STUDENT SUMMER INTERNSHIP TECHNICAL REPORT

Implementation of Machine Learning and Deep Learning Algorithms to Facilitate and Automate Nuclear Power Plant Operations

DOE-FIU SCIENCE & TECHNOLOGY
WORKFORCE DEVELOPMENT PROGRAM

Date submitted:

December 10, 2021

Principal Investigators:

Roger Boza (DOE Fellow Student)
Florida International University

Ahmad Al. Rashdan, Ph.D. (Mentor / Principal Investigator)
Idaho National Laboratory

Ravi Gudavalli Ph.D. (Program Manager)
Florida International University

Leonel Lagos Ph.D., PMP® (Program Director)
Florida International University

Submitted to:

U.S. Department of Energy
Office of Environmental Management
Under Cooperative Agreement # DE-EM0005213

FIU | Applied Research Center
FLORIDA INTERNATIONAL UNIVERSITY

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor any agency thereof, nor any of their employees, nor any of its contractors, subcontractors, nor their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe upon privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or any other agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof.

ABSTRACT

This technical report provides an overview of the remote summer internship completed by Roger Boza with Idaho National Laboratory (INL) under the direct mentorship of Dr. Ahmad Al. Rashdan. Over the 10-week summer internship, Mr. Roger assisted on two distinct projects and worked in one. All of those projects aimed at facilitating and automating nuclear power plants operations. The three projects incorporated machine learning (ML) and deep learning (DL) state-of-the-art algorithms/architectures to perform computer vision (CV) tasks. The projects focused on specific CV tasks for automating manual staff operations that normally require considerable time and effort. The project Mr. Roger worked on, explainability of ml models for fire watch, aimed to analyze and explain the reasons why previously trained convolutional neural network (CNN) models made the types of predictions they did. A literature review was conducted to determine which approach was best suited for the analysis. The selected method, gradient class activation mapping (Grad-CAM), generated visual heatmaps that highlighted the importance of certain features in the images with respect to the predictions made by the CNNs. The first project that Mr. Roger assisted on, Automated Gauge Reading, aimed to automate the reading and monitoring of analog gauge instruments across nuclear power plants. In this project, CNNs and CV were used to help in the process on reading gauge values. The second project that Mr. Roger assisted on, Obstacle Detection for Drones, focused on improving a technology known as Route Operable Unmanned Navigation of Drones (ROUNDS), which was initially developed by INL. The main objective off the project was to use ML to detect obstacles in drone flight paths. All three projects saw the successful completion of their respective objectives. This report will discuss the explainability of ML models for fire watch as the main task for the internship.

ACKNOWLEDGMENTS

I wish to thank my mentor, Dr. Ahmad Al. Rashdan, for his invaluable stewardship throughout this research, as well as for sharing his domain expertise. I also wish to thank Brian Wilcken (INL) for his technical expertise and assistance, along with Mike Griffel (previous INL mentor) for his initial work on fire watch. Furthermore, I would like to extend my gratitude to interns Cesar Reyes and Blake Cecil for their collaboration on automated gauge reading and obstacle detection for drones, respectively.

This research internship was made possible thanks to the Applied Research Center (ARC) at Florida International University (FIU), the Department of Energy Office of Environmental Management (DOE-EM), and Idaho National Laboratory (INL). I personally thank thee organizations for making this research experience possible.

TABLE OF CONTENTS

LIST OF FIGURES

ACRONYMS

ANN	Artificial Neural Network
ARC	Applied Research Center
CNN	Convolutional Neural Network
CV	Computer Vision
DL	Deep Learning
DTD	Deep Taylor Decomposition
DOE-EM	Department of Energy Office of Environmental Management
FIU	Florida International University
Grad-CAM	Gradient Class Activation Mapping
INL	Idaho National Laboratory
LRP	Layer-wise Relevance Propagation
ML	Machine Learning
NPP	Nuclear Power Plants
ROUNDS	Route Operable Unmanned Navigation for Drones

1. EXECUTIVE SUMMARY

This research work has been supported by the DOE-FIU Science & Technology Workforce Initiative, an innovative program developed by the US Department of Energy's Environmental Management (DOE-EM) and Florida International University's Applied Research Center (FIU-ARC). During the summer of 2021, a DOE Fellow intern, Roger Boza spent 10 weeks doing a virtual summer internship with Idaho National Laboratory (INL) under the supervision and guidance of Dr. Ahmad Al. Rashdan (Senior Research and Development Scientist). The intern's project was initiated on June 1, 2021, and continued through August 5, 2021 with the objective of implementing machine learning (ML) and deep learning (DL) algorithms to facilitate and automate nuclear power plant operations.

2. AUTOMATED FIRE WATCH

2.1. Introduction

The U.S. Nuclear Regulatory Commission defines fire watch as “individuals responsible for providing additional (e.g., during hot work) or compensatory (e.g., for system impairments) coverage of plant activities or areas for the purposes of detecting fires or for identifying activities and conditions that present a potential fire hazard.”[7] Throughout the nuclear power industry, the required level of fire watch varies, depending on the plant conditions. At plants that require frequent fire watch activities (Figure 1), the cost of fire watch is substantial (i.e., can exceed $1M per month per plant). Due to recent sensing technology advancements, as well as the growing use of high-resolution cameras, great potential exists for automating fire detection in real-time via remote monitoring. This would not only lower costs, but also reduce human errors and improve safety.

Figure 1. During fire watch, someone must usually be present near the location of the potential hazardous fire, for as long as the hazardous condition exists. Replacing the person with a camera requires introducing forms of machine intelligence to recognize fire in a video stream and generate an alarm.[8]

2.2. Objectives

Investigate the explainability of previously trained mathematical models that can analyze every frame in a live video stream of an industrial environment and determine whether there is fire. The trained models were given the following requirements:

1) Detect as many fires as possible.
    a. Minimize false positives (i.e., detecting fire when there is no fire).
    b. Minimize false negatives (i.e., not detecting fire when there is fire).
2) Make predictions in real-time (20 ms or less).

2.3. Neural Network Models

Convolutional neural network (CNN) are the mathematical model templates most commonly used to classify images in machine learning (ML). They are good at extracting features from images (e.g., colors and shapes [in regard to fires]). These models analyze imagery data with the help of kernels (filters) and provide translation equivariant responses known as feature maps. Typically, CNNs have a fixed size architecture (Figure 2), but can be constructed in a manner that increases their width, depth, and resolution[1].

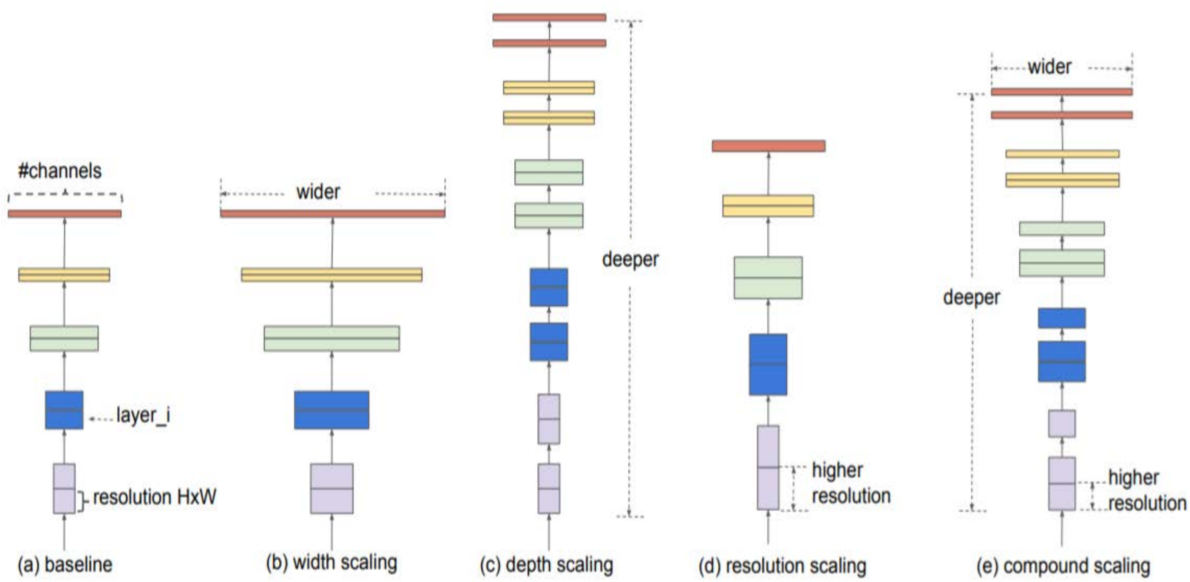

Figure 2. CNNs can have different structures that vary in terms of width and height.Model (a) is the most common type of CNN. Model (b) increases the width of the network to extract more features. Model (c) increases the depth of the network to extract abstract features (features of features). Model (d) increases the image resolution to get more details. Model (e) combines of all the various scaling approaches (i.e., [a]-[d]) and thus provides the most robust method.[1]

2.4. Model Training

Training successful CNNs usually requires large amounts of data (i.e., typically on the order of thousands), with diverse examples. This is because the model must learn which extracted features are important and be able to generalize them in order to make accurate predictions. A model that generalizes rather than memorizes (i.e., overfits) is said to be robust and is thus desired for industrial applications.

To train the models, over 12,000 images containing fire (6,000) and no fire (6,000) were previously collected from the Yahoo Flickr Creative Commons 100 Million (YFCC100m)[2] and Google's YouTube. This dataset was randomly split into data used to fit (train and tune) the models (Figure 3) and data used to test the resulting performance. The split partition was 80% for training and 20% for testing. Within the 80% that is used to train the model 10% was randomly selected for internal testing as the model continues to develop (validation). Many variations of CNN's were selected and trained on the same datasets.

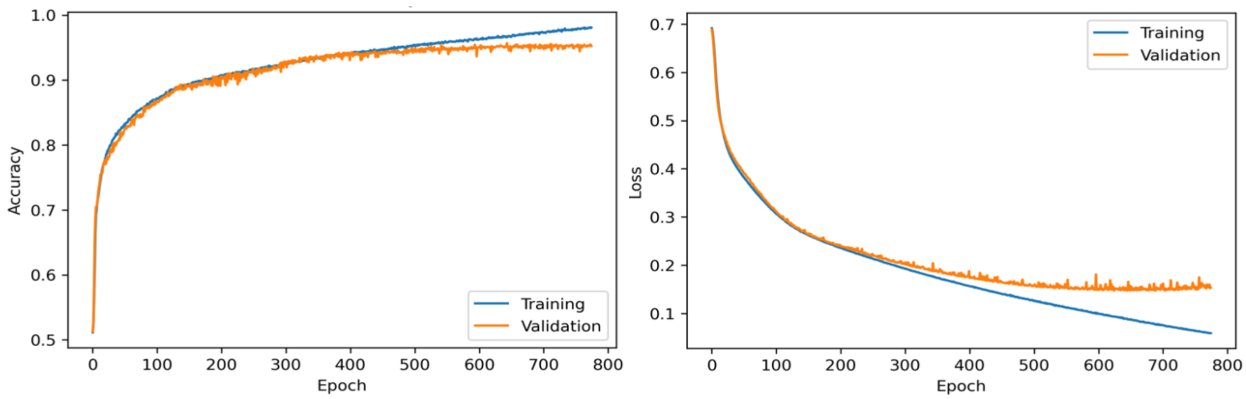

Figure 3. Model accuracy (left) and loss (right) metrics for training and validation datasets.The x-axis shows the number of epochs that are the training trials. As the model learns (i.e., the accuracy increases), the number of epochs increases. The loss metric helps the model determine when it's approaching an ideal state, thereby avoiding overfitting the model to the data.

2.5. Model Prediction Explainability

One challenge with image-based classification is knowing which areas of the image the decision was based on. Typically, the intermediate layers of a CNN, as shown in Figure 4, are considered a "black box" because they do not provide an explanation for the predicted output. A literature review of artificial neural network explainability showed that a variety of methods can be used to help explain the reasoning during the classification process.

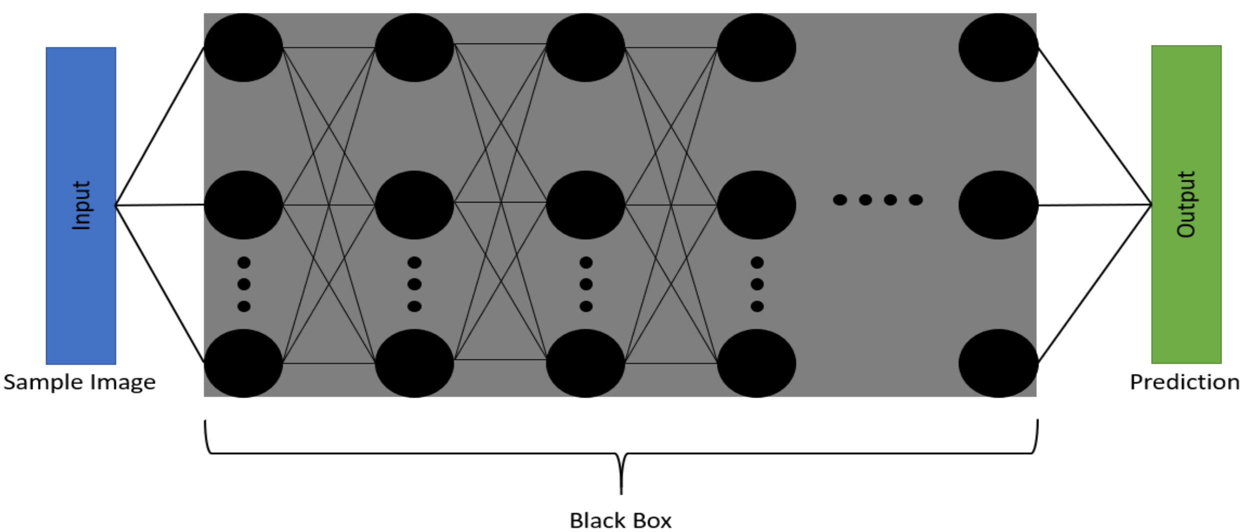

Figure 4. Example CNN layout, with input layer, intermediate layers, and output layer.The middle section, commonly referred to as a black box, contains most of the neural connections and weights for the model. It is very difficult to explain how and why this section's output led to the final decision.

2.6. Types of Explainability Methods

Sensitivity analysis [3] is based on the model's locally evaluated gradient, or some other local measure of variation. The most relevant features are those to which the output is most sensitive. However, sensitivity analysis, does not furnish an explanation of the function. It only produces a variation of it, and therefore unsuited to research.

Deep Taylor decomposition (DTD) [4] explains the model's decision by decomposing the function value f(x) as a sum of relevance score. DTD explains the nonlinear classification decisions made by the trained models in terms of the input variables. This method, based on the Taylor expansions, decomposes the output of deep neural networks. DTD can be applied to fully trained existing ANNs.

Layer-wise relevance propagation (LRP) [5] explicitly uses the feedforward graph structure of a deep neural networks to decompose the prediction. It computes a score for image pixels/regions that denote the importance of those particular regions to the trained model's final prediction. This technique propagates the prediction backwards using purposely designed local propagation rules.

Gradient class activation mapping (Grad-CAM) [6] produces visual explanations off the predictions made by CNNs. Grad-CAM uses the gradients with respect to the predicted class to create a coarse localization map that highlights the regions of greatest importance. This method uses global average pooling right before the final layer to analyze the pixel information, and is adaptable to previously trained CNN models.

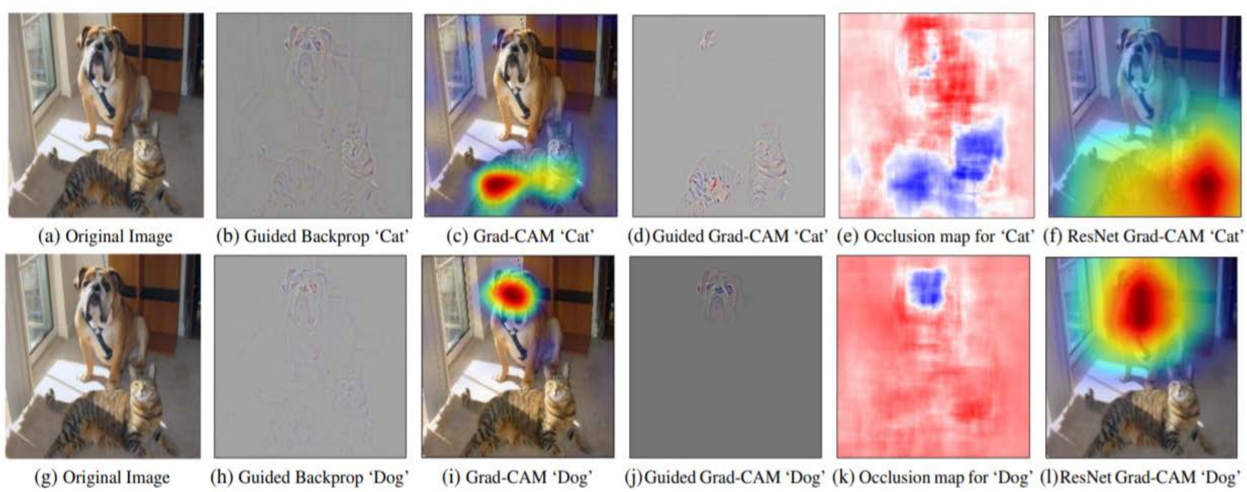

Figure 5. Prediction explainability visualization from Grad-CAM analysis.The pop row shows the Grad-CAM progression for determining the important features with respect to a "cat" classification. The bottom row shows the same progression, but with respect to a "dog" prediction.[6]

Of the methods and techniques uncovered in the literature review, Grad-CAM was selected as the best method for explaining CNN predictions. The previously trained models were modified so that the gradients could be exposed and analyzed. Furthermore, global average pooling was implemented and performed right before the final layer responsible for generating predictions. These modifications were done by enabling eager execution for the TensorFlow gradient tape. The models did not have to be retrained for the analysis to be completed.

2.7. Results

Grad-CAM was used to generate a heatmap visualization for the previously trained CNN models. Figure 6 shows the areas that most contributed to the fire classification decision. This helps in understanding whether the model predicted fire based on pixels that contained fire, or if it was coincidental because the image contained objects that resembled fire.

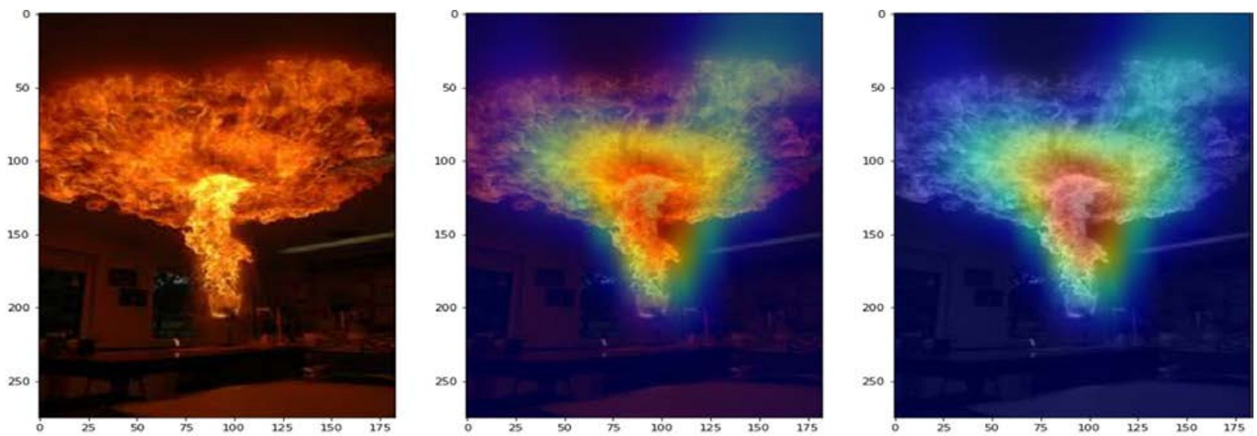

Figure 6. Grad-CAM verified the model is capturing firebased on the pixels that contain fire.(Left) Original image fed to the Grad-CAM algorithm. (Middle) heatmap visualization overlayed on the original image in red-green-blue format. (Right) Original image converted to grayscale then overlayed with the heat map for enhanced visibility (right).

2.8. Conclusion

The goal of this project was to investigate the explainability of previously trained mathematical models capable of recognizing fire in imagery data originating from an industrial environment, with few false positives/negatives. This research showed that some CNN models with high accuracy ignored the fire features and instead focused on other image areas showing traces of smoke, fire fighters, red engine trucks, etc. These models performed poorly on the testing data, and Grad-CAM was able to show why. The models that demonstrated high accuraccy and performed well on the testing data were shown to have high activation in the fire areas, as seen in Figure 7. This achievenment succesfully fulfills the objectives outlined for the project.

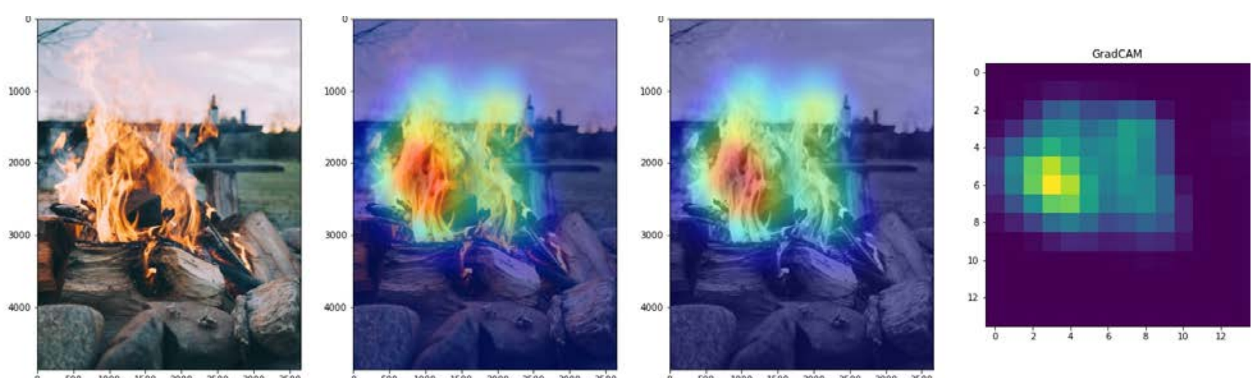

Figure 7. Grad-CAM analysis on testing image.(Left) Original image fed to CNN. (Right) Grad-CAM output. (Middle) Heat map visualizations showing the important features that led to a fire prediction.

3. REFERENCES

[1] M. Tan and Q. V. Le. "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks." 2019. [Online]. Available: https://arxiv.org/abs/1905.11946v3

[2] B. Thomee, D. A. Shamma, G. Friedland, B. Elizalde, K. Ni, D. Poland, D. Borth, and L. Li. "YFCC100M: the new data in multimedia research." 2016. [Online]. Available: https://doi.org/10.1145/2812802

[3] F. Pianosi, K. Beven, J. Freer, J. W. Hall, J. Rougier, D. B. Stephenson, and T. Wagener. "Sensitivity analysis of environmental models: A systematic review with practical workflow." 2016. [Online]. Available: https://doi.org/10.1016/j.envsoft.2016.02.008

[4] G. Montavon, S. Lapuschkin, A. Binder, W. Samek, and K. Müller. "Explaining nonlinear classification decisions with deep Taylor decomposition." 2017. [Online]. Available: https://doi.org/10.1016/j.patcog.2016.11.008

[5] S. Bach, G. Montavon, K. Müller, and W. Samek. "Layer-Wise Relevance Propagation for Deep Neural Network Architectures." 2016. https://doi.org/10.1007/978-981-10-0557-2_87

[6] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantan, D. Parikh, and D. Batra. "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization." 2016. [Online]. Available: https://arxiv.org/abs/1610.02391

[7] "Fire Protection for Operating Nuclear Power Plants." Apr. 2001, Accessed: Nov. 24, 2021. [Online]. Available: https://www.nrc.gov/docs/ML0109/ML010920084.pdf.

[8] Firewatch. 2011. Accessed: Nov. 24, 2021. [Online]. Available: http://www.nmslrhs.org/Photos/2011/07-02/slides/P7020155.html

SUMMER INTERNSHIP POSTER COMPETION

Each year, INL holds a poster competition regarding the cutting each technology and research in which the interns are engaged. This year, Roger Boza submitted a poster for each project in which he was involved. These posters were all submitted in the Nuclear Operations category. One poster, "Obstacle Detection for Drones Using Machine Learning," won first place (Figure 8).

Figure 8. Roger Boza holding the 1st place plaque for best Nuclear Operations poster at the 2021 INL intern poster competition. Image taken at FIU ARC.