Insights into Heterogenous Sensor Fusion and Passive RF Feature Extraction for xAI

Asad Vakil
Department of Electrical and Computer Engineering
Oakland University
Rochester, MI
avakil@oakland.edu

Robert Ewing
Air Force Research Laboratory
Dayton, OH
robert.ewing.2@us.af.mil

Erik Blasch
Air Force Research Laboratory
Arlington, VA
erik.blasch.1@us.af.mil

Jia Li
Oakland University
Rochester, MI 48309
li4@oakland.edu

NAECON 2024 - IEEE National Aerospace and Electronics Conference | 979-8-3503-6762-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/NAECON61878.2024.10670686

***Abstract*— Deep learning-based models have made significant contributions to many fields in recent years but lack robust explainability in their decision making and interpretability in their inference processes. While deep learning is capable of processing information that traditional methods might struggle with, such as nonlinear data, the lack of results explainability can be detrimental to the credibility in such models. Hence, there is a need to enhance the level of explainability, which can come from multimodal analysis. In this study, we implement a Convolutional Neural Network (CNN) for the fusion of Passive RF (P-RF) and Electro-Optical (EO) data to gain insights into how P-RF data can be utilized for target detection. The P-RF data first undergoes feature extraction via Short-Time Fourier Transform (STFT), Continuous Wavelet Transform (CWT), Wigner-Villle Distribution (WVD) and Constant-Q Gabor Transform (CQT). In previous experiments using the ESCAPE dataset, the multimodal design training was incentivized to utilize the P-RF data with I/Q histogram as the feature by purposefully restricting the available EO data. Prior experimentation with both Greedy Algorithms and Saliency Maps indicated that the fusion of P-RF and EO data still heavily focuses on the EO data, if possible, only relying on the P-RF data if it was necessary to detect the target. While P-RF has seen more use in vehicle detection for both autonomous driving and drone applications in recent years, its impact on sensor fusion based decision making is still under investigation. By expanding on the available P-RF data, this paper compares different features of P-RF data and their impact in the fusion using diverse counterfactual explanations (DiCE), as well as potentially increasing the reliability of the P-RF data for target detection.**

Keywords—Explainable AI, Sensor Fusion, EO, Passive RF

I. Introduction

The use of artificial intelligence (AI) has been extensively applied to many fields in recent years, including multimodal fusion [1]. While traditional data fusion methods are *transparent* and *understandable*, the innate blackbox nature of deep learning (DL) models is a major issue. While such DL models can process nonlinear data and excel where traditional algorithms cannot - by approximating the desired results, it could be difficult to consistently explain the processing results. Ignoring factors like the larger amount of training data needed for multimodal DL algorithms, the ability to understand what the model is processing is integral to the development, troubleshooting, and design.

Implementing DL training without having any means of understanding the model's decision-making process can lead to major issues [2]. Such problems are numerous, including inheriting biases from systems that are meant to inherently be impartial [3]. Others inherent design issues include exploiting features that were not meant to be included in the model training for the desired application [4]. These kinds of issues are easy to miss given the vastness of the training data needed for most multimodal DL algorithms, but have the potential to have devastating effects on applications such as autonomous driving, financial decision, medical healthcare [5], etc. Understanding that AI is not a comprehensive method that solves all problems is the first step to determine how to use the technology in a safe and productive manner. Using DL can provide an efficient or effective manner to enhance situation awareness [6].

Choosing to not use any form of DL algorithms because of the blackbox drawback is obviously not a reason to avoid utilizing the power of DL methods; especially in the case for data which is inherently difficult for human users to understand, and for higher dimensional fusion of multiple data sources. Such is the case for utilizing passive radio frequency data (P-RF) in data fusion. P-RF data provides various benefits for target detection, primarily with regards to the non-invasive nature of the modality. A P-RF based approach for multimodal fusion can augment visual data, cannot be visually obscured by camera angle, and is harder to detect for the purposes of electronic countermeasures.

Methods of using RF data for vehicle detection include leveraging RF signal strength [7], active RF signals [8], passive RFID [9] [10], and joint-sparse data level fusion [11]. However, using the raw I/Q data as P-RF for vehicle detection has not been extensively researched. In this paper, we present our work at using P-RF data in sensor fusion to achieve vehicle target detection. The contributions of this paper are as follows:

1. Processing P-RF data via Short Time Fourier Transform (STFT), Continuous Wavelet Transform (CWT), Wigner-Villle Distribution (WVD) and Constant-Q Gabor Transform (CQT). While these methods are considered conventional, they have not been used for *passive* RF nor for vehicle detection with that modality.

2. Utilizing Explainable AI to analyze the local and global impact on the fusion model for each of the five P-RF

features that are tested and compared, using a counterfactual based causal inferencing scheme, Diverse Counterfactual Explanations (DiCE).

II. Literature Review

2.1 EO/RF Sensor Fusion

The application for the heterogenous sensor fusion research presented in this paper, via deep learning (DL), is to accurately detect vehicle targets using EO and P-RF sensor inputs while also providing some level of explainability. The fusion of EO and RF data for the purposes of tracking has been used extensively [12] in similar applications, but the use of passive RF data [13] which is more challenging to implement without conventional methods such as Doppler radar. While most research has traditionally been focused on active RF sensors, the use of P-RF data comes with logistical and economic benefits, as it requires less power, is considerably harder to detect than active RF methods, and requires less hardware.

RF modalities excel in providing range, angular, and spectral resolution of collected information [14]. The benefits of combining RF data with higher spatial resolution of EO based sensors are extremely desirable for detection, classification, and tracking. There are several RF-based approaches that are used in applications such as tracking [15], localization [16], and detection [17]. Most EO modalities are intuitively much easier to implement, owing to a human's reliance on sight, such as full motion video (FMV) and infrared (IR) [18]. RF-based sensors can also provide repetitive coverage over a wide geographical area, and in doing so, can determine the precise distance and velocity of a target. While both EO and RF have limitations that from environmental operating conditions, the lack of overlap these modalities have in terms of these limitations makes the fusion of the two highly desirable.

While there has been P-RF research for the purposes of implementing vehicle target tracking, these methods typically involve RFID [19] or rely on sensor fusion of a blackbox method to utilize the data. Additional research uses P-RF for animal [20] and human occupancy detection [21,22] as well as wi-fi sensing for human activity analysis [23,24]. To our knowledge, there is a few reported techniques that rely on the use of the raw P-RF In-phase Quadrature (I/Q) data, owing to a) the inherent difficulty of utilizing the modality and b) other more active methods such as doppler already existing. In previous work we have utilized P-RF histograms of the I/Q data [25,26] in order to achieve data fusion, this paper explores other methods of expanding the use of the raw P-RF data.

2.2 Explainable AI (XAI)

With the use of artificial intelligence (AI) even reaching recreational use such as ChatGPT and Open AI, it is important to consider the need for explainability when considering uncertainty [27]. While proponents of AI might point out that requiring explainability will slow down AI processing, it is important when considering the moral and design considerations. Ignoring issues like whether or not an ante-hoc explainable AI (XAI) is truly faithful to the original model, the incomplete nature of saliency map visualizations, or the tradeoffs in performance, without XAI its impossible to understand what a blackbox model's decision-making process is. With XAI, there is both a design challenge and also morally concern, as it makes diagnosing errors extremely difficult. Problems in the training data, such as the very human choices [3,28] that the model's trained on can cause the human biases to be ingrained into the model's decisions. Other errors in training data might also cause the model to pick up on either wildly incorrect or "cheat" [4] its way to getting the correct results, which will not help it in its real-world applications. As such, especially as applications like autonomous driving, medical, and financial fields already use AI, the need to have some level of accountability [29] is of paramount importance.

As explainable AI (XAI) is still an emerging concept [30], there has yet to be any uniform adoption of interpretability assessment criteria for XAI. There are several criteria used to describe different approaches for providing explainability and the levels of explainability. These include (1) post and ante-hoc methods, which describe *where and when* the method itself is implemented in the model, (2) local or global, which describe *what* level of interpretability is being provided, or (3) model agnostic or model specific, which describe *how* versatile the method is.

Methods of XAI also vary based on the type of information being processed and the model used, such as *visualizations*. By implementing the post-hoc method, it's possible to correlate pixels on an input image to neuron activations, providing a level of understanding as to what parts of the image the model determined to be relevant features. While not as empirical as expressing impact via weights, visualizations provide a more user-oriented explanation that a human expert can understand. The visualization explanations are not necessarily restricted to users with knowledge of the field, as the input image and what regions are activated might provide intuitive changes that can be noticed. That being said, it is difficult to quantify the impact of the information that is provided by the visualizations.

Some more commonly used examples of Explainable AI methods include LIME [31] (Local Interpretable Model-Agnostic Explanations) and SHAP [32] (SHapley Additive exPlanations). These methods are model-agnostic explanations, and provide partial dependence plots, which are global and model-agnostic in nature. Others include the use of Greedy Algorithms to recreate the blackbox model's decision-making process, such as ExplainX.ai, an advanced version of ProtoDash [33].

For the purposes of this paper, we are presenting our work with Diverse Counterfactual Explanations (DiCE) [34] in order to explore the impact of different types of P-RF data in heterogenous sensor fusion for vehicle target differenetiation. DiCE [35] was first published in 2019, and provides explainability by using counterfactuals, hypothetical examples which show how the model obtains a different prediction. The

Grant FA9550-21-1-0224.

python-based DiCE framework generates and evaluates a diverse set of counterfactual explanations that are based on determinantal point processes. Each of those counterfactuals are then evaluated using metrics that enable the comparison of counterfactual-based methods to other local methods. While attribution-based methods like SHAP and LIME can provide a number of insights which are useful, as complementary post-hoc methods of XAI, given the larger pool of data and the limitations of each method, DiCE was chosen. This paper explores contrasting the different P-RF and EO modalities with respect to *each other*, such that the choice of choosing a counterfactual based causal inferencing using the DiCE method.

III. Experiment Design

3.1 The Escape Dataset

The *Experiments, Scenarios, Concept of Operations and Prototype Engineering* (ESCAPE) [36] dataset was published in 2019 (Figure 1) by the Air Force Research Laboratory (AFRL) Information Directory for the purposes of enabling multi-modal signature data-fusion research [37,38,39,40]. The ESCAPE dataset combines a variety of different sensors, including EO, P-RF, radar, acoustic, and seismic data in a common scenario for the application of advanced fusion. In each of the scenarios provided, there are one or more vehicles which attempt to "escape" detection. The motivation for the ESCAPE data is to develop any model using the multimodal data with an incentive to use more than just the EO sensors, especially when the EO input data is purposefully limited.

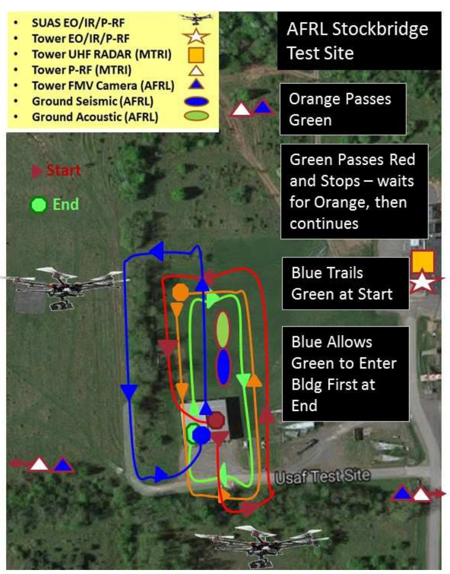

Fig. 1. ESCAPE data collection.

For the purposes of the scenarios and sensor data chosen from the ESCAPE data, the three that were used in this research are designated as Scenarios 1, 2, and 3, which correspond to the ESCAPE dataset's Scenarios 1, 2C, and 2D respectively (Figure 2). The total number of vehicle targets between the three scenarios is ten, and each scenario deals with a different number of targets (2, 3, and 5 targets respectively). The overall purpose of the dataset is that all the targets are designed to "escape" detection, by employing a number of different tactics that incentivize the fusion model to use multiple modalities as data inputs. The evasive scenarios come with the benefit of thereby challenging any model meant to differentiate between the potential vehicles when engaging in tracking. Similar vehicles move in such a manner that even human users might have difficulty in differentiating between them with just the EO visual information. All three scenarios involve multiple vehicles entering and exiting a garage, with multiple vehicles of similar make and build combined with dissimilar vehicles. The movements of these vehicle are also varied in order to present challenging discrimination opportunities that confuse tracking differentiation.

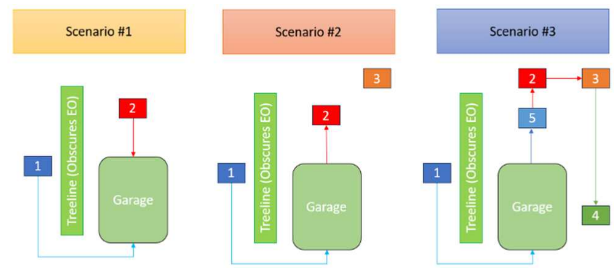

Fig. 2. Comparison of Scenarios #1, #2, and #3.

Scenario #1 (*Similar Vehicle Switch*) has two possible vehicle targets, both of which are of the same build and color as each other. For each of these scenarios, only one source of EO data was used, in order to maximize the need for the model to utilize the P-RF data rather than ignoring the P-RF input. The scenario starts as vehicle #2 travels into the garage in plain view of the EO sensors. As this happens, vehicle #1 travels into the garage from behind the tree line. While doing so, from the EO sensor's point of view, vehicle #1 is "hidden" due to visual obscuration that prevents the model from detecting its movements most of the time. Once vehicle #2 enters the garage, vehicle #1 then exits the garage, and the objective of the first scenario is to successfully determine when the "switch" is made. If the model incorrectly identifies the vehicle exiting as #2, then that means the model has failed and the vehicle has successfully "evaded" detection.

Scenario #2 (*Similar Vehicle Appearance*) is nominally more complicated than scenario 1 by comparison. In this scenario, there are three total vehicles and essentially follow the same pattern as Scenario 1, but only two of the three look visually similar. The difference is that rather than vehicle #3, which is visible, or vehicle #1, which is not possible to obtain at the video angle chosen switching in the garage, is that vehicle #2 that was parked in the garage the entire time. This makes it appear to an outside viewer that vehicle #1 enters and exits when in fact it is hidden inside of the garage. If the model determines vehicle #2 to be vehicle #1, then vehicle #1 will have thereby "*escaped*" detection. The EO input is insufficient on its own to make that determination, as the difference between similar vehicles, incentivizing the use of P-RF data

Scenario #3 (*Similar Vehicle Overtake*) is the most complicated of the three and chosen due to the complexity of the five vehicle targets, all of which are traveling at different speeds and with different makes. Four of these vehicles arrive out of the front of the garage, while the fifth vehicle arriving from out of view, thereby making the tracking at the end of the video input, linearly speaking, extremely difficult to conduct with only the EO input for that time frame. The variable speeds displayed by the five vehicle targets also presents an additional dimension of complexity with respect to tracking as the vehicles that are similar in design will overtake the other at different points within the scenario, making tracking a challenging process for Scenario 3 in general.

3.2 Data Preprocessing

The two sources of data are preprocessed in a number of different ways, with EO being preprocessed with Dense Optical Flow (DOF), and P-RF data being preprocessed into histograms and with a Fast Fourier Transform (FFT) over the original I/Q data. While different sources of data enhancement have been tested in previous research, the usage of different interpretations of the same data is ideal for the purposes of gaining insights with the decision level fusion.

To extract more useful features from the P-RF data, Wigner-Ville Distribution (WVD), Continuous Wavelet Transform (CWT), Constant-Q Gabor Transforms (CQB) [36], and Short-Time Fourier Transforms (STFT) are also implemented. WVD has been used for the better half of the century and provides a high-resolution time-frequency representation of the background P-RF data provided. The Constant-Q Transform and Short-time Fourier Transform are related algorithms that transform data series into the frequency domain. The STFT plots changing spectra as a function of time, while CQB plots in logarithmic scale, providing greater accuracy at lower frequencies but less accurate at higher frequencies. Lastly, CWT provides an overcomplete representation of the P-RF background data.

By expanding on the uses of P-RF data, it becomes possible to provide more features for the model to utilize. More feature incentivizes fusion, especially for scenarios in which EO sources alone are incapable of determining and locating the target vehicle. While most vehicles are largely within what the EO modality can see, when tested for vehicles that are behind the tree line or hidden in the garage, the need for the other modalities will ensure the model engages in data fusion.

3.3 Explainable AI and Fusion Model

While in prior research our group has utilized saliency maps and greedy algorithms to estimate each modality's impact on decision making [25, 26], on a local and global scale, for the purposes of this paper the use of *Diverse Counterfactual Explanations* (DiCE) is implemented to evaluate modality impact. The direct performance of the models for vehicle detection will also be compared using F1 score as a reference to each of the five models. The feature weights indicate the relative impact on the model's decision-making process, where values closer to 1 have a greater impact, while values closer to 0 respectively have a lower impact.

For the purposes of training, each CNN model is trained on the same number of epochs, with the only modifications in terms of layers and architecture being to accommodate the different number of modalities being input into the CNN. The objectives for each CNN remain the same, having the same labels for each model. Each model has three hidden layers and is trained on 50 epochs before being evaluated and having the weights of each model saved in order to implement DiCE. The P-RF Histogram data was originally included but later removed owing to its poor performance and lack of useful XAI insights that could be provided by DiCE.

TABLE I. SUMMARY OF MODELS TRAINED

Model	Input Modalities
#1	EO (DOF), All Five P-RF features
#2	EO (DOF), P-RF (CWT), P-RF (WVD)
#3	EO (DOF), P-RF (CQB), P-RF (STFT)
#4	EO (DOF), P-RF (CWT), P-RF (STFT)
#5	EO (DOF), P-RF (CQB), P-RF (WVD)

The five models tested are summarized in Table 1. The first model will serve as a baseline model which has all six modality features including one EO feature and five P-RF features. For the second to fifth models, these models will compare the impact of three different features, one of which will always be the EO (DOF) input, excluding the P-RF (Histogram) features. These five models are discussed further below in the results section.

IV. RESULTS

4.1 Baseline Model (#1)

As a baseline for the *other* four models tested, model #1 uses *all* six sources of information in its data fusion. In order to underline the strongest impact values provided by SHAP across the five models, the **highest** and *lowest* values are highlighted within Tables 2-6. The lowest local and global impact respectively for each different type of modality tested in the five models are marked in *italics*. Similarly, the highest local and global impact for each modality tested (across all five models) is marked in **bold**.

TABLE II. MODEL #1 (F1 SCORE: 0.97)

Modality	Local	Global
EO (DOF)	0.79	0.66
P-RF (CQB)	0.38	0.43
P-RF (CWT)	*0.33*	0.42
P-RF (Histogram)	0.14	0.29
P-RF (STFT)	*0.43*	*0.46*
P-RF (WVD)	*0.49*	0.57

As seen in Table 2, Model #1 performs relatively well, having the second highest F1 score compared to the other five models. The model follows the trend of placing the largest feature impact on the EO (DOF) data. This is unsurprising, as previous research [26] had similar results. The P-RF data impact results for Model #1 provided via SHAP indicate that the WVD and STFT modalities perform the strongest locally for the P-RF modalities in baseline Model #1, with the CQB coming in a close third. The P-RF (Histogram) results are notably lower than even the CWT's local impact. Owing to the poor performance of the P-RF (Histogram) modality, the histogram data was excluded in models #2-5 for fusion and SHAP testing.

4.2 P-RF Modality Comparison (Models #2-5)

As seen in below in Table 3, Model #2 compares the P-RF impacts of CWT and WVD as well as showing the EO (DOF) results. The CWT had a greater increase in local impact for Model #2 compared to the baseline Model #1 while the WVD has a noticeably larger increase in global impact. The EO (DOF) also had an increase in local and global impact, especially with respect to the global impact of Model #2.

TABLE III. MODEL #2 (F1 SCORE: 0.96)

Modality	Local	Global
EO (DOF)	0.81	**0.71**
P-RF (CWT)	0.37	**0.44**
P-RF (WVD)	**0.51**	0.*53*

There is a minor drop in F1 score for Model #2 when compared to baseline Model #1; however, this is to be expected as there is only *half* as much data for the model to work with. While not a direct indicator of a neural network's performance, having sufficient data to train the model on is an important consideration. Model #2 has the distinction of also having the highest global impact of the EO (DOF) input, as well as the highest P-RF (CWT) local impact with respect to the model (barely surpassing Model #4's local impact), and the highest WVD local impact and lowest global impact (as opposed to Model #5's 0.58 and Model #1's 0.57). While the SHAP values indicate a relatively higher reliance on EO data, the results indicate a high reliance on P-RF (CWT) data and stronger reliance on the P-RF (WVD) data.

TABLE IV. MODEL #3 (F1 SCORE: 0.92)

Modality	Local	Global
EO (DOF)	**0.83**	0.68
P-RF (CQB)	*0.37*	*0.39*
P-RF (STFT)	0.45	**0.51**

As seen above in Table 4, Model #3 suffers from the largest drop in performance after reducing the number of sources of fusion data. CQB sees a drop in local and global use, while STFT sees an increase, especially in global impact for Model #3 compared to the baseline Model #1. This model has the distinction of also having the highest local impact of the EO (DOF) and P-RF (STFT), as well as having the lowest local and global impact of CQB. Model #3 is an outlier, as only Model #1 has lower local and global scores for CQB, but these results are to be expected given that model #3 has the lowest F1 score and that the model impacts largely increased when fewer modalities are used.

TABLE V. MODEL #4 (F1 SCORE: 0.95)

Modality	Local	Global
EO (DOF)	0.8	*0.63*
P-RF (CWT)	0.36	0.43
P-RF (STFT)	**0.47**	0.49

As seen above in Table 5, Model #4 has the second lowest F1 score, but noticeably is where STFT reaches its highest local impact for its respective model, compared to the other five. The CWT modality has a notable increase in local impact compared to the baseline Model #1 and also is at its highest globally among the five models.

TABLE VI. MODEL #5 (F1 SCORE: 0.98)

Modality	Local	Global
EO (DOF)	*0.77*	0.64
P-RF (CQB)	**0.41**	**0.47**
P-RF (WVD)	0.5	**0.58**

As seen above in Table 6, Model #5 manages to narrowly outperform the other models in terms of F1 score, including the baseline Model #1. The SHAP values indicate the EO modality is noticeably at its lowest local and second lowest global impact, while both the CQB and WVD have their highest global impact scores between the five models, along with CQB also reaching its highest local impact score.

4.3 P-RF Results Overview

Based on the SHAP results, Model #5 makes the most use globally of the WVD, as well as the most use globally and locally for the CQB. The results of the SHAP values indicate a general trend for Model #1 having the majority of the lowest local and global impact scores, likely owing to having all six features extracted from the two different types of modalities used. Only Model #3 comes close to the baseline Model #1, with the lowest CQB local and global impact, followed by Model #5 with the lowest EO local impact, Model #4 with the lowest EO global impact, followed by Model #2 with the lowest WVD global impact.

By comparing the baseline Model #1's SHAP results and looking at different combinations of EO with different P-RF features, we can better explore the impact of the different methods of extracting P-RF data. In previous research [25, 26], the impact of the P-RF data was always based on which vehicle target was being tracked, where the vehicle target was with respect to the source of EO data, and if similar targets were in the field of view for the EO data. The results of this experiment indicate that by expanding on different P-RF features to exploit, the impact of the P-RF data can be further enhanced, thereby reducing the model's reliance on EO data.

V. Conclusion

This paper presents explainable fusion by comparing different P-RF features for the purposes of vehicle detection using the ESCAPE data. The initial results indicate that CQB and WVD P-RF features provide the best synergy with the EO (DOF) feature, for vehicle detection. While we are uncertain if it is the added benefit of both a frequency domain and time-frequency analysis that the model exploits, future research will need to be conducted to make that determination. This combination of modalities managed to achieve a higher F1 score as well as having the least dependence on the EO (DOF) data on a local level and second lowest globally of the five models based on the SHAP results. In future research, our group aims to increase the number of P-RF processing methods and explore the use of other XAI methods implemented accompanied by visualization for user understanding.

Acknowledgments

This research was supported in part by the AFRL Summer Fellows program and grant FA9550-21-1-0224. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. government.

References

[1] E. Blasch, T. Pham, C. Chong, W. Koch, et al., "Machine Learning/Artificial Intelligence for Sensor Data Fusion–Opportunities and Challenges," *IEEE Aero. and Elec. Sys. Mag.*, 36(7):80-93, 2021.

[2] R. V. Yampolskiy, "Predicting future AI failures from historic examples," *Foresight,* 21(1): 138-152, 2019.

[3] L. Jones, "A Philosophical Analysis of AI and Racism," *Stance: An International Undergraduate Philosophy Journal,* 13: 36-46, 2020.

[4] M. Nauta, R. Walsh, A. Dubowski and C. Seifert, "Uncovering and Correcting Shortcut Learning in Machine Learning Models for Skin Cancer Diagnosis," *Diagnostics,* 12(1): 40, 2021.

[5] T. Shaik, X. Tao, L. Li, H. Xie, J. D. Velásquez, "A survey of multimodal information fusion for smart healthcare: Mapping the journey from data to wisdom,Information Fusion, 102040, 2023.

[6] A. Munir, A. Aved, et a., "Situational Awareness: Techniques, Challenges, and Prospects," *AI*, 3, 55–77, 2022.

[7] N. Kassem, A. E. Kosba and M. Youssef, "RF-Based Vehicle Detection and Speed Estimation," in *2012 IEEE 75th Vehicular Technology Conference (VTC Spring)*, Yokohama, Japan, 2012.

[8] D. Roy, Y. Li, T. Jian, P. Tian, K. Chowdhury and S. Ioannidis, "Multi-Modality Sensing and Data Fusion for Multi-Vehicle Detection," *IEEE Transactions on Multimedia,* pp. 2280-2295, 2023.

[9] C. Li, "Automatic vehicle identification (AVI) system based on RFID," in *International Conference on Anti-Counterfeiting, Security and Identification*, Chengdu, China, 2010.

[10] O. Javaid, F. R. Yu and J. S. Huang, "Autonomous Vehicle Navigation and Communication by Passive Radio Frequency (RFID) Tags," in *DIVANet '21: Proceedings of the 11th ACM Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications*, New York, USA, 2021.

[11] R. Niu, P. Zulch, M. Distasio, E. Blasch, G. Chen, D. Shen, Z. Wang and J. Lu, "Joint sparsity based heterogeneous data-level fusion for multi-target discovery," in *IEEE Aerospace Conference*, Big Sky, MT, 2018.

[12] B. Kahler and E. Blasch, "Sensor Management Fusion Using Operating Conditions," in *IEEE National Aerospace and Electronics Conference*, Dayton, OH, 2008.

[13] A. Vakil, J. Liu, P. Zulch, E. Blasch, R. Ewing and J. Li, "A Survey of Multimodal Sensor Fusion for Passive RF and EO Information Integration," *IEEE Aerospace and Electronic Systems Magazine,* 36(7): 44-61, 2021.

[14] D. Shen, P. Zulch, M. Disasio, E. Blasch, G. Chen, Z. Wang, J. Lu and R. Niu, "Manifold learning algorithms for sensor fusion of image and radio-frequency data," in *IEEE Aerospace Conference*, Big Sky, MT, 2018.

[15] M. A. Shukoor, S. S. Mukeshbhai and S. Dey, "12-Bit Multiresonator Based Chipless RFID System for Low-Cost Item Tracking," in *IEEE International Conference on RFID Technology and Applications (RFID-TA)*, Delhi, India, 2021.

[16] R. Shahbazian and I. Trubitsyna, "Human Sensing by Using Radio Frequency Signals: A Survey on Occupancy and Activity Detection," *IEEE Access,* vol. 11, pp. 40878-40904, 2023.

[17] J. Liu, A. Vakil, R. Ewing, X. Shen and J. Li, "Human Presence Detection via Deep Learning of Passive Radio Frequency Data," in *IEEE National Aerospace and Electronics Conference (NAECON)*, Dayton, OH, 2019.

[18] E. Blasch, Y. Zheng, S. Liu, Z. Liu," Multi-modal Video Fusion for Context-aided Tracking," *Int'l. Conf. on Information Fusion,* 2020.

[19] Y. Zhang, X. Gong, K. Liu and S. Zhang, "Localization and Tracking of an Indoor Autonomous Vehicle Based on the Phase Difference of Passive UHF RFID Signals," *Sensors,* 21(9), 3286, 2021.

[20] W. D. Van Eeden, J.P. de Villiers, R.J. Berndt, W. A. J. Nel, "Micro-Doppler radar classification of humans and animals in an operational environment," *Expert Systems With Applications*, Vol. 102, 15 July 2018.

[21] J. Liu, H. Mu, et al., "Human Occupancy Detection via Passive Cognitive Radio," 20, 4248; *Sensors*, 2020.

[22] L. Yuan, H. Chen, et al., "3-D Indoor Positioning Based on Passive Radio Frequency Signal Strength Distribution," *IEEE Internet of Things Journal*, 10(15): 13933-13944, 2023.

[23] A. K. Koupai, M. J. Bocus, R. Santos-Rodriguez, R. J. Piechocki, R. McConvill, "Self-supervised multimodal fusion transformer for passive activity recognition," IET *Wireless Sens. Syst.* 12:149–160, 2022.

[24] M. Cominelli, F. Gringoli, L. Kaplan, M. Srivastava, F. Cerutti, "Accurate Passive Radar via an Uncertainty-Aware Fusion of Wi-Fi Sensing Data," *Int'l. Conf. on Information Fusion*, 2023.

[25] A. Vakil, E. Blasch, R. Ewing and J. Li, "MVI-DCGAN Insights into Heterogenous EO and Passive RF Fusion," in *9th International Conference on Soft Computing & Machine Intelligence (ISCMI)*, Toronto, Canada, 2022.

[26] A. Vakil, E. Blasch, R. Ewing and J. Li, "Finding Explanations in AI Fusion of Electro-Optical/Passive Radio-Frequency Data," *Sensors,* 23(3): 1489, 2023.

[27] A-L. Jousselme, P. de Villiers, A. de Freitas, et al., "Uncertain about ChatGPT: enabling the uncertainty evaluation of large language models," *Int'l. Conf. on Information Fusion*, 2023.

[28] E. Blasch, J. Sung, T. Nguyen, “Multisource AI Scorecard Table for System Evaluation,” *AAAI FSS-20: Artificial Intelligence in Government and Public Sector*, Washington, DC, USA, 2020.

[29] E. Blasch, P. de Villiers, G. Pavlin, A-L. Jousselme, P. C. G. Costa, K. B. Laskey, J. Ziegler, “Use of the URREF towards Information Fusion Accountability Evaluation,” *Int.l Conf. on Information Fusion*, 2021.

[30] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, et al., “Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,” Information Fusion, Vol. 58, Pages 82-115, 2020.

[31] M. Ribeiro, S. Singh and C. Guestrin, ""Why Should I Trust You?": Explaining the Predictions of Any Classifier.," *Conference of North American Chapter of the Association for Computational Linguistics*, San Diego, CA, 2016.

[32] S. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *NIPS'17: Proc. of the 31st International Conference on Neural Information Processing Systems*, New York, USA, 2017.

[33] K. S. Gurumoorthy, A. Dhurandhar, G. Cecchi and C. Aggarwal, "Efficient Data Representation by Selecting Prototypes with Importance Weights," in *International Conference on Data Mining*, 2019.

[34] S. A. Tariq, T. Zia and M. Ghafoor, "Towards counterfactual and contrastive explainability and transparency of DCNN image classifiers," *Knowledge-Based Systems,* vol. 257, no. 5, p. 109901, 2022.

[35] R. K. Mothilal, A. Sharma and C. Tan, "Explaining Machine Learning Classifiers through Diverse Counterfactual Explanations," in *arXiv*, arXiv, 2019.

[36] P. Zulch, M. Distasio, T. Cushman, B. Wilson, B. Hart and E. Blasch, "ESCAPE Data Collection for Multi-Modal Data Fusion Research," in *IEEE Aerospace Conference*, Big Sky, MT, 2019.

[37] D. Shen, E. Blasch, P. Zulch, M. Distasio, Ruixin Niu, J. Lu, Z. Wang, G. Chen, “A Joint Manifold Leaning-Based Framework for Heterogeneous Upstream Data Fusion,” *Journal of Algorithms and Computational Technology* (JACT), 12(4): 311-332, 2018.

[38] D. Garagić, D. Pelgrift, J. Peskoe, R. D. Hagan, P. Zulch and B. J. Rhodes, "Machine Learning Multi-Modality Fusion Approaches Outperform Single-Modality & Traditional Approaches," *IEEE Aerospace Conference*, Big Sky, MT, USA, 2021.

[39] D. Garagic, R. Ravier, J. Peskoe, T. Galoppo and P. Zulch, "Submodular Optimization via Reinforcement Learning for Active Control of Sensor Networks," *IEEE Aerospace Conference*, Big Sky, MT, USA, 2022.

[40] D. Roy, Y. Li, T. Jian, P. Tian, K. Chowdhury and S. Ioannidis, "Multi-Modality Sensing and Data Fusion for Multi-Vehicle Detection," IEEE Transactions on Multimedia, vol. 25, pp. 2280-2295, 2023.

[41] J. P. Grossman, G. F. Margrave, M. P. Lamoureux R. Aggarwala, “A robust algorithm for constant-Q wavelet estimation using Gabor analysis” SEG Int’l Exposition and 72nd Ann Mtg, Salt Lake, UT, 2002.