RADAR Emitter Classification with Optimal Transport Distances

Manon MOTTIER
Université Paris-Saclay, CNRS, CentraleSupélec
Laboratoire des signaux et systèmes
91190, Gif-sur-Yvette, France
manon.mottier@centralesupelec.fr

Gilles CHARDON
Université Paris-Saclay, CNRS, CentraleSupélec
Laboratoire des signaux et systèmes
91190, Gif-sur-Yvette, France
gilles.chardon@centralesupelec.fr

Frédéric PASCAL
Université Paris-Saclay, CNRS, CentraleSupélec
Laboratoire des signaux et systèmes
91190, Gif-sur-Yvette, France
frederic.pascal@centralesupelec.fr

***Abstract*—Identifying unknown RADAR emitters from received pulses is an important problem in electronic intelligence. It is a difficult problem, as agile RADAR emitters can have complex characteristics, and measurements are corrupted by various noises (non-Gaussian noise, missing pulses, etc.). In this paper, we introduce a new classification method based on optimal transport distances between collected RADAR pulses and a priori known emitter classes. Compared to previously proposed methods, this method does not require a training step, it can deal with a large number of classes, and it is easily interpretable. The method is tested on data obtained by a realistic RADAR scene simulator.**

***Index Terms*—Optimal Transport, Classification, RADAR recognition, Electronic warfare**

I. Introduction

RADAR emitter classification aims to identify the RADAR emitters present in a measured signal to gain electronic intelligence in a given environment. Recent advances in RADAR technologies make this task more difficult, as RADAR emitters exhibit more complex behaviors: agility in frequency and pulse repetition intervals, complex scanning patterns, etc. In this work, we assume that RADAR pulses have been deinterleaved, that is, the analyzed pulses are assumed to be emitted from a single emitter. Several methods exist to solve this problem, based on the analysis of the pulse repetition intervals [1], [2], deep learning [3], or hierarchical clustering with optimal transport distances [4].

For simple RADAR signals, time-frequency analysis is sufficient to identify the present emitter with great certainty [5]. Several supervised classification methods have been proposed to deal with more complex cases. Most of the methods are based on Deep Learning models and consider a small number of RADAR classes [6]–[10]. These methods require a large dataset for the training step, and adding a new class requires the classifier to be retrained. Additionally, a challenge in processing RADAR data is acquiring real data, even more labeled real data. Algorithms and methodologies are often developed using small datasets or simulated data. Most of the previous methods are based on simulated data, and their result performance strongly relies on the simulator's accuracy. Technological developments have modified the recognition process. The profiles of transmitters have become more and more complex, enhancing the existing panorama of transmitters of new types with more varied patterns. New methods have been developed in order to compare the group characteristics to a known database, also allowing to detect new transmitters [11], [12].

The article introduces a classification method based on an optimal transport distance between collected RADAR pulses and RADAR emitter models from a reference database containing more than 60 classes. The method does not require a training step, and it also allows the easy creation of new RADAR classes. The paper is structured as follows: Section 2 presents the data and their origin. Section 3 contains basic notions on optimal transport distances and introduces the proposed classification method. Simulation results are given in Section 4, while Section 5 concludes the paper and highlights interesting perspectives.

II. Data Description

Data are collected by a receiver, listening on a large bandwidth. Pulses are then segmented, analyzed, and described by four features:

- Frequency, F (f_n)
- Pulse width, PW (w_n)
- Level, G (g_n)
- Time of Arrival, TOA (t_n)

Additional features, such as frequency and amplitude modulations, are not considered in this work.

Fig. 1 shows a simulated signal gathering the pulses of five different emitters. In the top plot, the pulse level is plotted as a function of time, showing that several RADAR emitters can be active simultaneously. In the bottom plot, pulses are plotted based on their estimated frequency and pulse width. One can clearly see that a given emitter may emit on different frequencies (e.g., six frequencies for emitter (1). Moreover, errors in the pulse width estimation are important. Estimated pulse widths are truncated for low-energy pulses, mainly when the receiver is in a side-lobe of the emitter. Real data are challenging to acquire, so that the method will be validated on such simulated data. Finally, note that the proposed classification method requires separated pulses. Several deinterleaving methods have been proposed in the

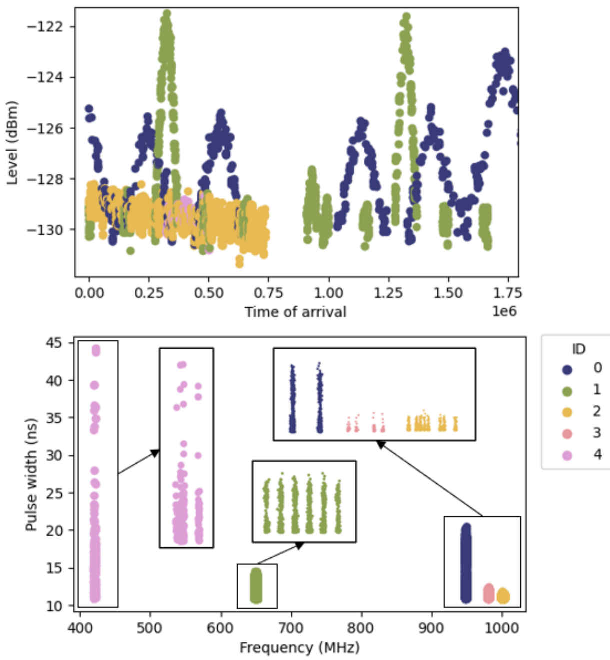

Fig. 1. Set of pulses contained five transmitters. Each color represents an emitter.

literature (see for instance [1]–[4], [7]). Here, we assume that the RADAR pulses have been correctly separated.

III. Identification Algorithm

The proposed methodology, outlined in Algorithm 1, is based on the development of a distance between a set of received RADAR pulses and a description of the characteristics of a RADAR emitter from a reference database. Classification is made by identifying the closest (in terms of distribution distance) RADAR emitters to the received data. In the case of emitters with simple characteristics (e.g., single frequency), simple distances such as a Euclidean distance between the average of the features of the received pulses and the features of a reference emitter can be used. However, such distances cannot deal with more complex, agile emitters, for which their averages cannot simply describe the distribution of the features.

Here, received data and classes are represented as probability distributions. There are several ways to define a distance between probability distributions. For instance, the Kullback-Leibler divergence or the total variation distance are frequently used. Nevertheless, these distances or divergences cannot be used in our case, as the distributions representing the data and the classes are discrete and, in general, will have disjoint supports. We will show in the sequel that the proposed optimal transport distances are well-suited to the problem at hand, as they can deal with distributions representing agile emitters and are robust to noise.

A. Optimal transport

Optimal transport makes it possible to find a mapping between an original mass distribution and a different target distribution [13], [14]. In this work, we focus on the part of this theory dealing with discrete probability distributions, useful for describing received data and different classes of typical RADARs.

In particular, we consider two discrete probability distributions $\nu = \sum_{n=1}^{N} a_n \delta_{x_n}$ and $\mu = \sum_{m=1}^{M} b_m \delta_{y_m}$, with $\mathbf{a} = (a_1, \dots a_N)^T \in \mathbf{R}_+^N$, $\sum_{n=1}^{N} a_n = 1$, and $\mathbf{b} = (b_1, \dots b_M)^T \in \mathbf{R}_+^M$, $\sum_{m=1}^{M} b_m = 1$. A transport plan $\mathbf{P}$ between ν and μ is defined by its coefficients P_{nm}, representing the amount of mass taken from x_n to y_m. With $c(\cdot,\cdot)$ a cost function, and $C_{nm} = c(x_n, y_m)$ the cost of transporting a unit of mass from x_n to y_n, the total cost $C(\mathbf{P})$ of a transport plan is

$$C(\mathbf{P}) = \sum_{n=1}^{N} \sum_{m=1}^{M} C_{nm} P_{nm} \tag{1}$$

The consistency of the transport plan $\mathbf{P}$ with ν and μ is guaranteed by $\mathbf{P1}_M = \mathbf{a}, \mathbf{P}^T\mathbf{1}_N = \mathbf{b}$.

The optimal transport plan $\mathbf{P}^\star$ is defined the minimization of the transport cost in (1) under the following constraints:

$$\mathbf{P}^\star = \underset{\mathbf{P} \in \mathbf{R}_+^{N \times M}}{\text{argmin}}\ C(\mathbf{P}) \text{ subject to } \mathbf{P1}_M = \mathbf{a}, \mathbf{P}^T\mathbf{1}_N = \mathbf{b} \tag{2}$$

In the following, this distance is applied to the similarity measure between emitters, and to the classification of a set of RADARs pulses. The cost function $c(x, y) = \|x - y\|_2$ will be used. The optimal transport distance between ν and μ is then defined by $d(\nu, \mu) = C(\mathbf{P}^\star)$.

Algorithm 1 Classification with Optimal Transport distance.

- Construction of a density probability distribution from the data: ν
- Construction of discrete measures from a RADAR emitter database: μ_j
- Computation of the transport cost between the probability distribution of the data and of each emitter model: $d_j = d(\nu, \mu_j)$
- Selection of the emitter model with the lowest transport cost between its probability distribution and the distribution of the data: $j^\star$

B. Measuring the similarity between emitters

The first application of optimal transport is to measure the distance between classes of typical emitters. Our initial RADAR database includes more than 60 different emitters, and some classes have very similar characteristics while others are easily distinguishable. Typical RADARs characteristics were randomly simulated to highlight the previous comment. As detailed above, we have chosen to work only with specific features: frequency f_n and pulse width w_n, as they are very significant and reliable. From this database, we build a measure μ_j describing each typical RADAR belonging to the reference database by:

$$\mu_j = \sum_{n=1}^{N} \alpha_n \delta_{f_n, pw_n} \tag{3}$$

with N the number of frequencies and pulse widths on which the RADAR transmits, α the proportion of frequency and pulse width and δ, the Dirac mass (with $\sum \alpha_n = 1$).

Fig. 2 shows an example of a representation of typically simulated emitters. Each stem represents a frequency used by the emitter for an emitter, with height representing the proportion of appearance of this frequency. Some emitters emit on very different frequency bands like transmitters B and C. They can easily be distinguished thanks to their frequency. On the other hand, other transmitters like A and B possess very close characteristics. Finally, we may have mixed scenarios: Transmitter E transmits on several frequencies, one common with transmitter D. In conclusion, RADARs with various observable characteristics will be easier to identify.

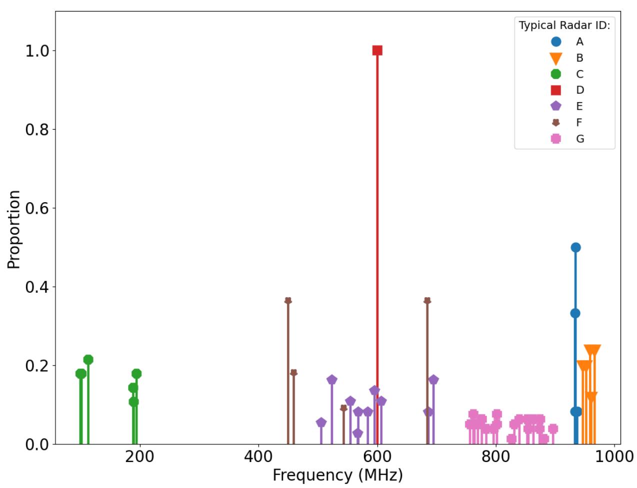

Fig. 2. Representation of some typical simulated RADARs. The graph shows the RADARs in 1 dimension in the frequency plane. Each color represents a typical simulated RADAR.

Fig. 3 shows the same emitters in the F-PW plane. One can notice that the transmitters A and B are clearly separated in the F-PW plane. This shows that incorporating additional characteristics improves the separability of the transmitters.

Fig. 4 shows the cost matrix between the previous simulated RADAR classes using frequencies and pulse widths. As previously explained, some classes are very close because they have similar characteristics. As an example, classes A and B are separated by a small distance, while classes B and C are larger. Therefore, it is expected that class B will be more frequently confused with class A than with class C.

Now, let us consider the optimal transport plan between emitters. Fig. 5 shows the optimal transport plan from emitter (from class) E toward emitters D, F, and G. The cost of moving

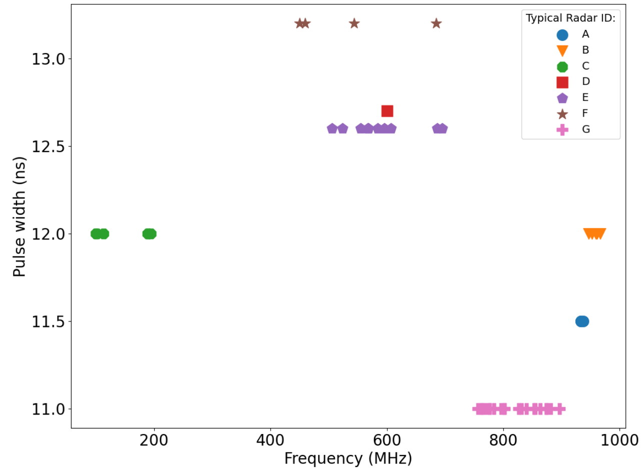

Fig. 3. Representation of some typical simulated RADARs. The graph represents the RADARs in the F-PW plane. Each color represents a typical simulated RADAR.

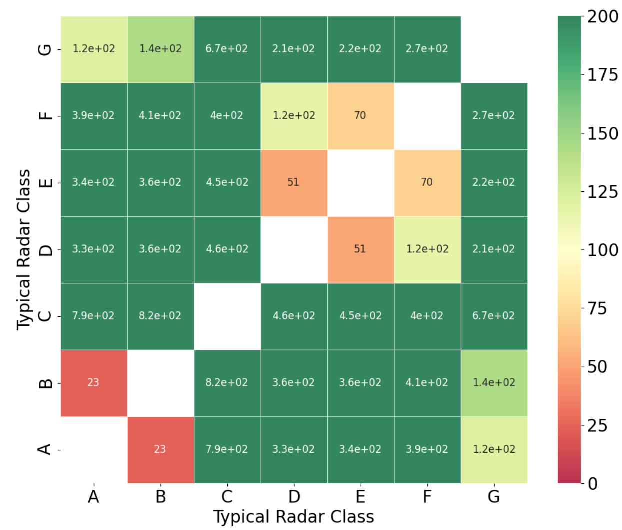

Fig. 4. Cost matrix between typical RADARs class builds from Frequency and Pulse Width. Green indicates high transport cost, which means that the class is very different, and red indicates low transport cost means that the class has similar characteristics.

the points from E towards those of D is low because the transmitter D has a very close pulse width and transmits on a single frequency; all points moved to a very close location. The displacement cost between transmitter E and G is four times higher because they do not transmit on the same frequency bands; they have very different pulse widths, so the points must make more significant displacements. Note that the optimal transport distance also considers the proportion of pulses at a given frequency and pulse width.

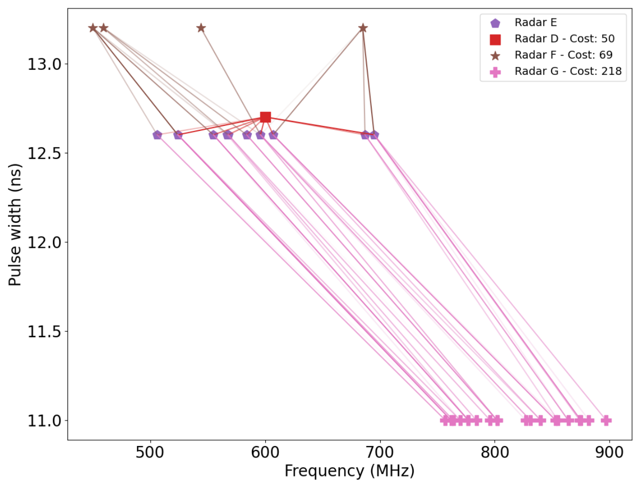

Fig. 5. RADAR E points displacements towards the points of RADARs D, F and G with the optimal transport in two dimensions.

C. Classification Decision

In order to compute the distance between a set of pulses and a RADAR class, received pulses are also modeled as a probability distribution, as follows:

$$\nu = \frac{1}{M} \sum_{m=1}^{M} \delta_{f_m, pw_m} \tag{4}$$

with M the number of pulses in the set. A set of pulses class is then assigned by identifying the closest RADAR class in the optimal transport distance sense:

$$j^\star = \underset{\mathbf{j} \in (1,\dots,J)}{\operatorname{argmin}} \; d\left(\mu_j, \nu\right). \tag{5}$$

with J the number of emitters classes. The algorithm identifies the transmitter $j^\star$ as the true class. To reduce the computation cost of computing the distances, the received pulses are binned in frequency and pulse width intervals. With sufficiently small bin sizes, perturbation of the distance is small enough so that the order of the distances between the data and the classes is conserved.

IV. Results

A. Classification example

Fig. 6 shows the result of our classification methodology applied to emitter 1 from Fig. 1. The plot on the left overlays the pulses and the three closest emitter classes. The blue dots fit very well with those on the data. The classifier correctly identifies the emitter present in the data. The plot on the right shows us the transport plan between the distribution of the data and each outputs [15]. Output 2 represents a single-frequency transmitter, so the data points are all sent to the same location. Output 3 represents a RADAR that transmits on six different frequency bands, so the data points are sent on the different Output 1 respecting the proportions of Output 1; this is why pulses around a given frequency are not all sent to the same point.

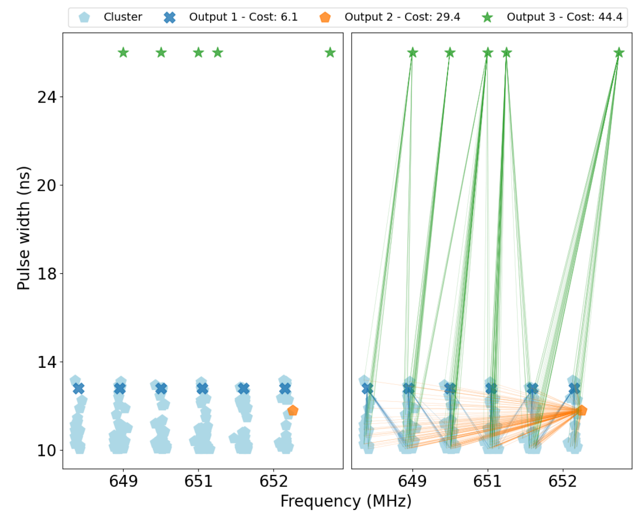

Fig. 6. Classification output for set of pulses 1 in two dimensions. The plot on the left overlap the set of pulses 1 with the first three classes identified by the algorithm. The plot on the right represents the transport plan between the data and those of the algorithm's outputs.

B. Performance evaluation

In order to evaluate the performance of the proposed methodology, we simulated a dataset containing the pulses of 3608 different emitters that must classified. The noise level on the different features varies form in each simulation. Tab. I shows global results of the classification methodology using one feature (frequency) and two features (frequency and pulse width). Using only the frequency, 94% of the transmitters are correctly identified. Thanks to the Optimal Transport cost, one can rank the classes as follows: the classes are sorted according to the associated cost. Thus, in 98.8% of the cases, the correct class is ranked among the two first one. This approach allows to introduce some flexibility in the classification decision. Indeed, when the cost of the (two/three) first classes are very similar, the data could be classified in all these classes, with almost the same confidence. Then, adding pulse width as a feature increases the identification rate to nearly 96%. However, the proportion of ranks higher or equal to three slightly increases to 0.4%. This is explained by the spread of the estimated pulse widths, as we can see in Fig 7, which is not taken into account. The figure shows an example of misclassification. The one-dimensional classification based on the frequency allows to correctly identify the true class, as the pulse frequencies are close to the unique frequency of the class. The plot on the right shows the transport plan between the different Outputs using frequency and pulse width. Here, the bias of the estimated pulse width, and the relative weights given to frequency and pulse width in the computation of the distance, make the signal closer to output 1 and output 2 than to the true class. Indeed, their pulse width is closer to the underestimated width of the received pulses.

TABLE I
EMPIRICAL RESULTS OF THE CLASSIFICATION MADE ON 3608 SET OF PULSES.

Correct classification position	Frequency	Frequency and Pulse Width
First	94.1%	95.8%
Second	5.7%	3.8%
Third	0.1%	0.1%
More	0.1%	0.3%

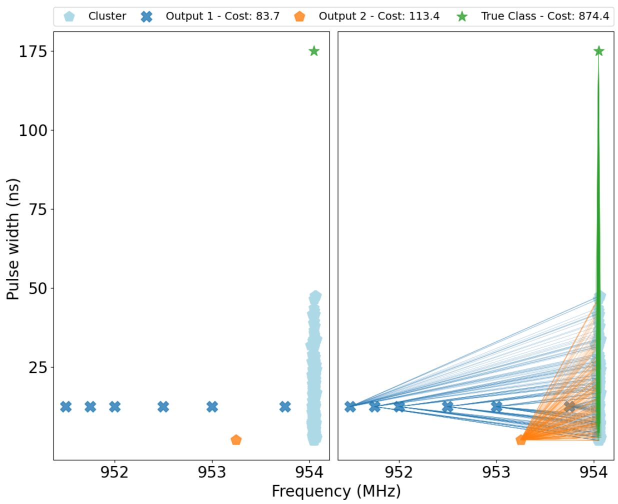

Fig. 7. Example of wrong classification made in two dimensions.

V. Conclusion and Perspectives

A. Conclusion

In this paper, one has developed a methodology to identify the emitters present in a RADAR signal. This work assumes that the different possible emitters have been previously deinterleaved. Then, considering an output that corresponds to (only) one emitter, the proposed method allowed to classify the transmitter thanks to optimal transport theory in two dimensions (frequency and pulse width). The results obtained on the simulated data are very encouraging and allow us to identify the class of transmitters confidently. Moreover, the methodology can handle a large number of classes to identify.

B. Perspectives

We assume that the sets of pulses coming from a previous deinterlacing step contain the pulses of a single transmitter. However, these sets may contain mixed pulses coming from several transmitters. We are currently developing a new methodology to deinterlace a RADAR signal better to deal with the imperfect pre-processing steps. Moreover, to enhance the classification results, several perspectives are of interest: First, one could add a third dimension in the optimal transport theory to better discriminate RADARs. As previously explained, some RADARs can have very similar characteristics in frequency and pulse width, and it is necessary to add other characteristics to differentiate them, *e.g.*, the pulse repetition period (PRI), which is the difference of time of arrival between successive pulses ($\delta_n = t_n - t_{n-1}$). However, note that the PRI cannot be directly integrated into the optimal transport as it requires pre-processing to be exploited. Indeed, missing pulses imply significant distortions of the probability density of PRIs. Then, as mentioned in this work, pulse width errors reduce the method performance. The robustness of this method with respect to such errors should be improved. Finally, to propose a complete classification method, this method should be capable of detecting emitters that not present in the database.

Acknowledgment

ATOS supported this work by providing the data and its RADAR expertise.

References

[1] H. Mardia, "New techniques for the deinterleaving of repetitive sequences," in *IEE Proceedings F (Radar and Signal Processing)*, vol. 136, no. 4. IET, 1989, pp. 149–154.
[2] D. Milojević and B. Popović, "Improved algorithm for the deinterleaving of radar pulses," in *IEE Proceedings F (Radar and Signal Processing)*, vol. 139, no. 1. IET, 1992, pp. 98–104.
[3] X. Li, Z. Liu, and Z. Huang, "Deinterleaving of pulse streams with denoising autoencoders," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 6, pp. 4767–4778, 2020.
[4] M. Mottier, G. Chardon, and F. Pascal, "Deinterleaving and clustering unknown radar pulses," in *2021 IEEE Radar Conference (RadarConf21)*. IEEE, 2021, pp. 1–6.
[5] A. Kawalec and R. Owczarek, "Radar emitter recognition using intrapulse data," in *15th International Conference on Microwaves, Radar and Wireless Communications (IEEE Cat. No. 04EX824)*, vol. 2. IEEE, 2004, pp. 435–438.
[6] J. Lunden and V. Koivunen, "Automatic radar waveform recognition," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 124–136, 2007.
[7] Z.-M. Liu and S. Y. Philip, "Classification, denoising, and deinterleaving of pulse streams with recurrent neural networks," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 4, pp. 1624–1639, 2018.
[8] Z. Geng, H. Yan, J. Zhang, and D. Zhu, "Deep-learning for radar: A survey," *IEEE Access*, vol. 9, pp. 141 800–141 818, 2021.
[9] L. Ding, S. Wang, F. Wang, and W. Zhang, "Specific emitter identification via convolutional neural networks," *IEEE Communications Letters*, vol. 22, no. 12, pp. 2591–2594, 2018.
[10] M. A. Nuhoglu, Y. K. Alp, and F. C. Akyon, "Deep learning for radar signal detection in electronic warfare systems," in *2020 IEEE Radar Conference (RadarConf20)*. IEEE, 2020, pp. 1–6.
[11] J. Liu, J. P. Lee, L. Li, Z.-Q. Luo, and K. M. Wong, "Online clustering algorithms for radar emitter classification," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 8, pp. 1185–1196, 2005.
[12] S. Apfeld and A. Charlish, "Recognition of unknown radar emitters with machine learning," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 6, pp. 4433–4447, 2021.
[13] C. Villani, *Optimal transport: old and new*. Springer, 2009, vol. 338.
[14] N. Bonneel, M. Van De Panne, S. Paris, and W. Heidrich, "Displacement interpolation using lagrangian mass transport," in *Proceedings of the 2011 SIGGRAPH Asia conference*, 2011, pp. 1–12.
[15] R. Flamary, N. Courty, A. Gramfort, M. Z. Alaya, A. Boisbunon, S. Chambon, L. Chapel, A. Corenflos, K. Fatras, N. Fournier, L. Gautheron, N. T. Gayraud, H. Janati, A. Rakotomamonjy, I. Redko, A. Rolet, A. Schutz, V. Seguy, D. J. Sutherland, R. Tavenard, A. Tong, and T. Vayer, "Pot: Python optimal transport," *Journal of Machine Learning Research*, vol. 22, no. 78, pp. 1–8, 2021. [Online]. Available: http://jmlr.org/papers/v22/20-451.html