This article has been accepted for publication in IEEE Geoscience and Remote Sensing Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LGRS.2023.3235243

Open-Set Patient Activity Recognition with Radar Sensors and Deep Learning

Geethika Bhavanasi, Lorin Werthen-Brabants, Tom Dhaene, *Senior Member, IEEE*, and Ivo Couckuyt

***Abstract*—Open-set recognition has achieved significant importance in recent years. For a robust recognition system, we need to identify the right class from a myriad of knowns and unknowns. In this work, we build and compare open-set recognition systems for patient activity recognition using compact radar sensors in a hospital setting. Radar sensors are an important part of a privacy-preserving monitoring system. Specifically, the proposed approach is based on a deep discriminative representation network trained using the large margin cosine loss and triplet loss. A probability of inclusion model in the embedding space based on the Weibull distribution is able to separate knowns from unknowns. This overall approach limits the risk of open space and enables us to easily identify any unknown activities. Our experiments show that the proposed approach is significantly better for open-set human activity recognition with radar when compared to the state-of-the-art open-set approaches.**

***Index Terms*—Open-set recognition, Human activity recognition, Extreme value theory, Large margin cosine loss, Triplet loss, Radar sensors, Deep learning.**

I. INTRODUCTION

OPEN-Set Recognition (OSR) is an area in Machine Learning (ML) concerning the identification of known classes, while simultaneously being able to correctly identify unknown classes. In Human Activity Recognition (HAR) for radar, traditional classifiers focus on a specific set of activities, which in a real-life scenario will likely not be the only activities one performs. To this end, in this work, we compare different OSR techniques and apply them to HAR for radar.

Radar sensors are widely studied for the recognition of human activities in many intelligent systems using ML, or more specifically Deep Learning (DL) techniques [1], [2]. In healthcare systems, patient activities are monitored using intrusive devices such as cameras, wearable or medical sensors, etc. However, radar sensors on the other hand are privacy-preserving and non-intrusive. Recently most advanced DL techniques and data sets became prominent in remote sensing [3], [4].

Much research has taken place to solve the OSR problem in recent years [5], [6]. Many of the approaches use traditional ML methods, but do not perform well for more complex data types (such as images, audio, graphs, etc.), where Deep Neural Networks (DNN) are better suited and have shown promising performance for OSR. OpenMax [7] was the first DL model where the unknown sample class probabilities are calculated using redistributed SoftMax probability distributions. There are also Deep Metric-Learning (DML) approaches that focus on separating the inter-class feature distributions [8] by learning a similarity metric.

Generative Adversarial Networks (GAN) are popular generative models to learn the decision margin between known and unknown samples [9], [10]. Following this idea, Generative OpenMax (G-OpenMax) was introduced in [11], where a conditional GAN is used to generate unknown samples and the classifier locates the decision margin based on the learning of these unknown samples. Most of the above research on OSR is on image-based activities and data. To the best of our knowledge, very little research is done on OSR using radar sensors with Micro-Doppler (MD) features. In [10], a GAN model is used to explore HAR based on MD signatures. A discriminative model-based OSR with Extreme Value Theory (EVT) is used to identify people based on gait characteristics of radar MD signatures in [12].

The objective of this research is to build an OSR approach to Patient Activity Recognition (PAR) using radar Range-Doppler (RD) maps, that can identify the normal daily patient activities (known) as well as recognize the abnormal activities (unknown). This use case is of interest to the medical community where the medical practitioners can be notified of abnormal activities performed by the patient in daily life, without invading their privacy. This in turn helps medical practitioners to take immediate medical actions to improve patient comfort. The main contributions of this work are as follows:

1) *Discriminative Representation Learning (DRL):* Two main loss functions Triplet Loss (TL) and Large Margin Cosine Loss (LMCL) are studied which aim to simultaneously reduce the intra-class variance while enlarging the inter-class margin.
2) *Deep Discriminative Representation Network (DDRN):* We propose a robust OSR approach for HAR with radars that combines DDRN with a probabilistic discriminant model based on the statistical EVT approach.
3) *Results & Analysis:* We investigate the performance of the proposed model by comparing the results with the DML (distance-based) approach on our DDRNs, OpenMax [7] and OpenGAN [9] approaches.

The rest of the paper is organized as follows. In Section II, the proposed open-set patient activity recognition approach is detailed. The experimental setup and the results are presented in Section III. Finally, in Section IV, we conclude with potential directions for future research.

This work was supported by Ghent University-imec and the Flemish Government under the ”Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen” programme. (Corresponding author: Geethika Bhavanasi.)

The authors are with Ghent University-imec, Department of Information Technology (INTEC), Internet Technology and Data Science Lab (IDLab) research group, iGent, Technologiepark-Zwijnaarde 126, B-9052 Ghent, Belgium (e-mail: geethika.bhavanasi@ugent.be).

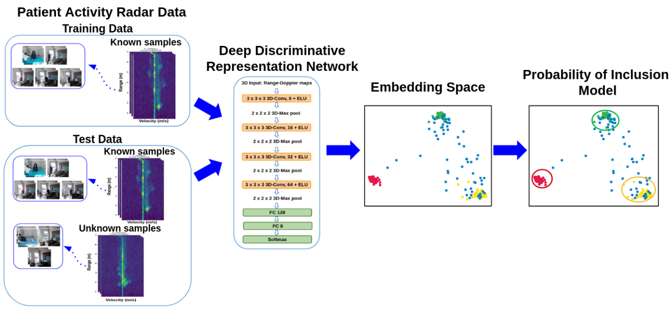


Figure 1: Schematic overview of the proposed open-set patient activity recognition framework with radar data Range-Doppler (RD) maps of known and unknown activities.

II. Methodology

In classification tasks, the model always tries to assign a sample to a known class. However, in real-life applications, we can never capture all possible outcomes and classes in the training data. Hence, classification models for OSR systems need also to be able to separate known and unknown activities. Fig. 1 & Algorithm 1 show the schematic overview of the proposed approach[1].

Algorithm 1 Overview: Open-Set Patient Activity Recognition.

Step 1. PARrad data set: Known & unknown activities (Section III-A).
Step 2. Deep Discriminative Representation Network (DDRN): Convolutional Neural Network (CNN) based on the TL or LMCL loss functions. (Section II-A & III-B).
Step 3. Probability of inclusion model: Statistical EVT in the embedding space (Section II-B Algorithms 2 & 3).
Step 4. Identify knowns and unknowns based on class probabilities from Step 3 (Section II-C).

A. Discriminative Representation Learning (DRL)

In representation learning techniques, the learned features are discriminative, that is they not only emphasize the correct classification but are also robust to deal with samples with large intra-class variance. To that end, a CNN is used as a DDRN with loss functions that are originally introduced for face recognition problems [8], [13]:

1) Large Margin Cosine Loss (LMCL): The LMCL which is also referred to as CosFace [13], represents the traditional softmax loss as a cosine loss by performing L2 normalization simultaneously on both the features and the weight vectors. The feature vectors are also scaled by a constant factor (s) and a cosine margin (m) is added to the cosine of the angle. The LMCL is formulated as:

$$L_{\text{LMCL}} = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s(\cos(\theta_{y_i,i})-m)}}{e^{s(\cos(\theta_{y_i,i})-m)}+\sum_{j\neq y_i}e^{s\cos(\theta_{j,i})}}. \tag{1}$$

[1]The code is made available at https://gitlab.ilabt.imec.be/sumolab/RadarDataOSPAR.git.

where $cos(\theta_{j,i}) = \boldsymbol{W}_j{}^T\boldsymbol{x}_i$, $\boldsymbol{W} = \frac{\boldsymbol{W}^*}{||\boldsymbol{W}^*||}$, $\boldsymbol{x} = \frac{\boldsymbol{x}^*}{||\boldsymbol{x}^*||}$, N is the number of training samples and $\boldsymbol{x}_i \in \mathbb{R}^d$ is the i-th sample feature vector corresponding to the ground-truth class of label $y_i \in 1, 2, ...C$, $\boldsymbol{W}_j \in \mathbb{R}^d$ is the weight vector of the j-th class in the fully connected layer, and θ_j is the angle between $\boldsymbol{W}_j$ and $\boldsymbol{x}_i$.

2) Triplet Loss (TL): TL encompasses the mining of effective triplets that can train the network to group similar activities and separate dissimilar activities based on the Euclidean distance in a vector space [8]. The mining of triplets is done with an online batch hard approach by selecting the five triplets per sample in the training set for which the Euclidean distance between the anchor and the positive sample is largest (hard positives) and the distance between the anchor and the negative sample is smallest (hard negatives). In this approach, several challenging triplets are selected among the batch. Through the online batch hard approach, the representation network learns more efficiently as the selected triplets are the hardest within a small subset of the data. The TL is defined as follows, with squared Euclidean distance $d(\boldsymbol{x}, \boldsymbol{y}) = \|f(\boldsymbol{x}) - f(\boldsymbol{y})\|_2^2$:

$$L_{TL} = \sum_{i}^{N}\Big[d(\boldsymbol{x}_i^a, \boldsymbol{x}_i^p) - d(\boldsymbol{x}_i^a, \boldsymbol{x}_i^n) + m\Big]_+, \quad \forall(f(\boldsymbol{x}_i^a), f(\boldsymbol{x}_i^p), f(\boldsymbol{x}_i^n)) \in \mathcal{T}. \tag{2}$$

where $f(\boldsymbol{x}) \in \mathbb{R}^d$ represents an embedding, $[\cdot]_+ = max(0, \cdot)$, m is the margin, $\mathcal{T}$ is the set of all possible triplets extracted from the training data set with cardinality N. A triplet consists of an anchor sample $\boldsymbol{x}^a$, a positive sample $\boldsymbol{x}^p$, and a negative sample $\boldsymbol{x}^n$.

Note that both loss functions LMCL and TL aim at samples of the same class to be closer to their class centers and features from different classes more separated. This makes it possible to tightly bound the support regions to the known classes to limit the open space risk.

B. Statistical Extreme Value Theory (EVT)

Statistical EVT [14] provides the limiting distributions (i.e., Extreme Value Distributions; EVDs) that occur for the maximum or the minimum of randomly distributed data. Any continuous independent and identically distributed random variables require three EVT models, based on whether the maximum or the minimum is needed, and if the data is upper- or lower-bounded [15]. A generated EVD will follow one of the three EVT model types: type I: Gumbel distribution, type II: Frechet distribution, and type III: Weibull distribution. These three types of EVDs can be unified into a three-parameter distribution known as the Generalized Extreme Value (GEV) distribution given as:

$$GEV(\boldsymbol{x}) = \begin{cases} \frac{1}{\beta}e^{-v^{-1/\alpha}}v^{-(1/\alpha+1)}, & \text{if } \alpha \neq 0 \\ \frac{1}{\beta}e^{-(t+e^{-t})}, & \text{if } \alpha = 0. \end{cases} \tag{3}$$

where $t = \frac{\boldsymbol{x}-\gamma}{\beta}$, $v = (1 + \alpha\frac{\boldsymbol{x}-\gamma}{\beta})$, and α, β and γ are the shape, scale and location parameters, respectively. In our work, we consider the Weibull distribution as it is upper bounded. After training the network, in the learned embedding space, we construct a probability of inclusion model based on the Weibull distribution which is suitable to determine the support regions for the known classes in order to limit the open space risk [12], [16]. The probability of inclusion model is constructed based on the Weibull distribution as given in Algorithm 2 & 3. In Algorithm 2, the Eq. 4 is a monotonically increasing function of the distance $d(\boldsymbol{x}_i^j, \boldsymbol{\mu}_j)$ and hence Eq. 5 is monotonically decreasing. This means that when there is an increase in the distance between the sample and its class center, the sample belonging to that class decreases. From this, we can see that the monotonically decreasing function is able to limit the open space risk.

Algorithm 2 Weibull distribution & Probability of inclusion model.

Input: Embeddings of the test samples, i.e., feature vector $\boldsymbol{x}_i^j \in \mathbb{R}^d$ with $j \in 1, 2, ..., C+1$, C the number of known classes, and $C+1$ the unknown class.

Output: $\Psi_j(d(\boldsymbol{x}_i^j, \boldsymbol{\mu}_j); \alpha_j, \beta_j, \gamma_j)$, i.e., the probabilities of the samples for each known class.

1. Perform clustering on $\boldsymbol{x}_i^j$ to get the class centroids $\boldsymbol{\mu}_i$.
2. Calculate distance $d(\boldsymbol{x}_i^j, \boldsymbol{\mu}_j)$.
3. Fit the Weibull distribution with parameters values from Algorithm 3 for each known class as:

$$F_j(d(\boldsymbol{x}_i^j, \boldsymbol{\mu}_j); \alpha_j, \beta_j, \gamma_j) = \begin{cases} \exp\Big\{ - \Big[- \Big(\frac{d(\boldsymbol{x}_i^j, \boldsymbol{\mu}_j) - \gamma_j}{\beta_j} \Big) \Big]^{\alpha_j} \Big\}, & \text{if } d(\boldsymbol{x}_i^j, \boldsymbol{\mu}_j) < \gamma_j \\ 1, & \text{if } d(\boldsymbol{x}_i^j, \boldsymbol{\mu}_j) \geq \gamma_j, \end{cases} \tag{4}$$

where $\alpha_j > 0$, $\beta_j > 0$ and γ_j are the shape, scale and location, respectively.

4. Finally, we calculate the class probabilities from the probability of inclusion model for each known class as:

$$\Psi_j(d(\boldsymbol{x}_i^j, \boldsymbol{\mu}_j); \alpha_j, \beta_j, \gamma_j) = 1 - F_j(d(\boldsymbol{x}_i^j, \boldsymbol{\mu}_j); \alpha_j, \beta_j, \gamma_j). \tag{5}$$

Algorithm 3 Estimate the Weibull distribution parameters.

Input: Embeddings of the correctly classified training samples, i.e., feature vector $\boldsymbol{x}_i^j \in \mathbb{R}^d$ with $j \in 1, 2, ...C$, and C the number of known classes.

Output: Weibull distribution parameters $(\alpha_j, \beta_j, \gamma_j)$ for each class in j.

1. Perform clustering and calculate distance $d(\boldsymbol{x}_i^j, \boldsymbol{\mu}_j)$ similar to step 1 and step 2 in Algorithm 2.
2. Consider 5-10% (based on number of samples) largest distance values as vector $\boldsymbol{x}_l$ from $\boldsymbol{x}_i$ for each class in j.
3. Perform Maximum Likelihood Estimate (MLE): We use the libMR library [17], applying MLE on $\boldsymbol{x}_l$ for each class to estimate shape α_j, scale β_j and location γ_j.

C. Decision making: known or unknown sample

The probabilistic discriminant model aims at predicting both the known activities (labels $1...C$) and the unknown activities which are assigned label $C+1$. A threshold δ is used to decide whether a sample belongs to one of the known activities or none of them. During the testing phase, the model predicts a sample $\hat{\boldsymbol{x}}$ with a label $\hat{y}$ as follows:

1) *Extreme Value Theory (EVT)-based approach:*

$$\hat{y} = \begin{cases} \underset{j \in \{1,..,C\}}{\text{argmax}} \ P(y=j \mid \hat{\boldsymbol{x}}), & \text{if } P(y=j \mid \hat{\boldsymbol{x}}) \geq \delta \\ \underset{j \in \{C+1\}}{\text{'Unknown'}}, & \text{Otherwise.} \end{cases} \tag{6}$$

where $P(y=j \mid \hat{\boldsymbol{x}}) = \Psi_j(d(\boldsymbol{x}_i^j, \boldsymbol{\mu}_j); \alpha_j, \beta_j, \gamma_j)$ is the class probability over the j-th class. A sample is assessed based on its class probabilities from the probability of inclusion model. If any of the class probabilities of the sample is larger than threshold δ then the sample is assigned to a class with the largest probability. In the other case, the sample is rejected as unknown.

2) *Distance-based approach:*

$$\hat{y} = \begin{cases} \underset{j \in \{1,..,C\}}{\text{argmin}} \ P(y=j \mid \hat{\boldsymbol{x}}), & \text{if } P(y=j \mid \hat{\boldsymbol{x}}) \leq \delta \\ \underset{j \in \{C+1\}}{\text{'Unknown'}}, & \text{Otherwise.} \end{cases} \tag{7}$$

where $P(y=j \mid \hat{\boldsymbol{x}}) = d(\boldsymbol{x}_i^j, \boldsymbol{\mu}_j)$ is the distances between the sample and the known class centers $\boldsymbol{\mu}_j$ over the j-th class. For a sample to belong to a known activity class, the minimum of the distances should be lower than the threshold δ. Then the sample is assigned to a class with the minimum distance. In the other case, the sample is rejected as unknown.

III. EXPERIMENTAL SETUP AND RESULTS

A. Data set

The Patient Activity Recognition with radar sensors (PARrad[2]) data set from [2] is considered. The PARrad data set contains data of 14 different activities of adult and elderly people. The activities are performed in two different environments (simulating) hospital rooms: Homelab and Hospital, where the patients are monitored by recognizing different activities in a non-intrusive way using two different FMCW radar sensors. The data set contains 21569 activities totaling 22 hours of data. This work focuses on the Hospital environment data contained in PARrad, with 13359 activities spread over 5 combined classes: walk, sit down, stand up, fall, and bed activities (roll in bed, lie in bed, sit in bed, get in bed and get out bed). The RD maps (with 98 Doppler bins and 63 range bins) radar features are used for the proposed model.

B. Deep Discriminative Representation Network (DDRN)

We adapted the CNN-RD network from [2] which uses RD maps as features. The RD maps consist of 40 frames, representing 3.7 seconds of radar data. The network is modified by appending two Fully Connected (FC) layers i.e., one FC layer with the number of neurons to match the appropriate embedding size and which is L2 normalized, and the other FC layer with the softmax non-linearity layer. Based on empirical results, we found the optimal LMCL hyperparameters scale

[2]The PARrad data set is available at https://sumo.intec.ugent.be/parrad.

factor (s) and margin (m) as 32 and 0.2, respectively. Similarly, for the TL, the hyperparameter margin (m) is set to 1.0. The neural network is trained for 2000 epochs using the Adam optimizer with a learning rate of 0.0001, making sure the classes are balanced for each mini-batch. We observed that the optimal embedding size is 8 as a further increase in size does not have a positive effect on the effectiveness of the network. Therefore, the final fully-connected layer consists of eight neurons. To this end, we have two DDRNs: one with TL and the other with LMCL. It takes around 8 hours on average to converge (using a GeForce GTX 1080 Ti graphic card).

For the given data set, the training data consists of only known activities (70% of the samples), while the validation and test data consist of both the known (each 15% of the samples) and unknown activities (the activities that are not included in the training data). The amount of samples for the validation and test data set varies according to the unknown samples as we performed leave-one-activity-out cross-validation. In the testing phase, the test samples are passed into the well-trained DDRN to extract 8-dimensional feature vectors. In the embedding space, using the probability of inclusion model we estimate the class probabilities (as given in Algorithms 2 & 3). Finally, based on the threshold we determine the label of the corresponding class i.e., one of the known class or unknown. It takes on average of 10 minutes to compute the probability of inclusion model. For the distance-based approach, it takes 5 minutes as it doesn't have to perform Weibull distribution.

C. Hyperparameter selection

In discriminative models, almost all OSR approaches use thresholds for making a decision either to categorize or reject the sample as unknown. Thus, the decision threshold δ is a critical parameter that plays a key role in the performance of the recognition in the OSR by having boundaries between the knowns and open space. In our work, the optimal threshold is calculated using the class probabilities of the samples which are obtained from the statistical EVT approach. In particular, the threshold is based on a 5-fold cross-validation on the validation data. The optimal threshold is set at the point of Equal Error Rate (EER) where the error rate of unknown activities incorrectly classified as knowns (i.e., False Acceptance Rate; FAR) and the error rate of known activities incorrectly classified as unknown (i.e., False Rejection Rate; FRR) are equal.

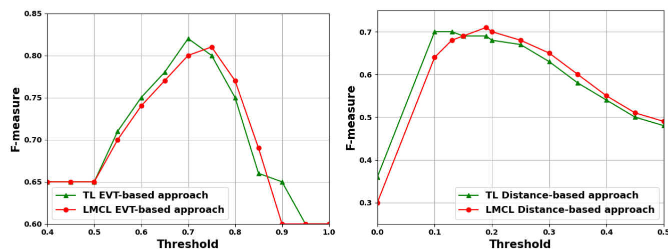
D. Results & Analysis:

The leave-one-activity-out cross-validation on the activities is considered, i.e., for each fold the network is trained with 3 of the known activities and leaves one activity as unknown as given in Table I. From Fig. 2, we can see the average optimal thresholds for both TL and LMCL for the EVT-based approach are 0.7 and 0.75, respectively. Similarly, for the distance-based approach, the average optimal thresholds for both TL and LMCL are 0.1 and 0.19, respectively. From empirical results, the optimal threshold for OpenMax is 0.4. The EVT-based approach (with both TL and LMCL) performance gain is almost 10% improvement f-measure over distance-based approach with optimal threshold, 24% over the OpenMax, and 20% over OpenGAN approaches. Fig. 3 shows the embedding space for one of the folds for both LMCL- and TL-trained networks. Macro-averaging open-set f-measure [18] is used to evaluate the performance. To quantify the difficulty of an OSR problem, the openness (O) metric [18] is used which is defined as: $O = 1 - \sqrt{\frac{C_{\text{train}}}{C_{\text{test}}}}$ where C_{train} and C_{test} are the number of training and test classes, respectively. The problem is completely closed when the openness is equal to 0. As we increase the openness, the performance of the approach decreases as shown in Figure 4, which is logical as a higher number of unknowns leads to a more challenging problem. Still, the performance of the model is respectable for 1 to 6 unknown activities, considering the fall & bed-related activities (roll in bed, lie in bed, sit in bed, get in bed and get out bed) from the PARrad data set. The closed-set f-measure is 98% (for TL and LMCL), a 16%-21% increase when compared to the more challenging open-set problem (see Figure 4). The following observations can be made:

1) *EVT- vs other approaches:* Overall, for EVT and distance-based approaches the discriminative networks trained with both LMCL and TL loss show similar performance in the identification of unknown samples (as shown in Table I). The proposed EVT-based approach outperforms the distance-based, OpenMax and OpenGAN.
2) *Activities:* There is some confusion between the activities: fall & sit and stand & walk. This can be explained as these activities have a similar higher velocity which makes it difficult for the model to identify them as separate activities.
3) *Openness:* As the number of unknown activities increases it is difficult for the model to correctly identify and reject them as unknowns, see Figure 4. In this case, the discriminative network trained with TL loss performs well compared to the network with LMCL loss for EVT-based approach.

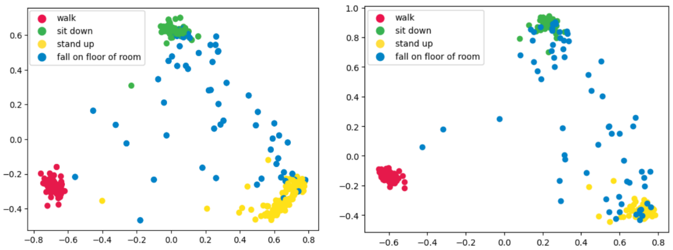
Table I: Leave-one-activity-out Cross-Validation (CV) using the macro-average f-measure.

Leave-one-activity-out Unknown activity	EVT-based LMCL	EVT-based TL	Distance-based LMCL	Distance-based TL
Fall	81%	82%	71%	70%
Walk	75%	74%	68%	68%
Sit down	77%	75%	68%	66%
Stand up	75%	75%	68%	68%



(a) EVT-based model vs Threshold (b) Distance-based model vs Threshold

Figure 2: The performance of the EVT-based and distance-based models with the change of threshold δ.



(a) DDRN trained with LMCL (b) DDRN trained with TL

Figure 3: Visualization of the embedding space learned features by the Deep Discriminative Representation Network (DDRN) trained with (a) the Large Margin Cosine Loss (LMCL) and (b) the Triplet Loss (TL) on the test set with known classes: walk, sit down & stand up. The unknown class is fall. The embeddings are projected to two dimensions using Principal Component Analysis (PCA).

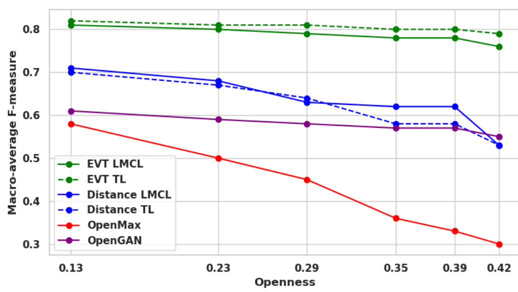


Figure 4: The macro-average f-measure of models as openness increases (known activities: walk, sit down & stand up).

IV. Conclusion

We propose a novel framework for patient activity recognition using radar data in the open-set setting. A statistical EVT approach and a deep discriminative representation network with two different loss functions are compared. The proposed approach outperforms a standard distance-based metric approach for radar sensor data. In our future work, we will focus on further investigating open-set recognition tasks with radar data in-depth with approaches such as GANs, G-OpenMax, etc. We will also investigate loss functions to further enhance the intraclass compactness and interclass discrimination of the learned deep features.

References

[1] W.-Y. Kim and D.-H. Seo, "Radar-based human activity recognition combining range–time–doppler maps and range-distributed-convolutional neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–11, 2022.

[2] G. Bhavanasi, L. Werthen-Brabants, T. Dhaene, and I. Couckuyt, "Patient activity recognition using radar sensors and machine learning," *Neural Computing and Applications*, p. 16, May 3, 2022.

[3] Q. He, X. Sun, Z. Yan, B. Li, and K. Fu, "Multi-object tracking in satellite videos with graph-based multitask modeling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022.

[4] X. Sun, P. Wang, Z. Yan, *et al.*, "Fair1m: A benchmark dataset for fine-grained object recognition in high-resolution remote sensing imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 184, pp. 116–130, 2022.

[5] X. Ma, K. Ji, L. Zhang, S. Feng, B. Xiong, and G. Kuang, "An open set recognition method for sar targets based on multitask learning," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[6] C. Geng, S.-J. Huang, and S. Chen, "Recent advances in open set recognition: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, pp. 3614–3631, 2021.

[7] A. Bendale and T. E. Boult, "Towards open set deep networks," presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, Jun. 1, 2016, pp. 1563–1572.

[8] E. Hoffer and N. Ailon, "Deep metric learning using triplet network," in *Similarity-Based Pattern Recognition*, A. Feragen, M. Pelillo, and M. Loog, Eds., Cham: Springer International Publishing, 2015, pp. 84–92.

[9] S. Kong and D. Ramanan, "Opengan: Open-set recognition via open data generation," *International Conference on Computer Vision (ICCV)*, pp. 793–802, 2021.

[10] Y. Yang, C. Hou, Y. Lang, D. Guan, D. Huang, and J. Xu, "Open-set human activity recognition based on micro-doppler signatures," *Pattern Recognition*, vol. 85, pp. 60–69, Jan. 1, 2019.

[11] Z. Ge, S. Demyanov, Z. Chen, and R. Garnavi, "Generative OpenMax for multi-class open set classification," arXiv:1707.07418, Jul. 24, 2017.

[12] Z. Ni and B. Huang, "Open-set human identification based on gait radar micro-doppler signatures," *IEEE Sensors Journal*, vol. 21, pp. 8226–8233, Mar. 2021.

[13] H. Wang, Y. Wang, Z. Zhou, *et al.*, "CosFace: Large margin cosine loss for deep face recognition," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5265–5274.

[14] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*. London, U.K, 2001.

[15] E. J. Gumbel, *Statistical Theory of Extreme Values and Some Practical Applications.* National Bureau of Standards, Applied Mathematics Div., 1954.

[16] L. P. Jain, W. J. Scheirer, and T. E. Boult, "Multi-class open set recognition using probability of inclusion," in *Computer Vision – ECCV 2014*, vol. 8691, Cham: Springer International Publishing, 2014, pp. 393–409.

[17] W. J. Scheirer, A. Rocha, R. J. Micheals, and T. E. Boult, "Meta-recognition: The theory and practice of recognition score analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 8, pp. 1689–1695, 2011.

[18] P. R. Mendes Júnior, R. M. de Souza, R. d. O. Werneck, *et al.*, "Nearest neighbors distance ratio open-set classifier," *Machine Learning*, pp. 359–386, 2017.