INL/CON-22-67663-Revision-0

A Review on the State of the Art of Machine Learning and Satellite Imaging: Detecting Scene Changes in Selected Nuclear Fuel Cycle Datasets

June 2022

Changing the World's Energy Future

Shiloh N Elliott, Ashley Jeanette Brockman Shields, Eduardo Antonio Trevino

Idaho National Laboratory

INL is a U.S. Department of Energy National Laboratory operated by Battelle Energy Alliance, LLC

INL/CON-22-67663-Revision-0

A Review on the State of the Art of Machine Learning and Satellite Imaging: Detecting Scene Changes in Selected Nuclear Fuel Cycle Datasets

Shiloh N Elliott, Ashley Jeanette Brockman Shields, Eduardo Antonio Trevino

June 2022

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

A Review on the State of the Art of Machine Learning and Satellite Imaging: Detecting Scene Changes in Selected Nuclear Fuel Cycle Datasets

Shiloh N. Elliott
Idaho National Laboratory

Ashley Shields
Idaho National Laboratory

Eduardo Trevino
Idaho National Laboratory

1. Abstract

The timely detection of clandestine nuclear facilities is one of the greatest challenges faced by the International Atomic Energy Agency's (IAEA). Idaho National Laboratory is currently applying machine learning (ML) to existing satellite imagery (SI) datasets to find facilities within the nuclear fuel life cycle, with primary focus placed on identifying critical predecessor (i.e., fuel fabrication and fuel enrichment) and successor (i.e., nuclear power plants) facilities. This could provide a satellite image methodology that the IAEA could leverage to discover clandestine facilities. The work presented in this paper describes the evolution of a workflow developed by this team for object detection related to critical infrastructure by expanding that workflow for the purpose of identifying nuclear fuel cycle components and automating dependency assessments. This will be done using two methods housed within a single pipeline. The first method involves implementing a DenseNet161 convolutional neural network to classify the images and explain the results using Local Interpretable Model-Agnostic Explanations (LIME). The second method implements You Only Look Once version 5 (YoloV5), to detect objects within images, provide a probability for the detection, and provide a bounding box that corresponds to the object of interest. The results of this work are anticipated to provide a clear picture of this portion of the nuclear fuel cycle and perform as a stand-alone tool for image assessment that can be expanded to additional fuel cycle components and implemented in international safeguards and national security domains. This capability addresses the IAEA's need to detect undeclared nuclear materials and activities within a state while encompassing the entire nuclear fuel cycle.

2. Introduction

Machine learning (ML) continues to advance its available algorithms and their effectiveness, facilitating increasing model accuracy, trustworthiness, explainability, and efficiency. The combination of ML with satellite imagery (SI) is advantageous to the International Atomic Energy Agency's (IAEA's) because it can be applied as a cost-efficient tool for facility detection and change detection. These technologies allow areas of interest to be monitored remotely and thus allow the agency to reallocate the resources required for boots on the ground or human-centric image assessments. This is triggered by the Department of Safeguards Operational Division's requests, based on their annual implementation plan for each state. In these plans, applications for the detection of undeclared nuclear facilities have not been emphasized. This is a deficiency that we believe can be addressed through implementation of ML algorithms, demonstrated by the work presented here.

3. Background Work & Related Work

Object detection in SI is an extensively investigated field. Traditionally these investigations have been limited to specific types of objects, such as ships [1] or airports [2]–[4]. The motivations behind these detection efforts varies from infrastructure expansion to shipping and military movement detection [1], [5]. The various methods currently utilized by research vary in both framework and hyperparameters utilized. The majority of research efforts do fall into the Regional-based Convolutional Neural Network (R-CNN) family of architectures [1], [2], [6], [7]. In the nonproliferation domain SI has been utilized to detect proliferation activities. This includes detection of increased enrichment activities near Iran's Natanz Enrichment Plant in 2001 and detection of Iraq's clandestine construction of the Tuwaitha nuclear research facility in the early 1990s. These discoveries were manual in nature requiring a human to look at physical images to detect potential proliferation activities [8]. Recent work in change detection has posited that additional development is needed for current algorithms to be widely applicable in the nonproliferation space [9]. Similar work has noted the difficulty of constructing a modeling capability based on subject matter expertise, reliable data, and an adaptable modeling framework [10]. To date, there is no published work that details a methodology for the nonproliferation domain capable of identifying facilities coupled with dependency analysis nor is a framework for automatically generating synthetic data within this context for disparate datasets. The work presented in the following methods offers a novel and efficient approach to automating the process of change detection and dependency analysis for nuclear reactors.

4. Methodology

4.1. Data Generation

The team is currently utilizing Maxar's DigitalGlobe data in model testing and training efforts. Open-source methods were used to obtain and collected latitude and longitude data for all known reactor facilities. A total of 330 sites were identified. Given the co-location of certain reactor facilities, we expect the final imagery count will be less than the initial site estimation. Data pulls are being conducted for each point collected. This process is still ongoing, current summary statistics by reactor types and status can be viewed in Figure 1. The team is utilizing visible spectrum imagery data from three Maxar satellites GeoEye-1, WorldView-2, and WorldView-3 (Figure 2). All three sensors produce high resolution imagery varying in resolution between 1.8 to 0.41 meters [11]–[13]. Imagery from 19 sites could not be collected. These sites were either obstructed by clouds, or there was a lack of imagery for the targeted satellites, or reactors were co-located within the same image.

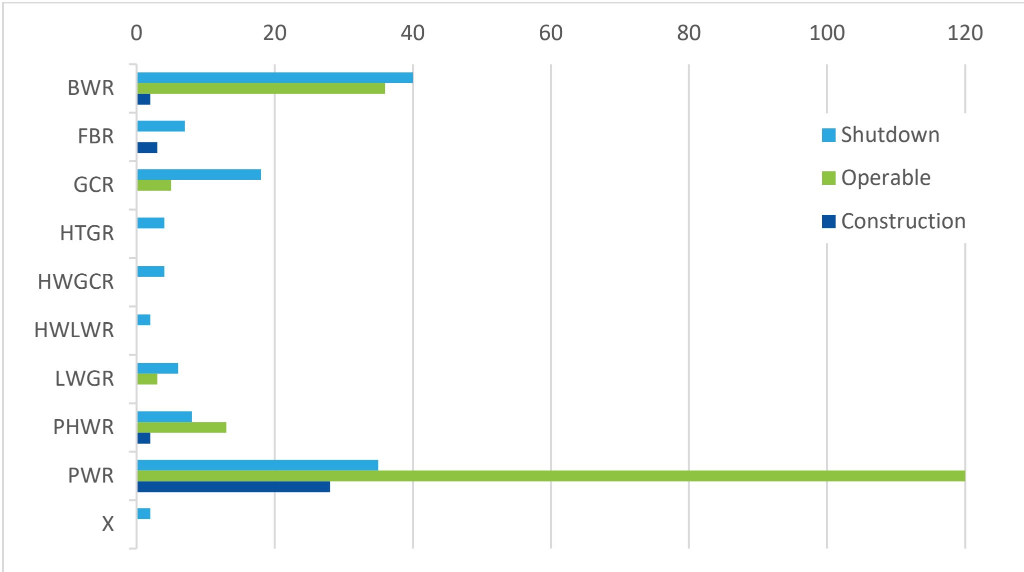

Figure 1. Chart depicting the operational status of the 10 reactor types of interest. BWR – Boiling Water Reactor, FBR – Fast Breeder Reactor, GCR – Gas-cooled Reactor, HTGR – High-temperature Gas Reactor, HWGCR – Heavy-water Gas-cooled Reactor, HWLWR – Heavy-water Light-water Reactor, LWGR – Light-water Graphite Reactor, PHWR – Pressurized-heavy-water Reactor, PWP – Pressurized-water Reactor, X – X-Energy

Figure 2. Comanche Peak Nuclear Power Plant in Texas. An example of WorldView-2 imagery (Imagery © 2022 Maxar).

4.2. Synthetic Data

Detecting nuclear power plants from SI datasets is heavily unbalanced (i.e., there are only a handful of facilities in the world). The ability to generate synthetic images, from our original data set, using different modalities of learning (e.g., varying resolution, camera height, and obstacles covering images), provides a more robust training data set for the deep-learning models to learn from. End-to-end deep-learning models designed for fast-object detection like YOLOv5, rely on several factors for performance, but most importantly, models rely on data. Since our model learns to predict bounding boxes from data, it struggles to generalize to objects in new or unusual aspect ratios, configurations, or other surface criteria that can change the data.

By generating synthetic data, we can combat these inconsistencies and limited data. To accomplish that, researchers use generative adversarial networks (GANs), a generative architecture originally proposed in 2014, by Ian Goodfellow et al. [14]. GANs typically consist of two models: a generative model that captures data distribution and a discriminative model that estimates the probability that a sample came from the training data as opposed to the generator [14]. Since the inception of the GAN architecture by Goodfellow et al., GANs have seen a rapid improvement in resolution and quality of image generation [14]–[17]. However, the synthesis process of GANs depends on absolute pixel coordinates, causing detail in an image to be glued to a coordinate, as opposed to surfaces of intended objects [18]. This poses as a potential dataset issue when using typical GANs. For this reason, Tero Karras et al. proposes alias-free GANs, a novel solution that traces the root cause to careless signal processing that causes aliasing in the generator [17]. Karras achieves this by interpreting all the signals in the network as continuous and applying small architectural changes to the GAN model to guarantee that unwanted information does not leak into the synthesis process [18]. To ensure that our synthetic image data is optimized for YOLOv5, plentiful for our dataset, and capable of generating unexpected or unrealistic image scenarios (improving classification during training), we are utilizing an alias-free GAN.

Figure 3. Atucha Nuclear Power Plant in Argentina, Lima, Buenos Aires. Image used in synthetic generation efforts. (Imagery © 2022 Maxar).

Figure 4. Completely Synthetic Nuclear Power Plants. A derivative product of WorldView-2 imagery (Imagery © 2022 Maxar).

4.3 DenseNet-161 Model

A preliminary assessment of the DenseNet-161 architecture resulted in an average accuracy of 86%, and the model has not yet been hyperparameter tuned or amended with synthetic data. Based on those initial results, we will proceed to implement the architecture for the final model. The DenseNet architecture [19] implements a densely connected convolutional neural network (CNN), where each node is fully connected to every other node in a series. This method is computationally expensive but performs well on complex data, such as the imagery datasets with diverse background pixels, as seen in nuclear domain datasets [19]. The literature supports that this architectural style is less prone to overfitting than alternatives, such as ResNet, and requires fewer parameters for model development [19].

4.4 Explainability

The team is currently utilizing LIME as an explainability framework. LIME is a model-agnostic framework that when applied to an image classification model is able to determine which portion of the image was utilized by the model to determine the classification class [20]. LIME accomplishes this task by dividing a given image into superpixels, defined regions within an image. These regions are then tested through linear regression to determine the importance of the region to correct classification of an image. The results of the regression model highlight areas of importance in an image with a higher weight while areas of less importance are denoted with a lower weight. An example of this can be seen in Figure 3 where blue indicates an area of great importance while red indicates an area of less importance in the image's classification.

Figure 5. Explainability analysis of a nuclear generation plant in Minnesota. Red areas indicate negative correlation, and blue areas indicate areas of positive correlation. (Imagery © 2022 Maxar)

4.5 YoloV5 Model

The You Only Look Once (YOLO) modeling family is based on the premise that an object detection task can be accomplished by dividing a single image into a number of bounding boxes with associated class probabilities. This approach allows the object detection task to be framed as a regression classification task. With each bounding box being assigned a probability of class ranking [21]. Since the creation of YOLO in 2015, there have been four acknowledged evolutions of the architecture. During these evolutions, improvements were made to each iteration: YOLOv2 introduced anchor boxes and improved performance, YOLOv3 provided a lighter faster model, and YOLOv4 introduced the idea of "Bag-of-Freebies" and "Bag-of-Specials" both of which enhance model performance without increasing computational modeling cost [22]–[24]. For our purposes, we will be utilizing YOLOv5 which is similar in architecture and design to YOLOv3 but has been integrated into the PyTorch ecosystem [25]. This integration allows for easy of model distribution and development of the modeling pipeline. This work is ongoing.

4.6 Proposed Framework

The proposed framework (Figure 4) describes our planned analytical workflow for nuclear reactor, nuclear fuel fabrication, and enrichment facility identification method. It begins with the input of imagery data for facilities of interest. This data is clipped and extracted from Maxar's DigitalGlobe. Once the data is acquired, the datasets will be assessed to ensure there are enough images present to construct a quality model. The example shown in Figure 6 uses a sample size of 500, but in practice, this number will be determined iteratively through the assessment of model results, including accuracy, loss, and model fit. If the dataset is determined to be too small, then synthetic data will be generated to artificially increase the sample size. In the case of having all "real" data or having partial synthetic data, the resultant dataset will be used to test and train a machine-learning model, and a model file will be produced. This exact same process is being conducted for nuclear reactors and will be conducted for fuel fabrication facilities and fuel enrichment facilities.

The resultant models will then be fed into an explainable decision framework using unknown data streams (or alternatively verification and validation data streams, depending on the maturity of the workflow). This new image will be assessed to identify if a nuclear reactor is present in the dataset, and LIME will be implemented to explain the decision we are making. If a reactor is detected in an image, its location will be queried to determine if the identified reactor is a known entity or if it is a previously undetected location through calculating its geospatial distance from known facilities. If the facility in question has yet to be identified, the workflow will call to the dependency models to assess the images for fuel fabrication facilities or fuel enrichment facilities. The results of each assessment will be explained using LIME, and the results will be concatenated to provide an overview of the state of the nuclear fuel cycle for the reactor of interest.

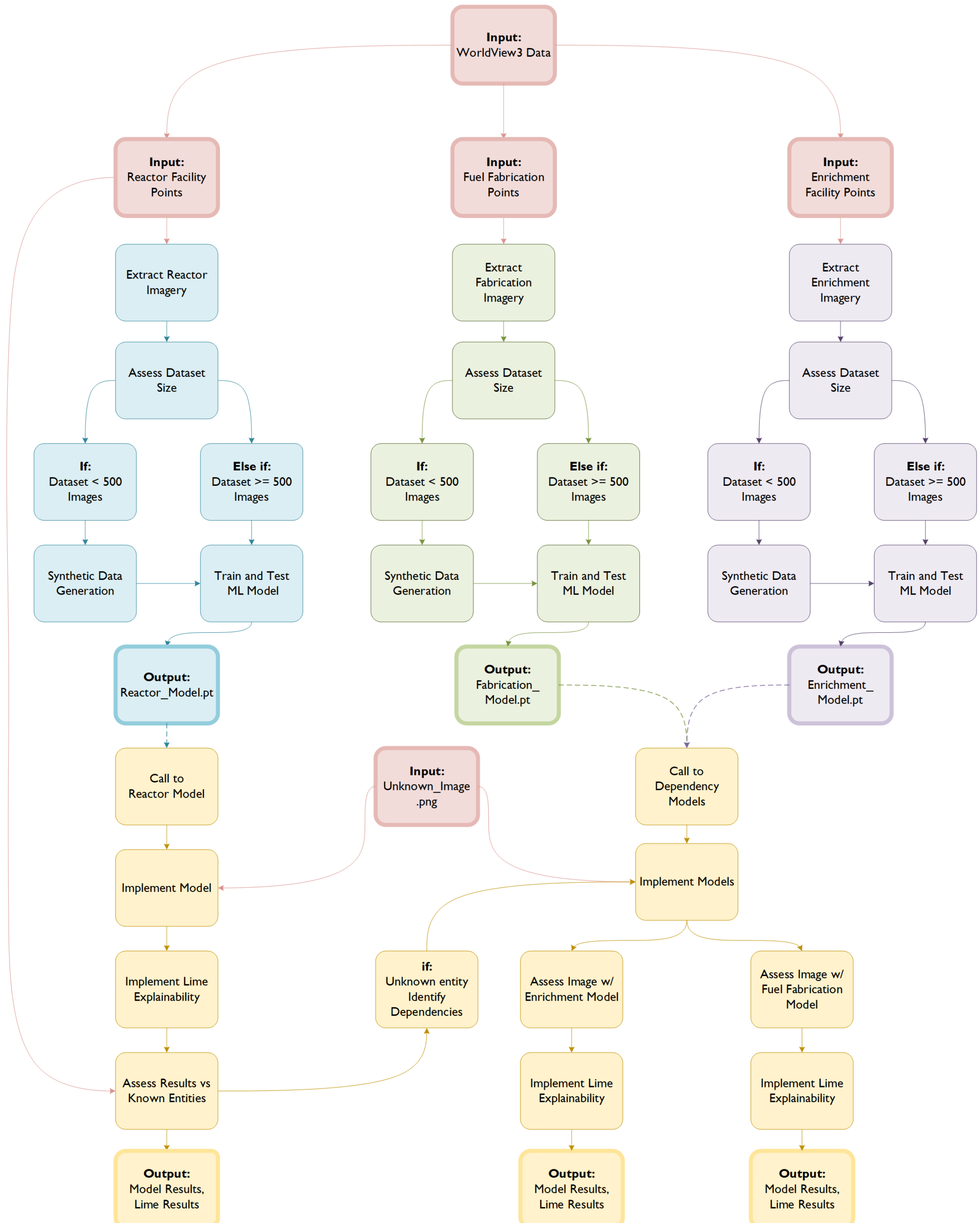

Figure 6. Proposed framework for nuclear reactor, nuclear fuel fabrication, and enrichment facility identification.

5. Conclusion

The work presented here offers a robust framework for approaching the IAEA's need for efficient change detection for remote monitoring of nuclear facilities. The advantage of this approach is its modular nature and the simplicity involved in expanding to facilities beyond the three highlighted here by amending the workflow with additional training data for additional features (e.g., uranium mining) and calling to the components developed for model development (e.g., Figure 6, Green) and image assessment (i.e., model implementation, LIME, and output results). This combined with a decision framework that includes synthetic data generation offers a new approach that is simple to leverage not only for nuclear fuel cycle facilities but also any physical features and their dependencies such as critical infrastructure analysis, infrastructure maintenance, and urban planning.

6. Future Work

In addition to potentially exploring other nuclear fuel cycle facilities, the authors realize that a potential adversary can disguise a physical structure, so an expected geometry and infrastructure is obscured. Another area of interest to explore is the use of hyperspectral imaging that could provide facility specific indicators irrespective of geometry for the purposes of detection. There is also potential to identify detection signatures through the fusion of multiple modeling techniques that could circumvent disguising efforts.

7. Acknowledgments

The authors would like to thank the Department of Energy National Nuclear Security Administration and Wayne Mei for their support of this work. Additionally, the team would like to thank Mark Schanfein and Gustavo Reyes for their mentorship and guidance. Finally, the team would like to thank Maxar's Digital Globe team for use of their data and support.

References

[1] S. Nie, Z. Jiang, H. Zhang, B. Cai, and Y. Yao, "Inshore ship detection based on mask r-cnn," in *International Geoscience and Remote Sensing Symposium (IGARSS)*, Oct. 2018, vol. 2018-July, pp. 693–696. doi: 10.1109/IGARSS.2018.8519123.

[2] P. Zhang, X. Niu, Y. Dou, and F. Xia, "Airport detection on optical satellite images using deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 8, pp. 1183–1187, Aug. 2017, doi: 10.1109/LGRS.2017.2673118.

[3] F. Chen, R. Ren, T. van de Voorde, W. Xu, G. Zhou, and Y. Zhou, "Fast automatic airport detection in remote sensing images using convolutional neural networks," *Remote Sensing*, vol. 10, no. 3, Mar. 2018, doi: 10.3390/rs10030443.

[4] S. Yin, H. Li, and L. Teng, "Airport Detection Based on Improved Faster RCNN in Large Scale Remote Sensing Images," *Sensing and Imaging*, vol. 21, no. 1, Dec. 2020, doi: 10.1007/s11220-020-00314-2.

[5] M. Mokhtarzade and M. J. V. Zoej, “Road detection from high-resolution satellite images using artificial neural networks,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 9, no. 1, pp. 32–40, 2007, doi: 10.1016/j.jag.2006.05.001.

[6] S. Zhang, R. Wu, K. Xu, J. Wang, and W. Sun, “R-CNN-based ship detection from high resolution remote sensing imagery,” *Remote Sensing*, vol. 11, no. 6, Mar. 2019, doi: 10.3390/RS11060631.

[7] H. Guo, X. Yang, N. Wang, B. Song, and X. Gao, “A Rotational Libra R-CNN Method for Ship Detection,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 8, pp. 5772–5781, Aug. 2020, doi: 10.1109/TGRS.2020.2969979.

[8] D. Albright, S. Burkhard, and A. Lach, “Commercial Satellite Imagery Analysis for Countering Nuclear Proliferation SATELLITE OBSERVATIONS OF EARTH FOR COUNTERING NUCLEAR PROLIFERATION,” *The Annual Review of Earth and Planetary Sciences is online at earth.annualreviews.org*, vol. 46, pp. 99–121, 2018, doi: 10.1146/annurev-earth-063016.

[9] J.-J. Han and N. K. Kim, “Accuracy Assessment of Change Detection Algorithm for Satellite Imagery in Support of Interpretation of Nuclear Activities.”

[10] J.-J. Han and N. K. Kim, “What Experts Care About in Satellite Observation : A Quantitative Analysis for Countering Nuclear Proliferation.”

[11] GeoEye, “GeoEye-1 Fact Sheet,” 2007. [Online]. Available: www.geoeye.com

[12] DigitalGlobe, “WorldView-2 Features Fact Sheet.” [Online]. Available: www.digitalglobe.com

[13] DigitalGlobe, “WorldView-3 Features Fact Sheet.” [Online]. Available: www.digitalglobe.com

[14] I. J. Goodfellow *et al.*, “Generative Adversarial Networks,” Jun. 2014, [Online]. Available: http://arxiv.org/abs/1406.2661

[15] T. Karras, S. Laine, and T. Aila, “A Style-Based Generator Architecture for Generative Adversarial Networks,” Dec. 2018, [Online]. Available: http://arxiv.org/abs/1812.04948

[16] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, “Analyzing and Improving the Image Quality of StyleGAN,” Dec. 2019, [Online]. Available: http://arxiv.org/abs/1912.04958

[17] T. Karras *et al.*, “Alias-Free Generative Adversarial Networks.” [Online]. Available: https://nvlabs.github.io/stylegan3

[18] T. Karras, T. Aila, S. Laine, and J. Lehtinen, “Progressive Growing of GANs for Improved Quality, Stability, and Variation,” Oct. 2017, [Online]. Available: http://arxiv.org/abs/1710.10196

[19] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 2261–2269, 2017, doi: 10.1109/CVPR.2017.243.

[20] M. T. Ribeiro, S. Singh, and C. Guestrin, “‘Why should i trust you?’ Explaining the predictions of any classifier,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2016, vol. 13-17-August-2016, pp. 1135–1144. doi: 10.1145/2939672.2939778.

[21] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection.” [Online]. Available: http://pjreddie.com/yolo/

[22] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” Apr. 2020, [Online]. Available: http://arxiv.org/abs/2004.10934

[23] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” Apr. 2018, [Online]. Available: http://arxiv.org/abs/1804.02767

[24] X. Huang *et al.*, “PP-YOLOv2: A Practical Object Detector,” Apr. 2021, [Online]. Available: http://arxiv.org/abs/2104.10419

[25] Ultralytics, “yolov5,” Jun. 25, 2020. https://github.com/ultralytics/yolov5 (accessed Jun. 06, 2022).