REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
		START DATE	END DATE
20230706	Final	20200801	20230831

4. TITLE AND SUBTITLE

Passive Four-Dimensional Imaging and Recognition in Low Light Levels with Visible Range Image Sensors

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
	N00014-20-1-2690	
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)

Javidi, Bahram

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Electrical and Computer Engineering Department, University of Connecticut, U-4157, Storrs, Connecticut 06269-4157; Email: Bahram.Javidi@UConn.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Office of Naval Research, One Liberty Center 875 North Randolph Street, Suite 1425 Arlington, VA 22203-1995	ONR	

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

We proposed to investigate object classification with novel passive sensing based on 3D integral imaging in low light levels with compact cameras in visible spectrum, and compare with LWIR cameras. The ability to detect, identify, and visualize targets and scenes passively in 3D in low light levels or with small compact cameras will be very important for applications of interest to US NAVY, including soldier centric applications.

15. SUBJECT TERMS

Passive 3D object recognition with visible range and Long Wave IR cameras, algorithms for 3D object recognition, object recognition in degraded environments

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	C. THIS PAGE		
U	U	U	SAR	18

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
Bahram Javidi	8604862867 Bahram.Javidi@UCONN.edu

PREVIOUS EDITION IS OBSOLETE.

STANDARD FORM 298 (REV. 5/2020)

Prescribed by ANSI Std. Z39.18

Bahram Javidi, University of Connecticut, Email: Bahram.Javidi@UConn.edu

Office of Naval Research (ONR)

Research Performance Final Report

Project Title: Passive Four-Dimensional Imaging and Recognition in Low Light Levels with Visible Range Image Sensors

Sponsor Name: Office of Naval Research

PI: Bahram Javidi
Board of Trustees Distinguished Professor
Electrical and Computer Engineering Department
University of Connecticut, U-4157
Storrs, Connecticut 06269-4157
email: Bahram.Javidi@UConn.edu

Award Number: Award # N00014-20-1-2690

Award Period Start: 01-Sep-2020

Award Period End: 31-Aug-2023

Accomplishments

What were the major goals and objectives of the project?

Object detection using conventional imaging and deep neural networks has been widely researched [1–5]. In conventional imaging, a high performance object detection can be achieved using deep neural networks. However, adverse conditions such as low photon count or occlusion deteriorate the performance of these systems. Photons are plentiful in long wave infrared (LWIR) spectrum, thus these domains have been used in low illumination conditions.

Thus, object classification in low light conditions has been investigated by imaging in the LWIR spectrum [6–8]. However, LWIR cameras are typically costlier, have much lower spatial resolution, and require bulkier optics compared to their visible range counterparts. Imaging in the visible spectrum can be an effective technique for object detection, however, the preeminence of conventional camera noise in photon starved environments degrades the classification performance.

In any imaging domain, adverse conditions such as occlusion degrade the object classification capabilities. Three-dimensional (3D) integral imaging (InIm) is a prominent 3D technique that works by capturing both angular and intensity information of the 3D scene [8–11]. 3D imaging is also contrasted with the 2D imaging which refers to the conventional 2D imaging obtained by camera and does not provide depth information. In low light, the 3D InIm reconstructed images have a higher signal-to-noise ratio (SNR) compared to conventional 2D imaging, since the 3D InIm is optimal in the maximum likelihood sense for read-noise dominant images [9,12].

We proposed to investigate object classification with novel passive sensing based on 3D integral imaging in low light levels with CMOS cameras, and compare with LWIR cameras. The ability to detect, identify, and visualize targets and scenes passively in 3D in low light levels or with mobile devices will be very important for applications of interest to US NAVY, including soldier centric applications. We will demonstrate that there are substantial benefits in using 3D passive imaging in low light levels with CMOS cameras or with LWIR cameras for visualization, seeing through occlusion and obscuration (such as trees, foliage, etc.), and real-time 3D object identification and tracking. The availability of 3D information in low light levels can offer new tools for information exploitation for the US NAVY war fighter. Our 3D sensing method is based on multi-view 3D integral imaging to record parallax and to extract range and 3D profile of targets. 3D information obtained from passive integral imaging with a camera array is inferred based on the different positions of the cameras. This is a passive sensing acquisition technology, applied in cases where active sensing technologies are inadequate, and non-controlled illumination conditions/scenarios are necessary.

Unlike LADAR which uses active illumination to measure time of flight, this multi modal passive sensing approach uses 2D image sensors with ambient light or thermal imaging. Fundamental questions have been addressed for reliable performance driven sensing, such as the light levels required for recording and reconstructing the 3D scene, the sensor requirements in terms of pixel parameters (size/ pitch/numbers), and parallax geometry. In addition, we have

experimented with LWIR cameras and LiDAR sensors. We have compared their performance with Integral Imaging systems using CMOS cameras in degraded environments to investigate the ability of these image sensors to visualize and detect scenes and objects in passive 3D in extremely low light levels which were not reported before.

The proposed approach contains many benefits compared with other 3D approaches. Stereo cameras suffer from depth estimation accuracy particularly in the presence of occlusions and degraded environment. RGB-D sensors are usually based on time-of-flight (ToF) technologies, or structured light (SL). A time of flight (ToF) 3D imaging system consists of a passive CMOS image sensor (camera) with an active modulated light source such as laser or light emitting diodes (LEDs). The active light source is used to illuminate the scene, and the reflected light from the scene is acquired using a dedicated sensor. The phase difference between the emitted and reflected lights is used to estimate the distance and depth information of the objects in the scene. Range cameras can be used to combine depth information with 2D intensity cues. ToF requires a high accuracy for wave delay measurement, in order to estimate the corresponding depth information. It has been typically used for indoor and short-range applications. Also, they may not perform well in the presence of occlusions and degraded environment.

Our investigation and experimental results demonstrate that for the experiments performed for low illumination and occlusion conditions, the performance of detection model on 3D InIm in the visible spectrum, on average, is better compared to the traditional 2D imaging systems in both visible and LWIR spectrum. It is also evident that for the traditional 2D imaging systems, object detection in the LWIR spectrum outperforms that in the visible spectrum. This explains the ubiquity of the LWIR cameras for imaging in low illumination and challenging environmental conditions. However, our experiments demonstrate that visible domain 3D integral imaging may outperform LWIR based 2D and 3D imaging systems in low light conditions plus occlusion. Compared to the traditional LWIR imaging, 3D visible domain integral imaging systems can provide object detection with a higher F1 score and a significantly higher precision score. [22-23]

What was accomplished towards achieving these goals?

We have performed the tasks and objectives of the project, including 1) developing dedicated algorithms and performing experiments on with passive integral imaging for object recognition in degraded environments with visible range image sensors. Degraded conditions include low light levels, occlusions and obscurations, fog, etc. We have designed and implemented the complete system for data collection and experimentation, developed dedicated algorithms using both convolutional neural networks and deep learning as well as comparison with statistical algorithms, and comparing our proposed approach with conventional systems for object recognition in low light levels and other adverse conditions. Figure 1 illustrates sample images of the scene that includes various objects to be classified and the presence of degraded environment. The scenes contained both hot and cold versions of all the objects.

The final aim of this project was to compare the performance of object detection in visible and LWIR spectrum. To reduce the bias during the training phase, the training scenes were captured simultaneously with visible and LWIR camera. With this, the exact same scenes are available for both LWIR and visible spectrum for training purpose. Each scene contained both hot and cold objects.

The most important discovery that we have made is that in adverse conditions the performance of passive 3D imaging in visible range using dedicated algorithms may exceed that of LWIR imaging. LWIR imagers are used extensively in DoD. Thus, this discovery can be transformative if additional tests and performance evaluations with a larger set of data including field tests confirm these results. Visible range cameras are far less expensive and more compact than LWIR cameras, and our findings can be of great interest and benefit to the DoD community.

We have developed dedicated algorithms and 3D passive imaging scenarios, and performed object detection and classification in low light environment and in the presence of obscurations in visible domain and longwave IR (LWIR) imaging. Preliminary experimental results are obtained. We have captured test images in low light conditions and occlusion using passive 3D integral imaging. Passive 3D integral imaging is chosen as it is optimal in the maximum likelihood sense for read-noise dominant images. This has advantage in low light conditions. In low light, denoising of the images is required before they can be used for 3D scene construction. Thus the object detection pipeline includes low-light enhancement of the captured elemental images, 3D reconstruction using integral imaging to mitigate the effects of adverse conditions, e.g. occlusion, and finally object detection and classification using convolutional neural network with deep learning.

We have used passive 3D InIm [16-23], and have compared its performance to active 3D imaging for object recognition in degraded environments. 3D InIm [22-22] is performed by recording multiple 2D elemental images of a scene from diversity of perspectives. These can be captured either through a single image sensor with lenslet array, camera array, or by a single camera mounted on a translation stage. The 2D elemental images are then integrated to obtain 3D information of the scene either through optical or computational reconstruction methods. The computational reconstruction is accomplished by backpropagation of optical rays through a virtual pinhole. The reconstruction depth can be set to any value within the depth of field of the captured elemental images.

In this project, object detection and classification has been performed using deep neural networks. However, we believe that using statistical techniques may have indicated superior performance of passive 3D integral imaging in visible range spectrum over its IWIR counterpart. A variety of the state of the art convolutional neural networks and deep learning algorithms have been used in this report to process the captured multi-dimensional images such as different versions of the You Look Only Once (YOLO) networks. The convolution network detection model determines the bounding boxes of the objects present in the image and classifies them simultaneously. The input image is divided in S X S grid cells. Each cell predicts the probability of an object being present. If an object is determined to be present in the cell, then it predicts the coordinate of boxes surrounding it and the confidence score for those boxes. Each bounding boxes predicted five parameters [x, y, w, h, c]. The (x, y) coordinate represents the central

location of object inside the box and (w, h) represents the height and width of bounding box and c is the confidence score. Each cell also predicts conditional class probabilities of the objects. The product of conditional class probabilities and confidence score of individual boxes gives the class-specific confidence scores for each box. These scores encode probability of class appearing in the box as well as how well the predicted box fits the object.

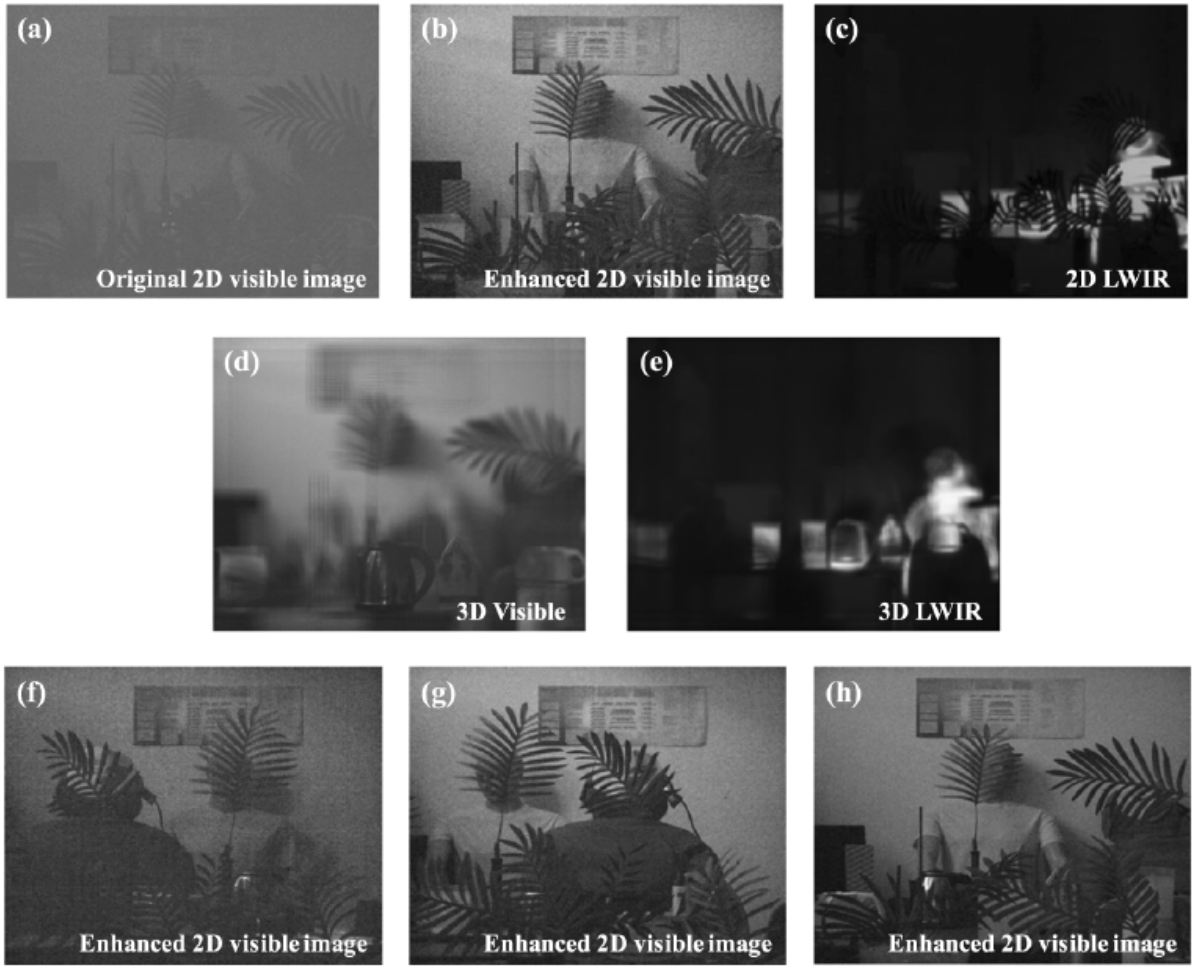

Fig. 1. (a) Sample low illumination 2D visible test image (3.3 Photons per pixel). (b) 2D image after enhancement using the dehazing and ICA pre-processing pipeline. (c) Sample low illumination 2D LWIR test image corresponding to the scene in (a). (d) Sample low illumination 3D visible test image corresponding to the scene in (a) reconstructed at the plane of the tall coffee pot (3.68 m). 2D elemental images used for its reconstruction were enhanced using the pre-processing pipeline as described in previous sections. (e) Sample low illumination 3D LWIR test image corresponding to the scene in (a). (f), (g), and (h) are additional sample 2D images after denoising enhancement in the visible domain. [22]

For visible spectrum, YOLO network was used to train object detection network for LWIR spectrum. Hot and cold objects have different contrasts in the LWIR spectrum. As such, two

detectors were trained for LWIR images. One only with hot objects, and one with both hot and cold objects. Our aim is to study the detection performance for both methodologies. Towards the end, we would like to compare the object detection results of both these methodologies for LWIR spectrum with that of visible spectrum.

The resolution between the captured images in LWIR and visible is different. Image resolution depends on both the sensor pixel density, and the operational wavelength. Visible range imaging produces much higher resolution than LWIR imaging. Visible range imaging has more than an order of magnitude higher resolution than LWIR imaging. Thus, resolution difference due to small wavelength is much better compared to that of the sensor pixel density. Visible camera has different sensor pixel density compared to the LWIR spectrum camera. We can mitigate the effects of different pixel densities. This can be accomplished by changing the pixels per object information of the images captured by the visible camera. This is valid if we assume that our scene is present in the paraxial regime, and that there is no significant angular separation between the two cameras. These conditions are a week form of the conditions assumed in the previous section. As such, we do not have problem satisfying these conditions. Once these conditions are satisfied, we can accomplish the task of modifying the pixels per object information by using spline based interpolation to transfer the pixel values from one frame to the other.

Both visible and LWIR training images were augmented by perspective diversity, mirror flip, and environmental degradation. The test images are to be collected in low illumination conditions. Thus, environmental degradation to be introduced in the training data set must reflect the approximate conditions in which test images would be captured.

In low illumination conditions, the visible spectrum camera captured images with very low photons per pixels. In photon-starved conditions, due to low number of photons, the captured images are dominated by camera noise. Our objective is to find a simple method to get an approximate estimate of the noise in low illumination conditions. This approximate estimate can help guide our effort to augment the data for neural network training. It should be noted, that we do not wish to find a rigorous way to replicate the exact conditions of low illumination imaging. In fact, replicating exact conditions of the test data set can lead to bias in the network.

In summary, we have investigated object detection and classification in low light conditions and in obscurations using passive 2D and 3D imaging by visible and LWIR imaging. It was found that for our purpose, CNN deep learning using YOLO provided the best results. However, 2D imaging was unable to detect occluded object or low light objects. Here 3D integral imaging was used to see through the occlusion. With this, the YOLO network was able to detect objects in 3D even in occluded environments. We have tested various low light conditions to check the robustness of the object detection pipeline. It was observed that even with as low as 3-4ms exposure time (3-4 photons per pixel), we are able to detect objects with proper algorithms including pre-processing denoising of the elemental images, and 3D reconstruction via integral imaging. We plan to do extensive testing and rigorous analysis of the performance of object detection pipeline for both visible and LWIR spectrum, and comparing these two performance on both occluded and non-occluded scenes in low-light conditions.

The important discovery is that in adverse conditions the performance of passive 3D imaging in visible range using dedicated algorithms may exceed that of LWIR imaging. LWIR imagers are used extensively in DoD. Thus, this discovery can be very transformational if additional tests confirm these results.

Summary of Significant Results

In this project, our aim is to compare the object classification capabilities between visible and LWIR imaging systems in adverse environmental conditions such as occlusion and low illumination. This comparison considers both conventional 2D and 3D InIm in both the visible and LWIR spectra. In our experiments, both cold and hot objects are considered in the scene. A CMOS image sensor is used for implementing the passive 3D InIm system in the visible spectrum. We have used the 'You Look Only Once version neural network for object classification, although other state of the art deep learning algorithms may be used. We compare the performance of the neural network for these systems using miss-rate, F1 score, and precision score. Various denoising models are considered and compared. Our experiments in low illumination and occlusion conditions indicate that for the experiments performed the performance of 3D InIm detection model, on average, is better compared to the traditional 2D imaging systems in both visible and LWIR spectrum. Most importantly, visible domain 3D integral imaging may outperform LWIR based 2D and 3D imaging systems. The LWIR based detectors performed very poorly when used to detect cold objects whereas 3D InIm in visible spectrum performed well. CMOS image sensors are ubiquitous with their performance rapidly increasing and their cost rapidly decreasing. The reported experimental results may indicate an alternative approach to LWIR systems for imaging in low light conditions. [22]

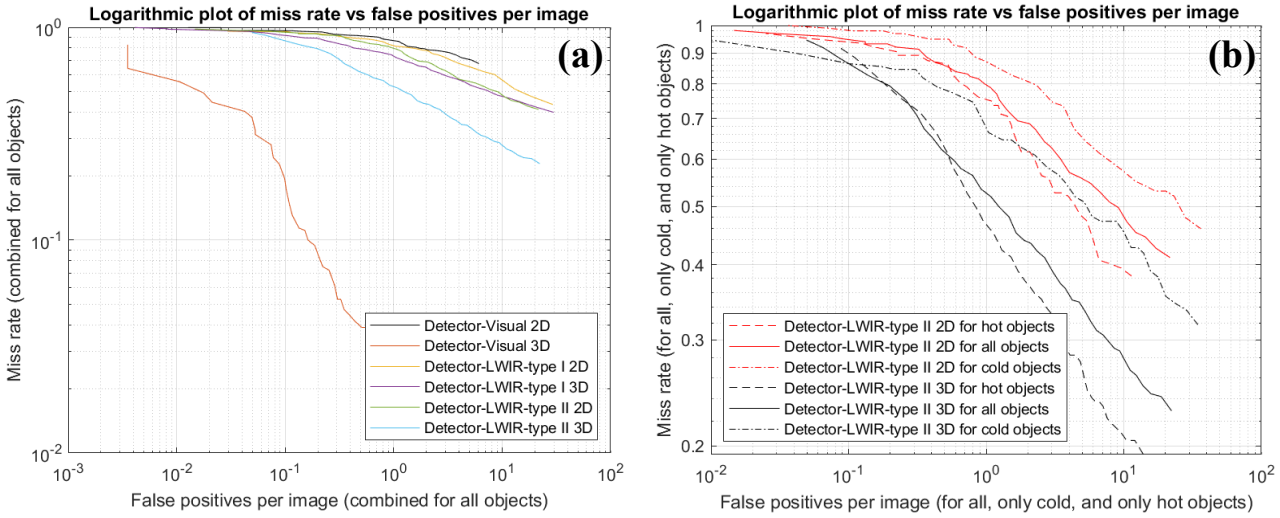

Fig. 2. (a) Average miss rate versus the number of false positives per image for all detection modalities (visible and LWIR) in both 2D and 3D. The plot is displayed on a logarithmic scale. This plot represents the average over all the objects in the scenes. (b) Average miss rate vs. false positives per image plot for the best LWIR detector (type II) compared between all, only cold, and only hot objects in both 2D and 3D. We denote the detector trained on the visible spectrum images as 'Detector-Visual', the LWIR detector trained only on hot objects is denoted as

'Detector-LWIR-type I', and the LWIR detector trained with both hot and cold versions of the objects is denoted as 'Detector-LWIR-type II'. [22]

As three separate detectors are considered in this work, we denote the detector trained on the visible spectrum images as 'Detector-Visual', the LWIR detector trained only on hot objects is denoted as 'Detector-LWIR-type I', and the LWIR detector trained with both hot and cold versions of the objects is denoted as 'Detector-LWIR-type II'. In each case, both conventional 2D imaging and 3D imaging using integral imaging are considered. Network detection threshold was set to 0.5 and the intersection over union ratio (IOU) threshold for object detection was also set to 0.5. The prediction is considered correct if the intersection over union ratio (IOU) for the predicted bounding box and the ground truth exceeds a predetermined threshold, set to 0.5 for our experiments. IOU is used to measure the overlap ratio between the detected and the ground truth bounding box. Its value varies from 0 (no overlap) to 1 (total overlap). We quantify the object detection performance using metrics such as precision, recall, F1 score, and miss rate. Precision and recall are defined as Precision = TP / (TP + FP), and Recall = TP / (TP + FN). Here, TP is the number of true positives, FP is the number of false positives, and FN is the number of false negatives. The miss rate quantifies the rate of true objects that skip detection. It is defined as Miss Rate = FN / (TP + FN). The F1 score is a higher order metric that combines precision and recall values to provide a more intelligible denotation to the performance of the network. It is defined as F1 = 2 * Precision * Recall / (Precision + Recall). The value of F1 score ranges from 0 to 1 with larger values being construed as that corresponding to a better performing network.

The average results for the five objects in Fig. 1 have been tabulated in Table 1 below. The scenes contained both hot and cold versions of all the objects. For each object, the ratio of number of occurrences of its hot and cold versions in the scenes was approximately 55:45.

Table 1. Average performance scores for object detection on all objects using YOLOv2 network

	Average Precision	Average Recall	Average F1	Average Miss Rate
Detector-Visible 2D	0.6429	0.1475	0.2400	0.8525
Detector-Visible 3D	0.9610	0.8222	0.8862	0.1778
Detector-LWIR-type I 2D	0.4130	0.1434	0.2129	0.8566
Detector-LWIR-type I 3D	0.3139	0.2494	0.2779	0.7506
Detector-LWIR-type II 2D	0.4341	0.2113	0.2843	0.7887
Detector-LWIR-type II 3D	0.2830	0.5784	0.3801	0.4216

It must be noted that the results for the LWIR spectrum are different for hot and cold objects. As such, the average results for LWIR spectrum mentioned in Table 1 are dependent on the ratio of number of occurrences of hot and cold versions of the objects present in the scenes. Table 2 and 3 below provides the results of the LWIR detectors separately for hot and cold objects. Table 2 considers only the hot objects in the LWIR spectrum, and Table 3 considers only the cold objects in the LWIR spectrum.

Table 2. Average performance scores for object detection on hot objects using YOLOv2 network

	Average Precision	Average Recall	Average F1	Average Miss Rate
Detector-LWIR-type I 2D	0.4222	0.2275	0.2957	0.7725
Detector-LWIR-type I 3D	0.3139	0.3477	0.3299	0.6523
Detector-LWIR-type II 2D	0.4623	0.2934	0.3590	0.7066
Detector-LWIR-type II 3D	0.3183	0.6595	0.4294	0.3405

Table 3. Average performance scores for object detection on cold objects using YOLOv2 network

	Average Precision	Average Recall	Average F1	Average Miss Rate
Detector-LWIR-type I 2D	NaN	0	NaN	1.0
Detector-LWIR-type I 3D	NaN	0	NaN	1.0
Detector-LWIR-type II 2D	0.3043	0.0714	0.1157	0.9286
Detector-LWIR-type II 3D	0.1889	0.3727	0.2508	0.6273

** NaN values for average precision and average F1 scores arise when the detector does not make a single prediction for those objects throughout all the scenes.*

The experimental results show that the LWIR based detectors perform very poorly when used to detect cold objects. This should be expected since cold objects provide little to no information in the LWIR domain. Their utility is most pronounced in the detection of hot objects. Varying the detector threshold allows us to plot the average miss rate with the number of false positives per image. The average plot for all the objects has been shown in the Fig. 2(a). For more clarity, Fig. 2(b) shows the plot for the best LWIR detector (type II) compared between all, only cold, and only hot objects. [22]

For type I and II LWIR detectors, the precision for 3D InIm with LWIR drops below that of the 2D imaging system. This is in contrast to the visible range detector, where the precision for 3D InIm significantly exceeds that of the 2D imaging system. We speculate that this may be attributed to the spatial features in both domains. Objects typically have numerous features in the visible domain, while maintaining a smooth profile in the LWIR domain. In visible domain 3D reconstruction brings forth the numerous features concealed behind the occlusion, thus boosting the detector precision. In contrast, 3D reconstruction in the LWIR domain offers minimal boost to the number of available features, while introducing the blurring effects of the occluding objects. Its combined effect may be responsible for the slight drop in the precision of the LWIR detectors. However, it should be noted that for LWIR 3D InIm the recall scores (0.2494, and 0.5784) are much higher than that of their 2D counterparts (0.1434, and 0.2113), thus suggesting that many more objects are getting detected albeit with low average precision. Thus although the

precision scores for LWIR 3D are smaller compared to that of LWIR 2D, the F1 scores are higher and miss rates are significantly lower for 3D InIm system.

Comparing the performance of the visible range imaging system with the LWIR based imaging system (with type II detector) provides different results for 2D and 3D object detection. For 2D imaging system, the LWIR object detector outperforms the visible object detector. Average F1 score for 2D LWIR (0.2843) is higher than that of 2D visible (0.2400), and the average miss rate of LWIR (0.7887) is lower compared with the visible (0.8525). For traditional 2D imaging in low illumination and adverse environmental conditions, the LWIR based imaging systems offer better performance compared to the visible spectrum imaging systems. However, incorporating the 3D InIm for imaging significantly improves the performance of the visible imaging systems under the degradations considered. In LWIR spectrum, the objects of interest are already well segmented from the background, and thus 3D InIm does not give significant boost to LWIR compared that to the visible spectrum images. With 3D InIm, the average precision value for visible spectrum object detection (0.9610) is significantly higher than that for LWIR object detection (0.2830). The same is observed with average F1 score, with visible systems having an average score of 0.8862 compared to the LWIR systems with average score of 0.3801. The miss rate for 3D visible (0.1778) is significantly smaller compared to the LWIR system (0.4216). With 3D InIm, it can be observed that, visible spectrum imaging system is able to detect objects with a higher F1 score, lower miss rate, and a significantly larger precision score compared to the LWIR system.

We have experimentally investigated object classification performance using 3D integral imaging with a deep neural network in adverse environmental conditions of low light illumination and partial occlusions. The experimental investigations were performed in visible spectral band with a CMOS camera, and in long wave IR domain with a LWIR camera, for both cold and hot objects in the scene. LWIR imaging has been a common approach to detect objects in low light levels. The ability to perform reliable object classification in low light and degraded environment in the visible spectral band with ubiquitous CMOS image sensors has many advantages including providing more than an order of magnitude improved spatial resolution over LWIR camera. The YOLO neural network was used for object detection and classification in the visible and LWIR domains. 3D integral imaging mitigated the effects of adverse environmental conditions providing improved performance. Experimental results show that for the experiments we performed, 3D integral imaging in the visible spectrum gives higher F1 score, a lower miss rate, and a significantly higher precision score than 1) 2D visible spectrum object detection, as well as, 2) both 2D and 3D object detections in LWIR spectrum. Our results demonstrate that for certain applications, such as in low illumination and adverse environmental conditions, passive 3D imaging systems in the visible domain may offer a viable alternative to the costlier, bulkier LWIR imaging systems. While these experiments are not comprehensive, they demonstrate the potential of 3D InIm in the visible spectrum for low light applications. [22-23]

The deep learning algorithms have evolved and gotten better since we started this project three years ago. It should be noted that even though we have not used the state of the art deep learning algorithms, the performance of our proposed approach outperforms conventional image

classification systems. We anticipate that using state of the art deep learning algorithms will produce improved outcome.

References:

1. D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov, "Scalable object detection using deep neural networks," IEEE Conference on Computer Vision and Pattern Recognition, 2155–2162 (2014).
2. Y. Tian, P. Luo, X. Wang, and X. Tang, "Deep learning strong parts for pedestrian detection," IEEE International Conference on Computer Vision, 1904–1912 (2015).
3. P. F. Felzenszwalb, R. B. Girshick, D. Mcallester, and D. Ramanan, "Object detection with discriminatively trained part-based models," IEEE Trans. Pattern Anal. Mach. Intell. 32(9), 1627–1645 (2010).
4. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," IEEE International Conference on Computer Vision, 770–778 (2015).
5. V. Bevilacqua, A. Brunetti, G. D. Cascarano, A. Guerriero, F. Pesce, M. Moschetta, and L. Gesualdo, "A comparison between two semantic deep learning frameworks for the autosomal dominant polycystic kidney disease segmentation based on magnetic resonance images," BMC Med. Inf. Decis. Making 19(S9), 244 (2019).
6. J. L. Pezzaniti, D. Chenault, K. Gurton, and M. Felton, "Detection of obscured targets with IR polarimetric imaging," Proc. SPIE 9072, 90721D (2014).
7. J. S. Tyo, B. M. Ratliff, J. K. Boger, W. T. Black, D. L. Bowers, and M. P. Fetrow, "The effects of thermal equilibrium and contrast in LWIR polarimetric images," Opt. Express 15(23), 15161–15167 (2007).
8. S. Komatsu, A. Markman, A. Mahalanobis, K. Chen, and B. Javidi, "Three-dimensional integral imaging and object detection using long-wave infrared imaging," Appl. Opt. 56(9), D120–D126 (2017).
9. A. Markman and B. Javidi, "Learning in the dark: 3D integral imaging object recognition in very low illumination conditions using convolutional neural networks," OSA Conti. 1(2), 373–383 (2018).
10. D. Aloni, A. Stern, and B. Javidi, "Three-dimensional photon counting integral imaging reconstruction using penalized maximum likelihood expectation maximization," Opt. Express 19(20), 19681–19687 (2011).
11. X. Shen, A. Carnicer, and B. Javidi, "Three-dimensional polarimetric integral imaging under low illumination conditions," Opt. Lett. 44(13), 3230–3233 (2019).
12. B. Tavakoli, B. Javidi, and E. Watson, "Three dimensional visualization by photon counting computational integral imaging," Opt. Express 16(7), 4426–4436 (2008).
13. G. Lippmann, "Epreuves reversibles donnant la sensation du relief," J. Phys. 7(1), 821–825 (1908).

Bahram Javidi, University of Connecticut, Email: Bahram.Javidi@UConn.edu

14. N. Davies, M. McCormick, and L. Yang, “Three-dimensional imaging systems: a new development,” Appl. Opt. 27(21), 4520–4528 (1988).
15. H. Arimoto and B. Javidi, “Integral Three-dimensional Imaging with digital reconstruction,” Opt. Lett. 26(3), 157–159 (2001).
16. F. Okano, H. Hoshino, J. Arai, and I. Yuyama, “Real-time pickup method for a three-dimensional image based on integral photography,” Appl. Opt. 36(7), 1598–1603 (1997).
17. M. Martinez-Corral, A. Dorado, J. C. Barreiro, G. Saavedra, and B. Javidi, “Recent advances in the capture and display of macroscopic and microscopic 3D scenes by integral imaging,” Proc. IEEE 105(5), 825–836 (2017).
18. A. Stern and B. Javidi, “Three-dimensional image sensing and reconstruction with time-division multiplexed computational integral imaging,” Appl. Opt. 42(35), 7036–7042 (2003).
19. E. H. Adelson and J. R. Bergen, “The plenoptic function and the elements of early vision,” Computational Models of Visual Processing 1, 3–20 (1991).
20. J. Liu, D. Claus, T. Xu, T. Keßner, A. Herkommer, and W. Osten, “Light field endoscopy and its parametric description,” Opt. Lett. 42(9), 1804–1807 (2017).
21. G. Scrofani, J. Sola-Pikabea, A. Llavador, E. Sanchez-Ortiga, J. C. Barreiro, G. Saavedra, J. Garcia-Sucerquia, and M. Martinez-Corral, “FIMic: design for ultimate 3D-integral microscopy of in-vivo biological samples,” Biomed. Opt. Express 9(1), 335–346 (2018).
22. Pranav Wani, Kashif Usmani, Gokul Krishnan, Timothy O’Connor, and Bahram Javidi, "Lowlight object recognition by deep learning with passive three-dimensional integral imaging in visible and long wave infrared wavelengths," Optics Express 30, 1205-1218 (2022).
23. Kashif Usmani, Timothy O’Connor, Pranav Wani, and Bahram Javidi, “3D object detection through fog and occlusion: passive integral imaging vs active (LiDAR) sensing,” Optics Express, 31, 479-491 (January 2023).

What opportunities for training and professional development did the project provide?

Several PhD graduate students [Gokul Krishnan, Timothy O’Connor, Pranav Wani, KAshif Usmani] worked on this project. The students have been coauthors on publications related to the project [please see the list of publications].

The PhD students had multidisciplinary background including electrical engineering and computer science, signal processing, statistics, convolutional neural networks, deep learning, and optical imaging.

Timothy O'Connor was supported under a GAANN Fellowship from Education Department at no cost to the ONR project. He has since completed his PhD degree.

Pranav Wani, Kashif Usmain, and Gokul Krishnan have completed the PhD exams and are writing their PhD theses. They will graduate in 2024.

The students were involved in theoretical, experimental, and numerical simulations in this project. They worked together to build the experimental system, perform object recognition experiments, collect and reconstruct 3D data videos, develop the dedicated algorithms, participate in presentations about the project, write journal and conference papers, etc. Please see the list of publications for student's participation in scholarly activities.

We have collaborated with professors of computer science at University Jaume I in Castelon, Spain to investigate optimum approaches for object recognition with deep learning, at no cost to this ONR project. I collaborated with a professor from Ben Gurion University in Israel on Compressive imaging for defending deep convolutional neural networks from adversarial attacks at no cost to this ONR project.

How were the results disseminated to communities of interest?

The results of the project were published in the premiere Journals and conferences. Many of the journals [Optics Express] are high impact factor open access journals thus the results are widely disseminated. Several journal articles were top downloads of the journals, thus received wide publicity. Optics Express Journal has an Impact Factor of 3.6.

Also, the results have been presented at various conferences including the OSA (Optica) Imaging and Applied Optics Congress. We did not attend many conferences due to COVID-19. Below is a list of publications.

Journal Papers:

1. Gokul Krishnan, Yinuo Huang, Rakesh Joshi, Timothy O'connor, And Bahram Javidi, "Spatio-temporal continuous gesture recognition under degraded environments: Performance comparison between 3D Integral Imaging (InIm) and RGB-D sensors," Optics Express Journal, Vol. 29, No. 19, pp. 30937- 30951 (2021).

2. T. O'Connor, A. Markman, B. Javidi, "Overview of three-dimensional integral imaging-based object recognition in low illumination conditions with visible range image sensors," SN Applied Sciences, Springer Nature, 2:1724, 23 September 2020.

3. B. Javidi, A. Carnicer, J. Arai, T. Fujii, H. Hua, H. Liao, M. Martínez-corral, F. Pla, A. Stern, L. Waller, Q. H. Wang, G. Wetzstein, M. Yamaguchi, and H. Yamamoto, "Roadmap on 3D integral imaging: sensing, processing, and display," Optics Express, **28**(22), pp. 32266-32293

(2020). **Top Download of Optics Express**

4. Kashif Usmani, Timothy O'Connor, Pranav Wani, and Bahram Javidi, "3D object detection through fog and occlusion: passive integral imaging vs active (LiDAR) sensing," Optics Express, 31, 479-491 (January 2023).

5. Pranav Wani, Kashif Usmani, Gokul Krishnan, Timothy O'Connor, Bahram Javidi, "Lowlight object recognition by deep learning with passive three-dimensional integral imaging in visible and long wave infrared wavelengths," Optics Express 30, issue 2, 1205-1218 (January 2022).

6. Kashif Usmani, Timothy O'Connor, and Bahram Javidi, "Three-dimensional polarimetric image restoration in low light with deep residual learning and integral imaging," Opt. Express 29, 29505-29517 (2021)

7. Vladislav Kravets, Bahram Javidi, and Adrian Stern, "Compressive imaging for defending deep neural networks from adversarial attacks," Optics Letters 46, 1951-1954 (2021).

8. K. Usmani, G. Krishnan, T. O'Connor, and B. Javidi, "Deep learning polarimetric three-dimensional integral imaging object recognition in adverse environmental conditions," Optics Express, 29 (8), 12215-12228 (August 2021).

9. B. Javidi, F. Pla, J. M. Sotoca, X. Shen, P. Lattore Carmona, M. Martinez, R. Fernandez, G. Krishnan, "Fundamentals of Automated Human Gesture Recognition using 3D Optical Imaging: A Tutorial," Advances in Optics and Photonics, Volume 12, 2020. December 14, 2020. **[Top Download of Advances in Optics and Photonics]**

Conference Papers:

13. Gokul Krishnan, Yinuo Huang, Rakesh Joshi, Timothy O'Connor, and Bahram Javidi, "An overview of continuous gesture recognition performance comparison using three-dimensional integral imaging and RGB-D sensors," Imaging 2022 congress, Sponsored by Optica (OSA), 11-15 July 2022, Vancouver, Canada.
14. Vladislav Kravets, Bahram Javidi, Adrian Stern, "Compressive Sensing Methods for Defending Deep Learning 3D Classifiers," Imaging 2022 congress, Sponsored by Optica (OSA), 11-15 July 2022, Vancouver, Canada.
15. Kashif Usmani, Gokul Krishnan, Timothy O'Connor, and Bahram Javidi, "Overview of Object Detection in Low Light Using Deep Learning and Polarimetric Three-Dimensional Integral Imaging," Imaging 2022 congress, Sponsored by Optica (OSA), 11-15 July 2022,

Vancouver, Canada.

16. Gokul Krishnan, Rakesh Joshi, Timothy O'Connor, Filiberto Pla, Bahram Javidi, "An overview of hand gesture recognition in degraded environments using three-dimensional integral imaging and deep neural networks," OSA Imaging and Applied Optics Congress, Vancouver, Canada, July 19-22, 2021.
17. Pranav Wani, Kashif Usmani, Gokul Krishnan, Timothy O'Connor, Bahram Javidi, "Object Classification in Photon-Starved Conditions using 3D Integral Imaging: Performance Comparison Between Visible and Longwave Infrared Imaging," OSA Imaging and Applied Optics Congress, Vancouver, Canada, July 19-22, 2021.
18. Kashif Usmani, Timothy O'Connor, Peter Marasco, Bahram Javidi, "Visible and long-wave infrared imaging in degraded environments using three-dimensional polarimetric integral imaging," OSA Imaging and Applied Optics Congress, Vancouver, Canada, July 19-22, 2021.

INVITED SEMINARS

1. Optica (Optical Society of America) Student Chapter at University of Jaume I, Castellon, Spain, March 2023.
2. ICFO (The Institute of Photonic Sciences), Barelona, Spain, March 2023
3. Distinguished Lecture, Electrical & Computer Engineering Department, Northwestern University, Evanston, Illinois, May 2023.

Prizes and Awards for Prof. B. Javidi received during the Project Period:

1. *Awarded The Optical Society (OSA) Emmett Leith Medal, 2021*
2. *Named by The International Society for Optics and Photonics (SPIE) as a Luminary, August 2021*
3. *Fellow, American Association for the Advancement of Science (AAAS), 2013*

Other Recognition of Research:

The following papers were Top Download of journals:

Gokul Krishnan, Rakesh Joshi, Timothy O'Connor, Filiberto Pla, and Bahram Javidi, "Human Gesture Recognition under Degraded Environments using 3D-Integral Imaging and Deep Learning," Optics Express, 28, #13, 19711-19725 (2020). **[Top Download of Optics Express]; Over 16000 Downloads.**

B. Javidi, A. Carnicer, J. Arai, T. Fujii, H. Hua, H. Liao, M. Martínez-corral, F. Pla, A. Stern,

L. Waller, Q. H. Wang, G. Wetzstein, M. Yamaguchi, and H. Yamamoto, "Roadmap on 3D integral imaging: sensing, processing, and display," Optics Express, **28**(22), pp. 32266-32293 (2020). **Top Download of Optics Express**

Technology Transfer

None.

Bahram Javidi, University of Connecticut, Email: Bahram.Javidi@UConn.edu