Domain adaptation for target classification using micro-Doppler spectra in radar networks

Svenningsson, Peter; Fioranelli, Francesco; Yarovoy, Alexander

Publication date
2021

Document Version
Final published version

Published in
2021 IEEE 24th International Conference on Information Fusion (FUSION)

Citation (APA)
Svenningsson, P., Fioranelli, F., & Yarovoy, A. (2021). Domain adaptation for target classification using micro-Doppler spectra in radar networks. In *2021 IEEE 24th International Conference on Information Fusion (FUSION): Proceedings* (pp. 1-8). Article 9626930 IEEE. https://ieeexplore.ieee.org/document/9626930

Domain adaptation for target classification using micro-Doppler spectra in radar networks

Peter Svenningsson, Francesco Fioranelli and Alexander Yarovoy
Microwave Sensing, Signals and Systems (MS3)
Delft University of Technology
Delft, The Netherlands
{p.o.svenningsson, F.Fioranelli, a.yarovoy}@tudelft.nl

***Abstract*—In this paper, the classification of human activity from micro-Doppler spectrograms measured by a radar network is considered. To cope with differences between the training and test datasets due to changes in the set of participants, signal-to-noise ratio and polarimetry, domain adaptation is proposed. To realize this, linear mapping between the two domains is assumed and estimated by one of two methods, expectation-maximization or empirical estimates of statistical moments. The performance of the methods is evaluated on experimental data measured by a multi-static radar network. The proposed methods increase the classification accuracy by 5–15 percentiles on the recorded dataset.**

***Index Terms*—radar target classification, micro-Doppler signature, domain adaptation, multi-static radar network**

I. Introduction

A radar sensing network comprises radar nodes that are able to observe a target under different observation conditions such as the orientation to the target, polarimetry and signal-to-noise ratio (SNR). Target classification may benefit from the diversity in observation conditions and the increased number of measurements provided by the radar nodes. However, there exist difficulties in creating a classification model which generalizes well to the data from all sensing nodes due to the unavailability of labeled data under all possible observation conditions. This work explores the utilization of unlabeled data from each sensing node to adapt the classification model to the specific observation conditions of the sensing node and proposes a novel domain adaptation method based on expectation-maximization.

The performance of a classification model will generally decrease as the test conditions become different than the conditions the training data was drawn from [1]. In a general setting this problem is referred to as domain shift. Such differences can appear in many practical applications including radar networks, where it plays an important role due to changes in the radar nodes' observation conditions such as the SNR of the received signal, polarimetry and the multi-static scattering angle. Consistent with the nomenclature used in [1], the distribution to which the classification model was fit is referred to as the source distribution and the data distribution in which we are attempting to perform inference where some latent factor has changed is referred to as the target distribution. Techniques to alleviate the effects of domain shift are here referred to as domain adaptation. These techniques make use of the dissimilarity between the source and target domains to increase the classification performance in the target domain.

Recent work in image classification has explored techniques that use empirical estimates of feature moments in a target domain in order to increase the classification performance of convolutional neural networks (CNN) [2], [3], here referred to as standardization methods. In these works, the parameters of a batch normalization layer [4] are estimated from the target domain. This method showed positive results when applied to a CNN which was fit on a dataset of handwritten numbers, MNIST, and evaluated on a dataset of house numbers, SVHN [3].

Related work in pattern recognition has also explored the application of expectation-maximization to fit a classification model jointly to the labeled data in the source domain and the unlabeled data in the target domain [5]. The method treats domain adaptation as a missing data problem and assumes that the source and target data are drawn from the same domain.

Previous works in domain adaptation which focuses on feature standardization in the target domain [2] [3] [4], assumes that the data in the target domain can be characterized by its mean and variance. Sampling bias may cause some classes to be missing from the target domain data and the standardization-based techniques will then provide an incorrect estimate of data distribution in the target domain. In surveillance applications such as drone detection, the class distribution in the target domain may be significantly different than in the source domain as the drone class may be missing from the target domain data. Therefore domain adaptation methods that can mitigate the effect of sampling bias are in demand.

This work focuses on the classification of human activity from micro-Doppler spectrograms derived from the radar data; however, the presented methodology is generally applicable for classification tasks. In literature, many classification algorithms have been proposed for radar target classification using the micro-Doppler spectrum [6] such as Markov models of micro-Doppler characterizations [7] and learnt representations [8]–[10] In this work, we validate our methodology the following models: a support-vector machine with a radial basis kernel [11] and a Gaussian naive Bayes classifier [12].

The above-mentioned models will be considered as discriminative models which estimate the conditional distribution

$p(y \mid x)$ of the class variable y given the features x. We will consider that the model was fit in a source domain with the conditional distribution $p_{\mathcal{S}}(y \mid x)$ and is evaluated in a target domain with distribution $p_{\mathcal{T}}(y \mid x)$ and that there exists some function $f(\cdot)$ which maps the target distribution to the source distribution. The mapping $f(\cdot)$ is either parameterized as a translation which is estimated by expectation-maximization or parametrized as a linear function and estimated by matching statistical moments between the source and target domains. We show on experimental data that we are able to increase the classification performance in the target domain by utilizing this mapping.

This work investigates the effectiveness of domain adaptation on data gathered from a coherent pulsed multi-static sensing network. It is investigated how domain adaptation improves the fused classification of a radar sensing network. The latent factors which differentiate the target and source domain are: the bistatic angle to the target, the receiver polarimetry and changes to signal-to-noise ratio (SNR). Recent research indicates that the variation in micro-Doppler signature between individuals is significant [13]. Therefore, adaptation to previously unseen individuals is investigated as well.

In summary, the contributions of this paper are:

- The first published investigation of domain adaptation for target classification using micro-Doppler spectra.
- A novel method for domain adaptation based on expectation-maximization which is verified on experimental data from a multi-static radar network and provides positive results in terms of increasing classification accuracy.

The signal model is presented in Section II while the proposed domain adaptation method and evaluation methodology are presented in Section III and IV respectively. The experimental data and its characterization is presented in Section V. Results and concluding remarks are found in Section VI and VII.

II. Signal model

A radar receiver measures the backscatter signal $s(t)$ with an instantaneous frequency spectrum $\hat{s}(t,\omega)$. The instantaneous spectrum is here characterized by 9 features that describe the micro-Doppler signature. Let μ_ω, σ_ω denote the first two moments of the frequency response at any frequency and time. The size of the frequency silhouette is defined as the proportion of $\hat{s}(t,\omega)$ with a larger response than $\mu_\omega - 0.3\sigma_\omega$. The contour of $\hat{s}(t,\omega)$ is defined by the highest and lowest frequencies with a response larger than $\mu_\omega + 0.3\sigma_\omega$. Frequency peaks are extracted from the upper contour. The features then comprise: the size of the frequency silhouette, the first two moments of the contour, peak height, Doppler centroid and Doppler Bandwidth as defined in [12]. For reproducibility, code to generate these features has been made available[1].

[1] https://github.com/petersvenningsson/feature-extraction-micro-doppler

Input: $X_{\mathcal{S}}$, $X_{\mathcal{T}}$, $\mathcal{M}$

$X_{\mathcal{S}}, y_{\mathcal{S}}$ The source domain data and labels
$X_{\mathcal{T}}$ The unlabeled target domain data
$\mathcal{M}$ The classification model
μ_k, Σ_k, w_k The source domain feature mean, covariance matrix and class distribution of class k
$T_{n,k}$ Probability of sample n belonging to class k
$\hat{\theta}_\mu$ Estimated feature translation
L Total likelihood of the target domain data

Output: $\hat{y}_{\mathcal{T}}$ Target domain class predictions

while $|L^{(t)} - L^{(t-1)}| << 1$ **do**

$$T^{(t)}_{n,k} = \frac{w_k \mathcal{N}\left(\mathbf{x}_n | \mu_k - \hat{\theta}^{(t-1)}_\mu, \Sigma_k\right)}{\sum_{j=1}^{K} w_j \mathcal{N}\left(\mathbf{x}_n | \mu_j - \hat{\theta}^{(t-1)}_\mu, \Sigma_j\right)}$$

$$\hat{\theta}^{(t)}_\mu = \frac{\left(\sum_{n=1}^{N}\sum_{k=1}^{K} T^{(t)}_{n,k} \log w_k (\mu_k - x_n)^T \Sigma_k^{-1}\right)}{\left(\sum_{n'=1}^{N}\sum_{k'=1}^{K} T_{n',k'} \log w_k \Sigma_k^{-1}\right)}$$

$$L(X_{\mathcal{T}})^{(t)} = \prod_{n=1}^{N}\sum_{k=1}^{K} w_k \mathcal{N}(x_n \mid \mu_k - \theta_\mu, \Sigma_k)$$

$$t \leftarrow t + 1$$

Fit $\mathcal{M}$ on $X_{\mathcal{S}}$, $y_{\mathcal{S}}$

$\hat{y}_{\mathcal{T}} \leftarrow \mathcal{M}(X_{\mathcal{T}} - \hat{\theta}^{(t)}_\mu)$

return $\hat{y}_{\mathcal{T}}$

Fig. 1: The proposed domain adaptation method based on expectation-maximization.

III. Method

The proposed method of expectation-maximization based domain adaptation is described below as well as the standardization method which is used as a benchmark. In both these methods, we assume that the classification model has high performance in the source domain. Therefore we estimate a function $f(\cdot)$ which maps the target domain to the source domain such that,

$$p_{\mathcal{S}}(y \mid x) \approx p_{\mathcal{T}}(y \mid f(x)). \tag{1}$$

A. Expectation-maximization

In the proposed method we assume that the feature-distribution for each class is multivariate normal,

$$p_{\mathcal{S}}(x \mid y_k) = \mathcal{N}(x \mid \mu_k, \Sigma_k), \quad k \in \{1, \cdots, K\}$$

with the parameters μ_k, Σ_k estimated from labeled data drawn from the source domain for the K classes. The choice of the analytical form of the distribution is motivated by that the sample distributions are approximately unimodal and symmetrical. We also assume that the domain shift can be parameterized as a translation in each feature dimension,

$$f(x) = x - \theta_\mu, \tag{2}$$

where θ_μ denotes the translation parameter. We can then describe $p_{\mathcal{T}}(x)$ as a mixture of Gaussians,

$$p_{\mathcal{T}}(x) = \sum_k w_k \mathcal{N}(x \mid \mu_k - \theta_\mu, \Sigma_k),$$

with total likelihood,

$$L(X_{\mathcal{T}}) = \prod_{n=1}^{N} \sum_{k=1}^{K} w_k \mathcal{N}(x_n \mid \mu_k - \theta_\mu, \Sigma_k), \tag{3}$$

where $w = p_{\mathcal{S}}(y)$ and $X_{\mathcal{T}}$ denotes the N data samples drawn from the target domain.

With the objective to find θ_μ which maximizes (3) we iteratively estimate the class-membership T of each data point in $X_{\mathcal{T}}$,

$$T_{n,k}^{(t)} = \frac{w_k \mathcal{N}\left(\mathbf{x}_n \mid \mu_k - \hat{\theta}_\mu^{(t-1)}, \Sigma_k\right)}{\sum_{j=1}^{K} w_j \mathcal{N}\left(\mathbf{x}_n \mid \mu_j - \hat{\theta}_\mu^{(t-1)}, \Sigma_j\right)} \tag{4}$$

and the maximum likelihood estimator of $\theta_\mu^{(t)}$ given $\theta_\mu^{(t-1)}$,

$$\hat{\theta}_\mu^{(t)} = \frac{\left(\sum_{n=1}^{N} \sum_{k=1}^{K} T_{n,k}^{(t)} \log w_k (\mu_k - x_n)^T \Sigma_k^{-1}\right)}{\left(\sum_{n'=1}^{N} \sum_{k'=1}^{K} T_{n',k'} \log w_k \Sigma_k^{-1}\right)}, \tag{5}$$

where (4) and (5) defines the expectation and maximization steps respectively.

The estimation of $f(\cdot)$ and the subsequent generation of class predictions in the target domain is described in Algorithm 1. In each iteration, the total likelihood increases monotonically and may converge to a local minima or a saddle point [14]. The termination criteria is here defined as an increase in total likelihood lower than 10^{-6} between two iterations.

In summary, the function $f(\cdot)$ which maps the data from the source domain to the target domain is parameterized as a simple translation in each feature dimension. The translation is estimated by maximizing the total likelihood of a statistical model which assumes that the feature distribution of each class is normally distributed.

B. Standardization

The proposed domain adaptation method is compared to the standardization method which estimates $f(\cdot)$ by matching the mean and variance of the source and target domain data. The standardization method assumes that the function $f(\cdot)$ can be parameterized independently for each feature as,

$$f(x) = \frac{x - \theta_\mu}{\theta_\sigma}, \tag{6}$$

where θ_μ and θ_σ defines a translation and a scaling of the feature respectively and that the mapping between the source and target domains can be defined as,

$$p_{\mathcal{S}}(x) = p_{\mathcal{T}}(f(x)). \tag{7}$$

The parameters in (6) are estimated by the sample mean and sample variance in both domains given the relationships

$$\mu_{\mathcal{S}} = \frac{\mu_{\mathcal{T}} - \theta_\mu}{\theta_\sigma}, \quad \sigma_{\mathcal{S}}^2 = \frac{\sigma_{\mathcal{T}}^2}{\theta_\sigma^2}, \tag{8}$$

which follow from (7), where $\mu_{\mathcal{S}}$, $\sigma_{\mathcal{S}}^2$ and $\mu_{\mathcal{T}}$, $\sigma_{\mathcal{T}}^2$ denotes the mean and variance of the source and target domains respectively.

In summary, a model which is fit from data in the source domain is evaluated on data in the target domain by first mapping the features to the target domain. The mapping $f(x)$ is estimated by the sample mean and sample variance of these domains using the relationships in (8).

IV. Evaluation

A classifier is trained in a source domain and evaluated on data drawn from a target domain in which some latent factor has changed. The domain adaptation method described in Section III is evaluated by the increase in accuracy in the target domain, in comparison to the standardization method and model evaluation without adaptation. The three latent variables are considered here are:

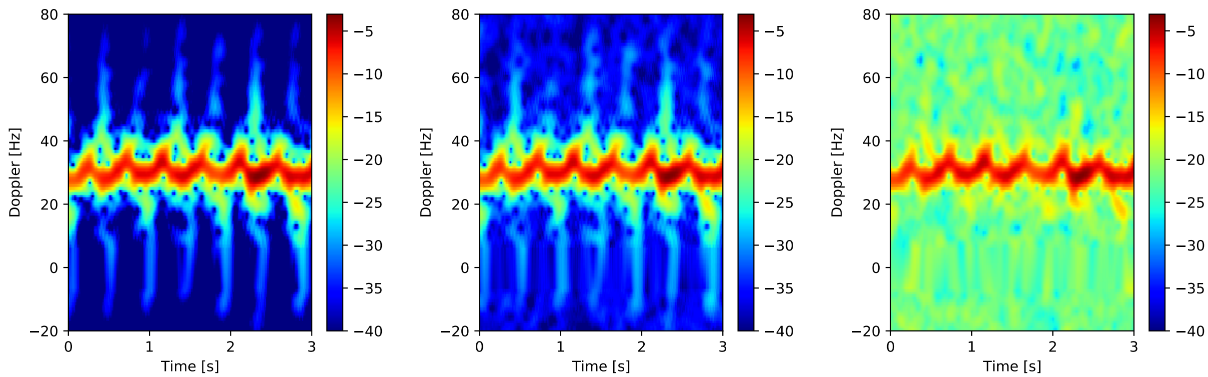
- The target domain consists of data drawn from individuals not seen in the source domain, here referred to as leave-one-participant-out (LIPO) validation.
- The target domain consists of samples with lower SNR than the source domain.
- The target and source domain consists of data measured from two different radar nodes, signifying changes to polarimetry and the bistatic angle.

With the aim to evaluate the domain adaptation method quantitatively, a large number of classification models are evaluated. The performance of the domain adaptation method is characterized as the mean performance of the classification models which are fit to different feature sets and combinations of the relevant latent variables. For instance, when domain adaptation across the three sensing nodes is considered there exists 6 combinations of the latent factor. To ensure sufficient model flexibility only feature sets larger than 6 are considered and stratified 5-fold validation is used where applicable. For each experiment 3×10^4 models are generated.

The effect of sampling bias in the target domain is evaluated by removing a number of classes from the target domain data used by the domain adaptation method. The model is then evaluated on the complete dataset in the target domain.

V. Experimental setup and data processing

The data is recorded using the multi-static coherent pulsed radar NetRAD, developed by University College London and the University of Cape Town. The sensing network comprises three nodes as visualized in Fig. 3. Node 1 is a monostatic transceiver that receives in V polarization and transmits in V polarization a linearly modulated up-chip with bandwidth 45 MHz, center frequency 4.2 GHz 18 dBi gain, 23 dBm transmit power and PRF 5 kHz. Node 2 is co-located with the transceiver and records H polarization. Node 3 is a bistatic node that records co-polarized data. The three nodes operate under the same clock cycle which is synchronized by a wired connection. Each node uses an independent local oscillator, with residual drifts between different oscillators corrected in post-processing as described in [15].



(a) No added noise (b) 20 dB SNR (c) 5 dB SNR

Fig. 2: Spectrograms containing the micro-Doppler signature of a person walking towards the monostatic radar node. The SNR of the received signal has been reduced by the addition of complex Gaussian noise. The amplitude of the signal is visualized in logarithmic scale (dB).

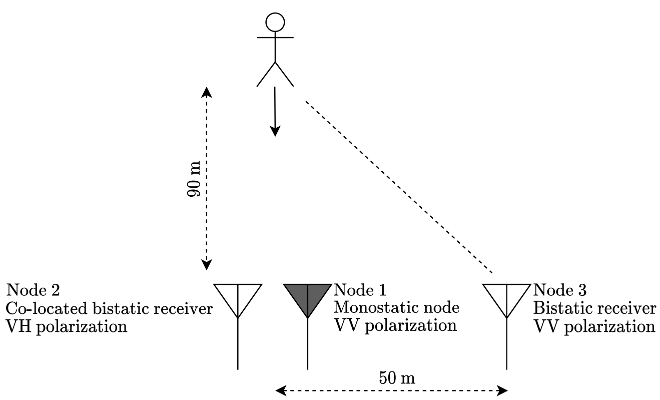


Fig. 3: A diagram of the coherent pulsed multi-static radar sensing network used to record the dataset.

A measurement sequence consists of one individual walking towards Node 1. The class variable specifies what the person is carrying. A non-exhaustive list of the seven available classes is: *walking*, *walking while carrying a backpack* and *walking while carrying a steel pipe*. Four individuals recorded 10 measurements of 10 s sequences for each class. Each sequence is subdivided into 3 s non-overlapping samples. The dataset was recorded in 2016 and has been used in previous research [12], [11] A complete description of the dataset is found in [12].

A short-time Fourier transform is used to generate the instantaneous frequency spectrum of the radar measurements which describe the micro-Doppler signature of the movement. A 0.3 s Hamming window function is used with 95 % overlap. Stationary clutter in the environment generates a strong response across several Doppler frequencies. These are replaced by linearly interpolating from the nearest unaffected Doppler frequencies as performed in [16]. When applicable, the SNR is reduced by adding complex Gaussian noise to each pulse in the measurement as described in [11]. A visualization of a spectrogram with and without added noise can be shown in Fig. 2.

VI. EXPERIMENTAL RESULTS

Results from the evaluation as outlined in Section IV are presented below. Results from the multi-static radar network fusion scenario are found in Section VI-A while in Section VI-B–VI-E results regarding the adaptation of individual nodes is presented.

A. Multi-static radar network fusion scenario

A classifier is trained on data from the mono-static Node 1. The presented domain adaptation methods are used to adapt the classifier to each radar node. Decision-level fusion is used to generate a prediction from the radar network which is evaluated in the LIPO setting. The results found in Fig. 4 show that the domain adaptation methods are able to increase the performance of the multi-static radar network by 5–15 percentiles. The proposed expectation-maximization based method outperforms standardization when fewer than three classes are present in the target domain data due to sampling bias.

B. Adaptation to new individuals LIPO

A classifier is trained on data from three individuals and validated on one individual not found in the training data. The results are shown in Fig. 5. Both domain adaptation techniques are able to increase the average performance of the classification models by 8–10 percentiles when all 7 classes are present in the target domain, with standardization outperforming expectation-maximization. When fewer classes than 3 are present in the target domain, the expectation-maximization based approach outperforms standardization. Note that standardization is detrimental to the classification performance when fewer than three classes are present in the target domain.

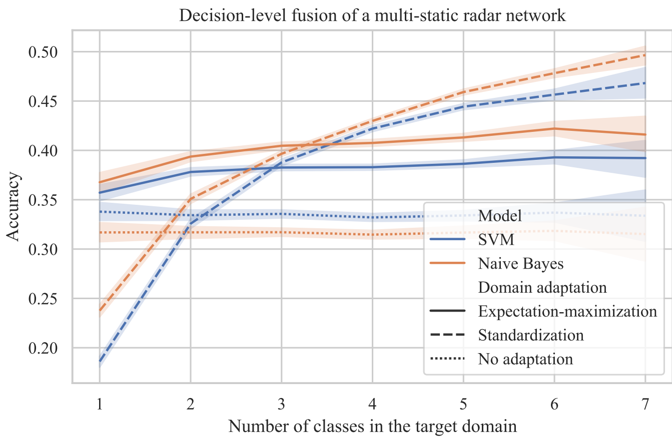


Fig. 4: Domain adaption applied to a multi-static radar network. A classifier is trained on data from the mono-static Node 1 and the classifier is adapted to each radar node. The performance of the network is evaluated by decision-level fusion evaluated in the LIPO setting. The performance of the domain adaptation is characterized as the mean accuracy of a large number of models with the shaded region indicating a 95 % confidence interval.

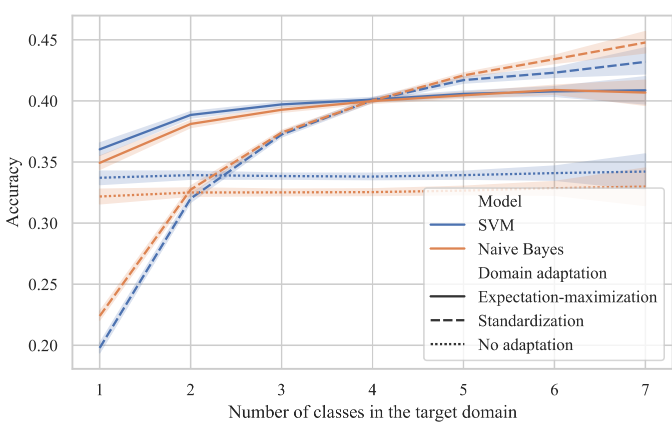


Fig. 5: The effectiveness of domain adaptation when the individuals in the target domain are not found in the source domain. Expectation-maximization outperforms standardization when fewer than half of the classes are present in the target domain. The shaded region indicates a 95 % confidence interval of the mean performance of the classification models.

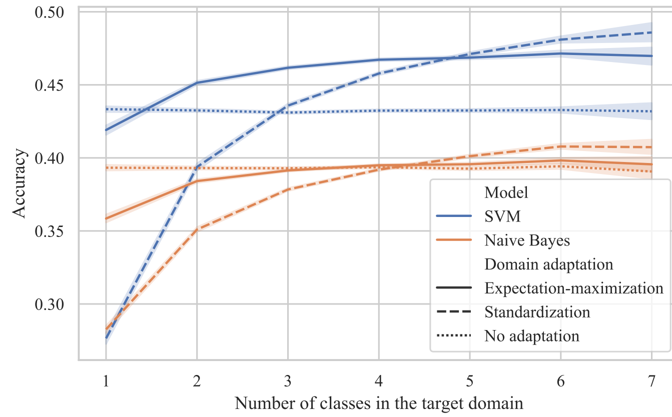


Fig. 6: The performance of domain adaptation when a classification model is transferred between two radar nodes. Indicating a change in polarimetry and or bistatic angle. The graph shows the mean accuracy over 3×10^4 models fit with varying feature sets and source/target node combinations.

C. Adaptation to changes received polarization and multi-static scattering angle

A classification model is fit on data from a source node and evaluated on data from a different target node. The domain shift is then a consequence of changes to received polarization and bistatic angle between the source and target domain. The results shown in Fig. 6 indicate that the expectation-maximization method outperforms the standardization method when fewer than half of the classes are present in the target domain. When all classes are present in the target domain, the domain adaptation methods provide an increase in performance of 5 percentiles for the SVM model and provide no significant increase for the Naive Bayes model.

The impact of domain adaptation when a model is transferred between sensor nodes is also evaluated using LIPO validation. The results are shown in Fig. 7. In this setting, the differences between the source and target domain are greater and the domain adaptation techniques are more able to increase the classification performance than in the setting without LIPO validation. Consistent with the LIPO results in Section VI-B, the expectation-maximization technique outperforms standardization when less than half of the classes are present in the target domain. When all classes are present in the target domain, the domain adaptation methods are able to increase the classification accuracy by 8–12 percent with standardization outperforming expectation-maximization.

D. Adaptation to reduced SNR

The SNR is reduced by adding complex Gaussian noise to each recorded pulse. A model is fit in a source domain with no added noise and evaluated in a target domain where the SNR has been reduced. The performance of the domain adaptation methods are shown in Fig. 8 when all classes are present in the target domain. Both domain adaptation

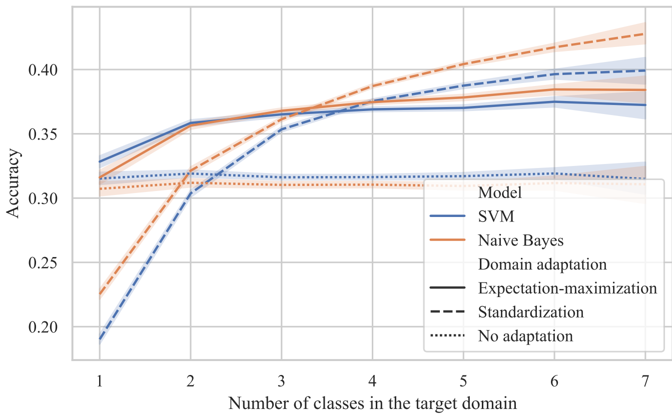


Fig. 7: The performance of domain adaptation when a classification algorithm is transferred between two radar nodes and validation is performed in the LIPO setting. The domain adaptation algorithms show stronger performance in the LIPO setting as the source and target domains are less similar.

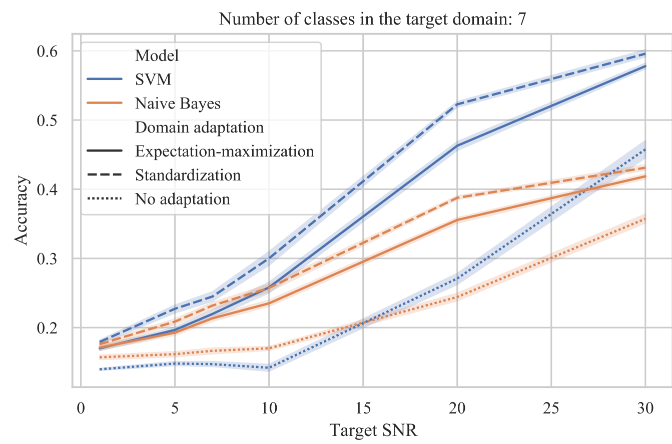


Fig. 8: The performance of the domain adaption methods when the SNR in the target domain has been reduced by the addition of complex Gaussian noise. All seven classes are present in the target domain. The shaded region indicates a $95\,\%$ confidence interval of the mean performance of the classification models.

methods give a performance increase of 5–20 percentiles. In this setting, the performance of the SVM model shows higher improvement than the Naive Bayes model. Consistent with results in Sections VI-B and VI-C, the standardization method shows stronger performance than the expectation-maximization based method when all classes are present in the target domain. At low SNR values the domain adaptation methods provide negligible improvement.

As shown in Fig. 9 the two domain adaptation methods provide a performance increase of 5–15 percentiles when the number of classes in the target domain is limited to two. The expectation-maximization method outperforms the standardization method only for high SNR in the target domain.

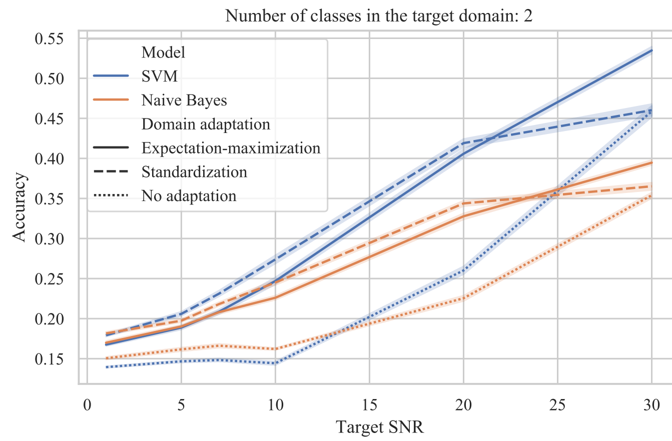


Fig. 9: The accuracy of the three classification algorithms when evaluated on data with artificially reduced SNR. The models are fit to data with no added noise. Two classes are present in the target domain.

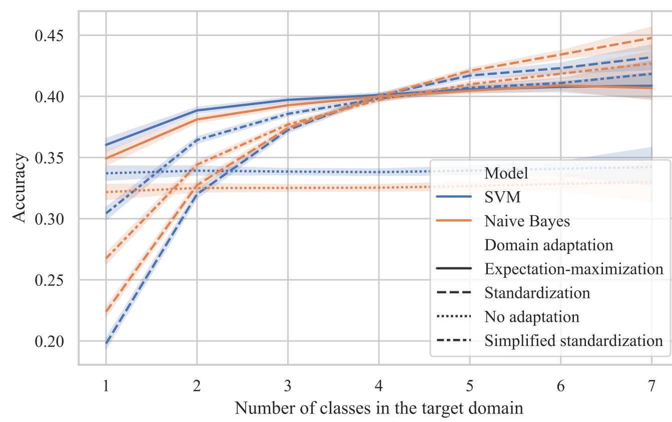


Fig. 10: The performance of the domain adaptation methods when individuals in the target domain are not found in the source domain with the inclusion of simplified standardization which only estimates a feature translation.

E. Direct comparison of expectation-maximization and standardization

The expectation-maximization and standardization approaches to domain adaptation utilize two different parametrizations of the function $f(\cdot)$ as defined in (6), (2). To make an analysis of the estimation procedures directly comparable, a simplified standardization method is evaluated using the parametrization $f(x) = x - \theta_\mu$. The LIPO validation scenario in Section VI-B is reproduced with the addition of the simplified standardization method. The results presented in Fig. 10 show that the simplified standardization method performs 2–4 percentiles better than the standardization method when less than four classes are present in the target domain and performs 1–3 percentiles worse when more than five of the seven classes are present in the target domain.

VII. Conclusion

A novel method for domain adaptation is proposed based on expectation-maximization and its performance is compared to the standardization method. Both methods are applied to data acquired by a multi-static radar network where variations in the radar nodes' observation conditions cause differences in the micro-Doppler features of the same target while observed by several radar nodes. It has been shown that domain adaptation methods that utilize unlabeled data to adjust the feature characterization of micro-Doppler spectrograms to the observation conditions can be used to increase the classification accuracy of the radar nodes as well as their fused predictions. These results are a step in the direction of creating portable radar target classification models in micro-Doppler spectra which is of particular importance when a target is observed by a multi-static radar network given the diverse set of observation conditions.

The results presented in Section VI show that the proposed method for domain adaptation and the standardization method are able to increase the classification performance for a variety of causes of domain shift. When sampling bias introduces differences in the number of classes present in the target and source domains, the proposed method based on expectation-maximization is able to provide an increase in performance, while the method based on standardization is detrimental to the classification performance.

The proposed expectation-maximization based approach is able to robustly estimate the domain shift by utilizing a statistical model of the distribution of the classes. In contrast, the standardization method characterizes the source and target distribution only by their mean and variance and is therefore less robust when sampling bias is present.

The standardization method utilizes a more flexible parameterization to describe the mapping from the source to the target domain which translates and scales the features while the mapping used in the expectation-maximization based approach only translates the data. When all classes are present in the target domain, the standardization method outperforms the expectation-maximization method. These results indicate that scaling the features to match the domains' second moments is beneficial if the variance can be correctly estimated.

Matching the second moment of the source and target distributions has a detrimental influence on the classification performance when sampling bias causes a limited number of classes to be present in the target domain. If only a few classes are present in the domain, the variance is underestimated as its dependency on the separation between the classes is not fully accounted for.

The expectation-maximization method does not correctly estimate the feature translation when all classes are present in the target domain. This may occur if the expectation-maximization algorithm has converged to a local minima or a saddle point. This effect could be mitigated by initializing many expectation-maximization optimization procedures with different initial states as performed in [5].

The expectation-maximization based approach utilizes a parameterization of the mapping between the two domains which is a simple translation in each feature. If the parameterization is extended to include changes to the covariance matrix, there exists no closed-form solution of the parameters in the maximization step. However, it may be possible to find an approximate solution. It may also be beneficial to find parameterizations that are driven by expert knowledge in how the features change between domains, such as attenuation of Doppler velocities for non-radial motion. These are topics for future work.

Acknowledgment

Research was sponsored by the Army Research Laboratory and was accomplished under Grant Number W911NF-2020-185. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

The researchers also acknowledge and are grateful to the UCL Radar Group for providing the experimental data.

References

[1] W. M. Kouw and M. Loog, “A review of domain adaptation without target labels,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 3, pp. 766–785, 2021.

[2] Y. Li, N. Wang, J. Shi, J. Liu, and X. Hou, “Revisiting batch normalization for practical domain adaptation,” *Pattern Recognition*, vol. 80, 03 2016.

[3] S. Seo, Y. Suh, D. Kim, J. Han, and B. Han, “Learning to optimize domain specific normalization for domain generalization,” *Computer Vision – ECCV 2020*, vol. 12367, pp. 68–83, 2020.

[4] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in *International conference on machine learning*. PMLR, 2015, pp. 448–456.

[5] T. van Laarhoven and E. Marchiori, “Domain adaptation with randomized expectation maximization,” *arXiv preprint arXiv:1803.07634*, 2018.

[6] Y. Lin and J. Le Kernec, “Performance analysis of classification algorithms for activity recognition using micro-doppler feature,” in *2017 13th International Conference on Computational Intelligence and Security (CIS)*. IEEE, 2017, pp. 480–483.

[7] W. Van Eeden, J. P. De Villiers, R. Berndt, W. A. Nel, and E. Blasch, “Micro-doppler radar classification of humans and animals in an operational environment,” *Expert Systems with Applications*, vol. 102, pp. 1–11, 2018.

[8] M. S. Seyfioğlu, A. M. Özbayoğlu, and S. Z. Gürbüz, “Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1709–1723, 2018.

[9] M. S. Seyfioğlu and S. Z. Gürbüz, “Deep neural network initialization methods for micro-doppler classification with low training sample support,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 12, pp. 2462–2466, 2017.

[10] S. Z. Gurbuz and M. G. Amin, “Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring,” *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 16–28, 2019.

[11] F. Fioranelli, J. Patel, S. Z. Gürbüz, M. Ritchie, and H. Griffiths, “Multistatic human micro-doppler classification with degraded/jammed radar data,” in *2019 IEEE Radar Conference (RadarConf)*. IEEE, 2019, pp. 1–6.

[12] F. Fioranelli, M. Ritchie, S. Z. Gürbüz, and H. Griffiths, "Feature diversity for optimized human micro-doppler classification using multistatic radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 2, pp. 640–654, 2017.

[13] V. Papanastasiou, R. Trommel, R. Harmanny, and A. Yarovoy, "Deep learning-based identification of human gait by radar micro-doppler measurements," in *2020 17th European Radar Conference (EuRAD)*. IEEE, 2021, pp. 49–52.

[14] C. J. Wu, "On the convergence properties of the em algorithm," *The Annals of statistics*, pp. 95–103, 1983.

[15] W. A. M. Al-Ashwal, "Measurement and modelling of bistatic sea clutter," Ph.D. dissertation, UCL (University College London), 2011.

[16] J. J. De Wit, R. I. Harmanny, and P. Molchanov, "Radar micro-doppler feature extraction using the singular value decomposition," in *2014 International Radar Conference*. IEEE, 2014, pp. 1–6.