCA loss, we tried each
66 of the methods in [8, 9, 10, 11, 12, 13] including mixup/cutmix regularization, adversarial training,
67 pseudo-labels, multi-adversarial and cyclic adversarial domain adaptations. However, none of these
68 methods amounted to any significant results. We are not sure if these methods failed as a consequence
69 of being used along with metric learning or because the data are in the format of spectrograms. We
70 leave investigations into the matter for our future work.

Table 1: Model Accuracies (%). (GRL) The model trained with GRL based UDA. (GAN) The model
trained with GAN based UDA. (ST) The model trained with student-teacher based UDA

	Source-Temporal	Server-Test	Conference-Test	Office-Test
GRL	99.37	97.80	99.40	94.96
GAN	99.61	98.70	99.70	97.47
ST	97.77	86.99	95.50	85.07

71 **References**

- 72 [1] Iwr1642 single-chip 76-ghz to 81-ghz mmwave sensor integrating dsp and mcu evaluation
73 module iwr1642boost this product has been released to the market and is available for purchase.
74 for some products, newer alternatives may be available.
- 75 [2] Prabhu Janakaraj, Kalvik Jakkala, Arupjyoti Bhuyan, Zhi Sun, Pu Wang, and Minwoo Lee. Star:
76 Simultaneous tracking and recognition through millimeter waves and deep learning. In *2019*
77 *12th IFIP Wireless and Mobile Networking Conference (WMNC)*, pages 211–218. IEEE, 2019.
- 78 [3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image
79 recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*,
80 pages 770–778, 2016.
- 81 [4] Jun-Bo Liu, Ya-Ping Huang, Qi Zou, and Sheng-Chun Wang. Learning representative features
82 via constrictive annular loss for image classification. *Applied Intelligence*, 49(8):3082–3092,
83 2019.
- 84 [5] Feng Wang, Jian Cheng, Weiyang Liu, and Haijun Liu. Additive margin softmax for face
85 verification. *IEEE Signal Processing Letters*, 25(7):926–930, 2018.
- 86 [6] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François
87 Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural
88 networks. *The Journal of Machine Learning Research*, 17(1):2096–2030, 2016.
- 89 [7] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil
90 Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural
91 information processing systems*, pages 2672–2680, 2014.
- 92 [8] Hongyi Zhang, Moustapha Cisse, Yann N Dauphin, and David Lopez-Paz. mixup: Beyond
93 empirical risk minimization. *arXiv preprint arXiv:1710.09412*, 2017.
- 94 [9] Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, and Youngjoon
95 Yoo. Cutmix: Regularization strategy to train strong classifiers with localizable features. In
96 *Proceedings of the IEEE International Conference on Computer Vision*, pages 6023–6032,
97 2019.

- 98 [10] Ian J Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*, 2014.
- 99
- 100 [11] Dong-Hyun Lee. Pseudo-label: The simple and efficient semi-supervised learning method for
101 deep neural networks. In *Workshop on challenges in representation learning, ICML*, volume 3,
102 page 2, 2013.
- 103 [12] Zhongyi Pei, Zhangjie Cao, Mingsheng Long, and Jianmin Wang. Multi-adversarial domain
104 adaptation. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.
- 105 [13] Ehsan Hosseini-Asl, Yingbo Zhou, Caiming Xiong, and Richard Socher. Augmented cyclic
106 adversarial learning for low resource domain adaptation. *arXiv preprint arXiv:1807.00374*,
107 2018.