Explainable Hybrid Decision Level Fusion for Heterogenous EO and Passive RF Fusion via xLFER

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

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

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

Jia Li
Department of Electrical and Computer Engineering
Oakland University
Rochester, MI
li4@oakland.edu

***Abstract*— This paper presents an explainable late-stage decision fusion model for Electro-Optical (EO) and Passive Radio Frequency (P-RF) target detection via hybrid Explainable AI model. *Explainable* insights that are intuitive and empirical are provided by counterfactual explanations at the early stage of data flow, with a traditional algorithm, decision tree (DT), handling late-stage fusion. Results show that at both the local and global level, the DT explainability of fusion methods provides insights for EO and P-RF fusion methods at each level of fusion The usage of Histograms, Wigner-Ville Distribution (WVD) and Continuous Wavelet Transform (CWT) for the novel use of P-RF data provided insights into the eXplainable Late-stage Fusion of Electro-optical and Radio-Frequency (xLFER) usage of the modality for target detection. While WVD and CWT have been used extensively in RF signal processing, their use in P-RF data for target detection feature extraction has not been documented to our knowledge, nor with a hybrid Explainable AI model.**

Keywords—Explainable Artificial Intelligence, Heterogenous Sensor Fusion, EO, P-RF, Continuous Wavelet Transform, Wigner-Ville Distribution

I. Introduction

As the push for automation continues, the use of artificial intelligence (AI) becomes more prevalent across multiple industries and applications. While traditional signal processing classification methods are *transparent* and *understandable*, current deep learning (DL) classification algorithms lack explainability when handling nonlinear applications. An example is the trend for applications like autonomous driving where tasks include a neural network's ability to closely approximate detection of hazardous objects. As a result, the promised use of DL blackbox algorithms (e.g., internal operations not accessible) has become more prevalent in aspects of everyday life. Such blackbox algorithms; however, come with the disadvantage of not being inherently understandable. Most of these models will require vast amounts of data and find features to exploit based on the available data. This is not a problem in and of itself, but rather becomes one if the features learned are fundamentally incorrect. And being able to make the determination as to if it has been trained correctly is difficult in a blackbox system.

Understanding the explainability of automatic target detection is important for many applications such as medical diagnoses. Irrelevant details to diagnoses, for example, can prop up from unforeseen training data choices, such as correlating the appearance of a ruler with malignant tumors [1]. While it's certainly true that images of biological growths with a ruler next to them are *really likely* to be malignant, that's not likely the *intended features* that the model's creator wanted it to pick up on from the training data. Of course, such problems with blackbox DL methods do not stop there. Other issues from blackbox algorithms can include bias from what should *ideally* be an impartial judge [2], incorrect diagnoses of disease [3], or major general disfunction in its desired application [4]. It is one thing if the application is for something trivial. However, for more important applications, such as automated driving [5], financial, medical, or first-responder use, which carry grave consequences if poorly managed or trained, the need for *explainability* is clear.

Approved for Public Release; Distribution Unlimited: To that end, as deep learning (DL) methods and other blackbox algorithms continue being used, it is important now, more than ever, to provide explainability of such models. Both from a moral perspective and from a design perspective, explainability allows for better understanding of the model's shortcomings and potential flaws. The blackbox problem will always exist on some level, but striving to provide explainability will ensure that some level of understanding can be gained from the model's decision-making process.

This paper presents an explainable late-stage fusion of EO and RF (xLFER) data. The proposed xLFER framework is divided into individual models which independently process data in the earlier stages of fusion and output their decisions for decision-level fusion. While models in the early-stage fusion using blackbox methods to independently process the input data from single modalities, the decision level model uses a traditional model, decision tree, in order to provide a final decision based on the earlier model's outputs. Explanations from the early models include *heatmaps* and *counterfactual explanations* which afford traditional algorithms insights into the model's early-stage decision-making process. Comparison research was conducted with an early-stage fusion model and different traditional models (detailed further in Section III) that implement the decision level fusion.

To the best of our knowledge, no research involving target detection with EO and P-RF data has been conducted using *counterfactual explanations*. In the research presented, our counterfactual explanations are generated using DiCE [6] (Diverse Counterfactual Explanations). Our previous research using explainable AI involved *greedy algorithms* and *saliency maps*, without hybrid models for additional levels of explainability. The rest of the paper is as follows. Section II includes a literature review. Section III overviews the experimental design. Section IV describes the results and Section V provides conclusions drawn from the experiment.

II. Literature Review

A. EO/RF Sensor Fusion

The desired application for the xLFER model is the ability to accurately detect and track vehicle targets using EO and P-RF sensor inputs. The fusion of EO and RF data for the purposes of tracking has been used extensively in many similar applications [7] [8] [9]. However, the use of passive RF data is challenging to implement without conventional methods, such as Doppler radar. While the research focus has traditionally been 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 detail in the form of range, angular, and spectral resolution of collected information [10]. The many benefits of combining RF data with higher spatial resolution of EO based sensors make the fusion of the two sources of data extremely desirable for the applications of detection and tracking [11]. There are many RF-based approaches that are used in a wide variety of applications. These include the fields of tracking [12], human sensing [13], proximity [14], localization [15], and detection [16].

The major issues with using P-RF data come from the fact that there is no signal, generally speaking, that can be filtered out using available methods. The preprocessing of the P-RF data, along with ensuring the synchronization and correct sampling is an important factor for P-RF data exploitation. However, when correctly used in synch with another sensor, the resulting model can yield performance greater than what the sources of data can do so independently of one another.

While fusion with active RF modalities such as Radar with EO modalities is common, it is rare that P-RF and EO modalities are fused together. Passive RF as a modality comes with many benefits, such as reduced energy requirements and costs, and being inherently harder for countermeasures to be used against it for sensing. Moreover, it is even rarer that non-blackbox models are capable of effectively processing the P-RF data, and providing explanations related to its decision-making process are even rarer.

B. Explainable AI

With AI being popular in the year with ChatGPT's release, the importance of having interpretability and understanding of the many deep learning models cannot be understated. As explainable AI (xAI) is still an emerging concept, there has yet to be any uniform adoption of interpretability assessment criteria for xAI. There are many different criteria used to describe different approaches for providing explainability. These include (1) post and ante-hoc methods, which describe *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. Some examples include Bayesian Rule List, which is an ante-hoc explanation, LIME (Local Interpretable Model-Agnostic Explanations) and SHAP (SHapley Additive exPlanations) which are model-agnostic explanations, and partial dependence plots, which are global and model-agnostic in nature.

The type of explanations desired and what types can be applied are dependent on the model and desired application. For image processing, visualizations can be highly desirable. Visualizations are post-hoc xAI methods that include gradients, activation maximizations, deconvolutions, and decompositions. These techniques use tools such as generative models or saliency maps in order to determine activations produced on the last layer of a deep convolutional neural network (DCNN). From these activations, DCNNs can form a pixel-by-pixel mapping that highlights what factors provided the highest level of confidence in that decision, which then can be overlaid on top of the original sample.

The explanation insights are different based on the application; some might be more quantifiable while others might be more intuitive based for data scientists. In the case of visualization methods, they can provide a more user-oriented explanation that human users can understand. In most cases, visualization insight is hard to quantify, though metrics such as Fréchet inception distance can be used to quantify similarities between different images. There are other types of visual insights provided which are more easily shown in empirical form, using algorithms such as LIME, Counterfactual Explanations [17], or SHAP, that compare what data sources are more relied upon on at both local and global level, or to gain understanding from a human user.

C. Hybrid Blackbox Algorithms

The combination of deep learning models and more traditional models has been pursued for a number of reasons. While different algorithms have their uses, being able to combine the ability to fully utilize a blackbox model's learning capabilities with the explainability of a traditional model comes with a variety of different benefits. Being able to extract a *decision tree* [18] for example, from an otherwise blackbox approach is extremely desirable, as it makes ensuring the comprehensibility and reliability of such a model easier. While traditional methods might have issues with nonlinear data, if the decisions can be recreated using decision trees, it becomes extremely desirable for the purposes of explainability. The hybrid use of blackbox systems can take many forms, such as using decision tree pruning as backpropagation [19], generating decision trees from backpropagation [20], rule extraction via decision tree induction [21], and sensor-classification fusion methods [22]. Being able to combine deep learning's ability to approximate complex relationships for problems with the explainability of decision trees allows for greater understanding of the hybrid model's decision-making process.

III. Experiment Design

A. The ESCAPE Dataset

The ESCAPE [23] dataset was published back in 2019 by the Air Force Research Laboratory (AFRL) Information Directory in order to enable multi-modal signature data-fusion research. In this dataset, the Experiments Scenarios Operations and Protype Engineering (ESCAPE) dataset combines a vast variety of different sensors, including EO, P-RF, radar, acoustic, and seismic data in a common scenario for the application of advanced fusion. Within the scenarios provided, there are a number of vehicle targets which attempt to avoid detection, and therefore "escape", providing any model using the data an incentive to use more than just the EO sensors.

The ESCAPE data has been analyzed for different data fusion and learning techniques such as feature-level fusion [24]. Building on popular data sets such as the moving and stationary target recognition (MSTAR) synthetic aperture radar (SAR), many developments have used the data set with advances in deep learning [25]. Likewise, the data fusion community has advanced to utilize the deep learning methods for joint multimodal fusion analytics [26]. Recently, efforts have been developed using the ESCAPE data for explanation analysis [27], with comprehensive comparisons over novel deep learning methods [28]. Another example is from an EO and SAR fusion dataset that affords comparisons of deep learning methods that emphasize explainability in performance improvements [29, 30].

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. The number of vehicle targets between the three scenarios totals 10, and each scenario deals with a different number of targets. The overall purpose of the dataset is that all the targets are designed to "evade" detection, by employing a number of different tactics that incentivize the fusion model to use different modality data input. The evasive scenarios comes with the benefit of thereby challenging any model or algorithm meant to differentiate between potential targets when engaging in tracking, as similar targets being moved in a manner that even human users might have difficulty in differentiating between them with only visual information. The three scenarios all involve multiple vehicles entering and exiting a garage, with multiple vehicles of similar make and build are combined with dissimilar vehicles. The movements of the vehicles are also varied to present challenging discrimination opportunities that confuse tracking differentiation.

Figure 1: Scenario 1

Scenario 1 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 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 #2 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.

Figure 2: Scenario 2

Scenario 2 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 #3 that was parked in the garage the entire time. This makes it appear that vehicle #1 enters and exits when in fact it is hidden inside of the garage, thus "escaping" detection successfully. 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.

Figure 3: Scenario 3

Scenario 3 is the most complicated of the three and chosen due to the complexity of the five vehicle targets, all 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.

B. Early Level Fusion and Preprocessing

The two sources of data are preprocessed in a number of different ways, with EO being preprocessed with Dense Optical Flow (DOF) and thresholding, P-RF data being preprocessed into histograms and with a Fast Fourier Transform (FFT) over the original I/Q data. As the primary focus is on the P-RF data, a less intensive form of EO preprocessing, thresholding, is chosen as an alternative for xLFER to use. While different sources of data enhancement have been tested in previous research [9], the usage of different interpretations of the same data is ideal for the purposes of gaining insights with the decision level fusion.

In order to extract more useful features from the P-RF data, the implementation of Wigner-Ville Distribution (WVD) and Continuous Wavelet Transform (CWT) are 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 as shown in Figures 4 and 5. CWT provides an overcomplete representation of the RF signal. Expanding on the uses of P-RF data provides more features for the model to utilize when EO sources alone are incapable of determining and locating the target vehicle.

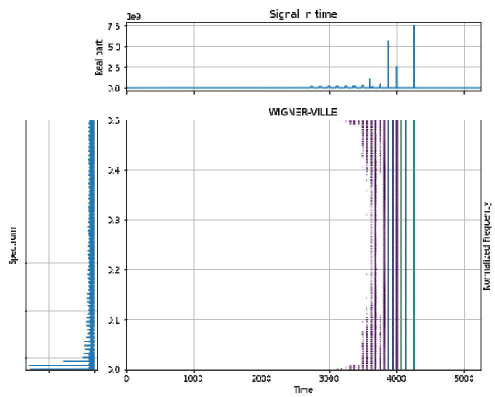

Figure 4: Wigner-Vile Distribution sample of P-RF data in Scenario 1

The generation of the WVD and CWT samples is taken from averaged bin files hosting the data in python. SciPy and the Time-Frequency analysis modules are utilized in order to process the P-RF data. From there, the rest of the data is independently processed for each modality, with the results being sent downstream for decision level fusion.

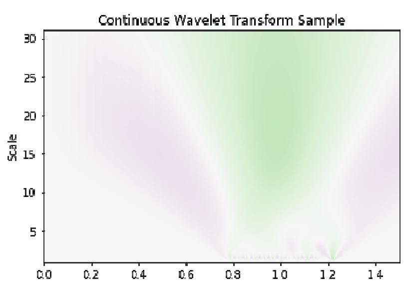

Figure 5: Continuous Wavelet Transform sample of P-RF data in Scenario 1

For the purpose of gaining explanations for the blackbox algorithms that will be used for the data, the use of Counterfactual Explanations is implemented. The performance of the models is kept separate from each other, with only the decisions the independent model's output being used to feed the decision-level fusion model. The decision-level fusion model from there uses the decisions made by the downstream models as training data.

C. *Late Fusion Design*

With the preprocessing completed and early models training is completed, the next stage is the late-stage decision level fusion. The xLFER model takes five separate sources of decisions (Histogram, CWT, WVD, DOF, and Thresholding as seen in Figure 6) and inputs them into the decision-level model. The late decision-level model, a decision tree, provides further insights using the generated decision tree's usage of the five inputs. The DT model's weights explain the model's decision-making process and thereby provide greater explainability than relying solely on visualization methods.

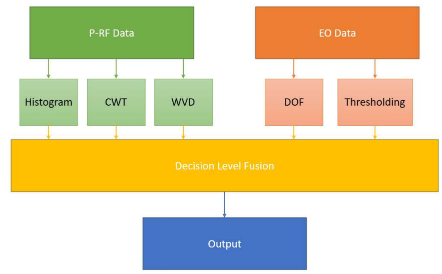

Figure 6: Late Fusion Model

One of the most useful features of this dataset, that our group has not been able to ideally capitalize on in previous research, is that the ground targets emit 13 frequency channels over a 4 MHz frequency band [23]. While it would be possible to preprocess the RF data, looking for those *specific signals* is not a part of the tracking and differentiation of different targets. Rather it is a signal of opportunity which *potentially* aids in differentiating between different targets. As this system is *passive* in nature, the system cannot focus on those exact signals, but it can process the RF data further to provide better visibility of the existence of these features, for the model to then utilize. In any real-life scenario, target vehicles will likely have some level of RF component; for example, emitted by wireless communication systems, mobile phones, and potentially self-driving car systems in the future, etc.

D. *Comparison Research*

With the usage of decision-level fusion, naturally the comparison research for this data will not be limited to just the initial Decision tree based fusion model. The comparison of the same input training data using Logistic Regression (LR), Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Gaussian Naïve Bayes (GNB), and a convolutional neural network (CNN) will be used to compare the results. Evaluation will be conducted via F1 Score to gain a better insight into the performances of each model. An F1 Score is calculated as the harmonic mean of both the precision and recall measurements. Precision measures how many retrieved values were relevant (Positive Predictive Value). Recall is a measurement of the completeness of the positive predictions (aka sensitivity).

IV. Results

A. *Early Level Fusion*

While the standalone P-RF modalities could be processed better, the EO based modalities, as seen below in Table 1, are able to achieve relative success independently. It is impossible for the EO modality on its own to achieve a perfect score, owing to how the ESCAPE dataset is designed, with several parts of the scenario purposefully obscuring the sight of the target. However, having a higher score is promising with respect to the information and its level of accuracy that it will provide to the decision-level model.

TABLE I. EARLY LEVEL FUSION COMPARISON

Base Modality	Preprocessing	F1 Score
EO	Thresholding	0.73
EO	Dense Optical Flow	0.88
P-RF	Histogram	0.27
P-RF	CWT	0.41
P-RF	WVD	0.56

As seen in Table I, there is a considerable improvement in the processing of the P-RF data. In previous research with P-RF data impact, the histograms were only able to achieve a modest amount of impact and accuracy. However, the usage of CWT and WVD has led to a drastic increase in the model's accuracy on the standalone models. This is reflected in the local and global impact, as described in later in part C.

B. Comparison with Similar Models

The comparison of Decision tree algorithm with Logistic Regression (LR), Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Gaussian Naïve Bayes (GNB) is used to showcase its performance with respect to other traditional methods, using the information from the EO (DOF), EO (Thresholding), PRF (Histograms), PRF (CWT), and PRF (WVD) models. While previous research has shown better results with earlier level fusion, the results overall were favorable for the decision level fusion approaches that were tested.

TABLE II. COMPARISON OF LATE DECISION LEVEL FUSION MODELS

Model:	F1 Score:
CNN	0.98
DT	0.92
LR	0.93
SVM	0.91
KNN	0.89
GNB	0.94

The Decision tree's (DT) performance aside, the improvement that the two other PRF models provide can be seen in the increased F1 scores. The KNN model had the least impact, with the SVM model showing a decent performance, while the LR and GNB were able to outperform the Decision tree's performance entirely, as seen above in Table II. The Decision tree's decision making was unsurprisingly focused on the EO sources, with Thresholding predictably having the lower impact of the two. The P-RF sources for CWT and WVD also predictably had a bigger impact than the decision-making process than the Histograms did, but otherwise is more or less in line with the weights the counterfactuals. While the DT model didn't fare as well as the LR and GNB models, the insights the tree weights provided were useful comparisons for the counterfactuals. While the baseline traditional models did not perform as well as the decision level fusion CNN, their performance was considerably improved from prior research with only P-RF histogram input.

C. Explainable AI

In order to gain explanations from the models tested, counterfactual explanations are generated using DiCE [13] (Diverse Counterfactual Explanations). Counterfactual explanations provide this information by showing feature-perturbed versions of the same sample based on different features, and thereby gaining insights into the model. This makes it possible to generate feature importance scores using a summary of the counterfactuals generated, determining its impact on both a local and global level.

TABLE III. COMPARISON OF EARLY LEVEL FUSION IMPACT

Modality	Local	Global
EO (Thresholding)	0.76	0.61
EO (DOF)	0.83	0.68
P-RF (Histogram)	0.16	0.32
P-RF (CWT)	0.37	0.41
P-RF (WVD)	0.53	0.57

The global importance scores per feature are estimated by aggregating the scores over individual inputs, while the local feature importance scores are computed for a given instance by summarizing a set of counterfactual examples around the point. As seen above in Table 3, the early-level fusion feature importance scores largely favor the EO modality. The results are largely in line with previous research with a greedy algorithm that assigned local and global feature impact scores [20], with the general trend being stronger EO impact than P-RF but with P-RF having a much larger impact on the global scale. Unlike the aforementioned research with ExplainX.ai, the use of counterfactuals relies on the model's decisions as opposed to approximating the original model via Greedy Algorithm.

V. CONCLUSION

In this work, we present a late-stage decision-level model that provides explanations for both early and later stages of data fusion. The usage of CWT and WVD for preprocessing provided noticeable improvements in the model's ability to utilize the P-RF data. With the improved accuracy of the P-RF models, the results indicate the models were able to exploit the preexisting signals to differentiate between different targets. The counterfactual explanations confirmed previous research into local and global impact with regards to behavior the fusion model used. Coupled with the results of the decision level fusion of the decision tree model, greater insights into the usage of P-RF data for target detection and differentiation is provided in this paper. While the decision tree model was not able to perform as well as our baseline, the decision-level fusion CNN, the traditional models were able to achieve a relatively close performance to the CNN. In future research, we would like to conduct further experimentation of other P-RF preprocessing methods and compare their impact on the fusion. It would further explore the explainability methods with visualizations for the explainability that can leverage the analysis from decision trees. With the determination of the features and sensor use, the methods can be explored for fog and edge paradigms [31] as well as hierarchal sensor management strategies such as with visual transformers [32].

ACKNOWLEDGMENTS

This research is partially supported by the AFRL Summer Faculty Fellowship Program. 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 United States Air Force or the U.S. Government.

REFERENCES

[1] M. Nauta, R. Walsh, A. D. and C. Seifert, "Uncovering and Correcting Shortcut Learning in Machine Learning Models for Skin Cancer Diagnosis," *Diagnostics,* vol. 12, no. 1, p. 40, 2022.

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

[3] E. Tjoa and C. Guan, "A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI," *IEEE Transactions on Neural Networks and Learning Systems,* vol. 32, no. 11, pp. 4793-4813, November 2021.

[4] R. V. Yampolskiy, "Predicting future AI failures from historic examples," *Foresight,* vol. 21, no. 1, pp. 138-152, 2019.

[5] Y. Li, A. W. Yu, T. Meng, B. Caine, J. Ngiam, D. Peng, J. Shen, B. Wu, Y. Lu, D. Zhou, Q. V. Le, A. Yuille and M. Tan, "DeepFusion: Lidar-Camera Deep Fusion for Multi-Modal 3D Object Detection," in *arXiv:2203.08195*, 2022.

[6] R. K. Mothilal, A. Sharma and C. Tan, "Explaining Machine Learning Classifiers through Diverse Counterfactual Explanations," in *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparencys*, New York, NY, United States, 2020.

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

[8] 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,* vol. 36, no. 7, pp. 44-61, 1 July 2021.

[9] E. Blasch, A. Vakil, J. Li and R. Ewing, "Multimodal Data Fusion Using Canonical Variates Analysis Confusion Matrix Fusion," in *2021 IEEE Aerospace Conference (50100)*, Big Sky, MT, USA, 2021.

[10] 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 *2018 IEEE Aerospace Conference*, Big Sky, MT, USA, 2018.

[11] E. Blasch, Z. Liu and Y. Zheng, "Advances in Infrared Image Processing and Exploitation using Deep Learning," in *Proceedings of the SPIE*, 2022.

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

[13] I. Nirmal, A. Khamis, M. Hassan, W. Hu and X. Zhu, "Deep Learning for Radio-Based Human Sensing: Recent Advances and Future Directions," *IEEE Communications Surveys & Tutorials,* vol. 23, no. 2, pp. 995-1019, 2021.

[14] Y. Nishida, "Proximity Motion Detection Using 802.11 for Mobile Devices," in *2007 IEEE International Conference on Portable Information Devices*, Orlando, FL, USA, 2007.

[15] H. Mu, J. Liu, R. Ewing and J. Li, "Human Indoor Positioning via Passive Spectrum Monitoring," in *2021 55th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, USA, 2021.

[16] J. Liu, R. Ewing, E. Blasch and J. Li, "Synthesis of Passive Human Radio Frequency Signatures via Generative Adversarial Network," in *2021 IEEE Aerospace Conference (50100)*, Big Sky, MT, USA, 2021.

[17] 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.

[18] G. Schmitz, C. Aldrich and F. Gouws, "ANN-DT: an algorithm for extraction of decision trees from artificial neural networks," *IEEE Transactions on Neural Networks,* vol. 10, no. 6, pp. 1392-1401, November 199.

[19] B. Kijsirikul and K. Chongkasemwongse, "Decision tree pruning using backpropagation neural networks," in *IJCNN'01. International Joint Conference on Neural Networks Proceedings (Cat. No.01CH37222)*, Washington, DC, USA, 2001.

[20] M. Zorman and P. Koko, "Hybrid NN-DT cascade method for generating decision trees from backpropagation neural networks," in *Proceedings of the 9th International Conference on Neural Information Processing, 2002. ICONIP '02.*, Singapore, 2002.

[21] M. Sato and H. Tsukimoto, "Rule extraction from neural networks via decision tree induction," in *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No.01CH37222)*, Washington DC, USA, 2001.

[22] 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.

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

[24] A. Vakil, J. Liu, P. Zulch, E. Blasch, R. Ewing and J. Li, "Feature Level Sensor Fusion for Passive RF and EO Information Integration," in *2020 IEEE Aerospace Conference*, Big Sky, MT, USA, 2020.

[25] U. Majumder, E. Blasch and D. Garren, Deep Learning for Radar and Communications Automatic Target Recognition, Artech, 2020.

[26] E. Blasch, T. Pham, C.-Y. Chong, W. Koch, H. Leung, D. Braines and T. Abdelzaher, "Machine Learning/Artificial Intelligence for Sensor Data Fusion–Opportunities and Challenges," *IEEE Aerospace and Electronic Systems Magazine,* vol. 36, no. 7, pp. 80-93, July 2021.

[27] A. Vakil, E. Blasch, R. Ewing and J. Li, "Finding Explanations in AI Fusion of Electro-Optical/Passive Radio-Frequency Data," *Sensors ,* vol. 23, no. 3, p. 1489, 2023.

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

[29] S. Low, O. Nina, A. D. Sappa, E. Blasch, N. Inkawhich, “Multi-Modal Aerial View Object Classification Challenge Results-PBVS 2023,” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 412-421, 2023.

[30] S. Low, O. Nina, A. D. Sappa, E. Blasch, N. Inkawhich, “Multi-Modal Aerial View Image Challenge: Translation From Synthetic Aperture Radar to Electro-Optical Domain Results-PBVS 2023,” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 515-523, 2023.

[31] A. Munir *et al.*, "FogSurv: A Fog-Assisted Architecture for Urban Surveillance Using Artificial Intelligence and Data Fusion," in *IEEE Access*, vol. 9, pp. 111938-111959, 2021.

[32] J. Wensel, H. Ullah, A. Munir, "ViT-ReT: Vision and Recurrent Transformer Neural Networks for Human Activity Recognition in Videos," in *IEEE Access*, vol. 11, pp. 72227-72249, 2023.