# An Optimal RetinaNet Model For Automatic Satellite Image-Based Missile Site Detection

Ashwin Yadav,<sup>#,\*</sup> Kamal Jain,<sup>#</sup> Akshay Pandey,<sup>#</sup> Eshta Ranyal<sup>#</sup> and Joydeep Majumdar<sup>\$</sup>

<sup>#</sup>Indian Institute of Technology Roorkee, Roorkee, Uttarakhand – 247667, India <sup>§</sup>DRDO – Defence Research & Development Laboratory (DRDL), Kanchanbagh, Hyderabad-500 058, India

\*Email: ashwiny77@gmail.com

### ABSTRACT

Satellite image processing is a manually tedious job and offers scope for automation as part of the information extraction process from satellite images. The process of information extraction involves object detection and one of the challenges is ascertaining the minimum number of images required to train the deep learning model to achieve a certain minimum accuracy. To the best of the authors' knowledge, work in missile site detection is relatively limited, with an existing exploration of the latest one-shot detection methods, such as RetinaNet, being absent. This work proposes an optimal deep learning model based on the RetinaNet framework and training on a minimal dataset A comparative analysis with previous work paves the road for future research in one-shot methods and optimally trained models. As part of the study, the key findings are that an optimal training scheme based on a minimal training dataset is possible. This step enables a reduction in training time for the development of an optimal missile site detection model is concerned. One of the many techniques to determine the minimal number of training images required to train the object detection model is plotting the number of training images versus the mean average precision. The same is validated in our work. Further, a hybrid scheme based on the two-model concept is tested wherein one model prioritizes Recall while the other prioritizes Precision. Thus a combination of both models to detect a set of targets provides an optimal framework for object detection. Lastly, the study finds that the single-stage RetinaNet algorithm offers the advantage of balancing speed and accuracy over erstwhile twostage and other single-stage methods.

Keywords: Missile site; Satellite imagery; Machine learning; Deep learning; RetinaNet

### 1. INTRODUCTION

Missile sites are an important military asset of any nation and there is a need to establishing means to detect and identify them as part of a fully automated customized application. There is little work in the open domain on the subject.<sup>1-3</sup> Thus this area needs to be explored, and a framework of object detection developed for detecting these missile sites via satellite images. This step will reduce the time in image analysis while enabling vital information to be available to commanders at all levels in the battlefield. Figure 1 shows the typical Missile sites in existence. We can observe that even though the various patterns appear broadly similar, there are vital differences resulting in each site possessing a unique feature set. This aspect renders a specific challenge in object detection and is discussed in detail subsequently. The study's objective is to assist or better replace the image analyst in detecting various missile sites in satellite imagery. The algorithm or application should reduce the time of detection of the sites while providing a reasonably high degree of accuracy which is a challenging task considering the heterogeneous nature of the target in terms of the varying patterns as depicted in Fig. 1. Figure 1 illustrates only a few possible patterns from many others (refer to www. climateviewer.org).<sup>4</sup> The challenges are summarized below:

- (a) The heterogeneous nature of the patterns gives rise to unique features corresponding to each site. Thus the feature set is not finite making the application of traditional object detection algorithms such as Scale Invariant Feature Transform (SIFT) and Orient FAST and Rotate BRIEF (ORB)<sup>5-6</sup> complex.
- (b) It isn't easy to extract the complete feature set or manually prepare a standard feature map corresponding to the targets for use in a traditional Machine Learning (ML) framework.
- (c) Choosing a suitable algorithm that is fast but relatively accurate from amongst traditional methods, ML and Deep Learning (DL) for the object detection task.

The importance of the training process and data has been discussed by Lin et al.<sup>7</sup> in their work on aircraft detection using the YOLO-v3 algorithm. In the present work too, the focus is on a systematic training process to maximize accuracy. The paper endeavors to address the challenges highlighted above in the context of the problem of missile site detection. The main contribution of this paper is:

Received : 19 June 2022, Revised : 14 September 2022 Accepted : 30 September 2022, Online published : 1 November 2022

- (a) Application of the single-stage RetinaNet DL algorithm to the missile detection problem and comparative analysis with previous work. Only two-stage methods are currently in use in the missile site detection problem.<sup>1-3</sup>
- (b) Development of an optimal hybrid model scheme trained on the concept of minimal training dataset framework and

suffer from the same shortcomings as that of traditional methods in terms of the prerequisite of the full feature set description of the target. This aspect is brought out in the study by Mahony *et al.*<sup>8</sup> We are therefore encouraged to apply the DL-based methods considering the heterogeneous nature of missile site patterns and the comparative advantages of the method therein. Traditionally research has focused on two-stage DL





(k) Figure 1. Typical missile site structures.

its testing on high and low-resolution images. The organization of the paper is as follows:

- 1. Section 2 describes the proposed framework for object detection.
- 2. Section 3 discusses the results, comparison with existing work, and analysis.
- 3. Section 4 summarizes the study.

# 2. PROPOSED FRAMEWORK

# 2.1 Background

Mahony *et al.*<sup>8</sup> carried out a comparative analysis of traditional vs. DL methods for object detection. ML-based methods, such as Histogram of Gradients (HoG)<sup>9</sup> using Support Vector Machine (SVM) classifier post feature extraction, also

methods such as Region-based Convolution Neural Network RCNN.<sup>10</sup> Liu *et al.*<sup>3</sup> have applied two-stage methods such as

(1)





Faster RCNN and one-stage YOLO-v3<sup>11</sup> to the problem. More recent single-stage detection methods such as RetinaNet<sup>12</sup> have not yet been applied to solve the problem of missile detection. Figure 2 illustrates the difference in architecture between the two-stage and one-stage methods. Thus in the current work, we apply the RetinaNet algorithm with its inherent advantages of balance between speed and accuracy compared to erstwhile single-stage and existing two-stage methods.

# 2.2 Methodology and Data Set Description

Figure 3 illustrates the methodology. The present work applies the RetinaNet algorithm<sup>12</sup> based on Fizyr's python code,<sup>16</sup> employing a pre-trained Pascal Visual Object Classes (VOC)<sup>17</sup> model applied in a transfer learning framework to a customized dataset. Ranyal *et al.*<sup>18</sup> use the RetinaNet algorithm in their work on detection of rail fasteners. The current work aims to train the model on the minimum dataset while maximizing the Mean Average Precision (mAP). The images used for training are from Google Earth, a bonafide source.<sup>19</sup> The training step is executed iteratively with a gradual increase in the mAP using the model with the highest mAP from the previous stage. A 70:30 ratio of training to validation images is maintained.

improved results compared to the individual models. The blue model provides higher recall while the red model provides higher precision (refer to Tables 4-5). The study explores a hybrid approach by combining both models for object detection on a set of images to overcome the limitations of the individual models. Results are illustrated and discussed ahead accordingly. The study compares results achieved with previous work.<sup>1,3</sup> The website www. climateviewer.org<sup>4</sup> provides the locations of the missile sites with images downloaded subsequently from Google Earth Google Earth offers a high-resolution image (<=43 cm Ground sample distance (GSD)) set compared to other costlier commercial satellite systems. 35% of the authors in the survey paper<sup>19</sup> have verified Google Earth as a bonafide resource. Testing on Sentinel 10 m GSD images sourced from the GEE platform<sup>20</sup> validates whether the model has learned the coarse resolution feature set. The GEE framework enables cropping of the image to user requirements, unlike standard satellite images downloaded from other sources. Figure 4 represents the images used for training the model, while Fig. 5 shows the GEE platform image used for testing purposes.



Figure 3. Methodology: Missile site detection.

Unseen images of missile sites not contained within the training or validation data form part of the testing process. Based on the results obtained, the model is tested on 10 m resolution Sentinel 2 images obtained from Google Earth Engine (GEE) framework.<sup>20</sup> The objective is to try whether the model can learn a feature set corresponding to a coarser (10 m compared to 2-3 m of Google earth) resolution. Testing on two models, one trained on a set of limited images (up to 210) and the second on an augmented dataset (up to 300) obtained using horizontal and vertical rotation operations, is carried out. These models, termed the blue and red models, are illustrated in Fig. 8. The study explores the possibility of combining these two models providing



Figure 4. Sample google earth image used for training.



Figure 5. Sample 10 m GSD Sentinel image (GEE platform) for testing.

## 2.3 Execution

Based on the methodology illustrated vide Fig. 3 following are the steps.

- (a) The testing process uses images from 30 missile sites<sup>4</sup> in the ratio of 70:30 training versus validation data format with a set of 7 images per site (i.e., a total of 210). Figure 6 depicts a sample image. The images are obtained at varying. Along with the inherent Feature Pyramid Network (FPN) feature of the RetinaNet model, this shall enable a fair degree of scale invariance. The Neural Network model used in the RetinaNet model is the ResNet50 architecture by Kaming He, *et al.*<sup>21</sup>
- (b) The tool of data augmentation in the form of horizontal and vertical rotation of selected images within the dataset



Figure 6. Sample set of training images per site.

enables an increase in the number of training images in steps of 30 (horizontal and vertical flip of 15 random pictures at a time) from 210 to a maximum of 300. Table 1 and Fig. 7 depict the number of images versus the maximum mAP achieved at each stage.

Table 1. Number of Images vs mAP achieved				
Number of training cum validation images	mAP(fraction)			
210	.5547			
240	.6047			
270	.6620			
300	.6701			

. .

- (c) The objective was to answer the question "What is the minimum number of images to train your model effectively." (Effectively implies a reasonably high mAP to detect missile sites in an unseen image.)
- (d) The next step is testing the models with varying mAPs on both categories of images (Refer to Fig. 8 and 9, respectively), with testing on coarse resolution only demonstrative. Results are based on testing on Google Earth Images (Refer to Fig. 7).
- (e) The last step is a comparison with previous work.<sup>1,3</sup>

# 3. RESULTS AND ANALYSIS

## 3.1 Results

Results are as presented below:

- (a) Results obtained vide Table1 and Fig. 7 demonstrates that it is possible to train the model with a minimum number of images while trying to maximize the mAP.
- (b) An unknown set of Google earth images test the Red model. The model detected the missile site in an image corresponding to an area of 8.9 x 5.2 on the ground. However, it could not detect the missile site in a Sentinel 2 10 m GEE image.
- (c) The Blue model was tested on the Sentinel 210 m resolution GEE platform RGB image. It can detect the missile site in an area of 6 x 8 km.
- (d) A random unknown set of 10 images obtained from the balance of 30 sites (excluding the ones used for training) of the total original 60 missile sites test the Blue and Red. Table 2 presents the Standard metrics of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) corresponding to each of the models in terms of numbers obtained. The values of metrics are the respective total based on cumulative testing on the ten image datasets for each model.
- (e) The hybrid schemes nos 1 and 2 in Table 2 above represent the optimum solution. In the Hybrid 1 model the blue model processes image sequence, 6 while red model processes the balance thereby achieving an overall balance between recall and precision. In the Hybrid 2 model image sequences, 6 and 9 are processed by the blue model with the balance being processed by the red model to achieve higher recall. (e.g. military target detection)

#### 3.2 Comparison with Previous Work

A comparison of the RetinaNet model with the work by



Figure 7. Images vs mAP plot.



Figure 8. Sample Red model detection in 8.9 x 5.2 km area, Google earth image.



Figure 9. Sample Blue model detection in 7 x 8 km area, GEE platform image.

Marcum *et al*<sup>1</sup> and Feng Liu<sup>3</sup> is carried out vide Table 3. The objective of this work is similar to<sup>1,3</sup> i.e., to aid or better substitute the image analyst in the detection of missile sites over large areas in the shortest possible time. Following important points in the context of the two existing works merit attention:

- (a) The vintage of the Marcum's paper<sup>1</sup> is October 2017 and RetinaNet was developed in February 2018. The paper only gives the base architecture of the network e.g., Resnet50 but does not describe the underlying Deep CNN framework in greater detail.
- (b) Marcum's approach<sup>1</sup> is based on a 4 GPU hardware architecture which is a single GPU in our case and hence

we compensate for the same accordingly.

- (c) Liu's paper<sup>3</sup> is of April 2019 vintage and applies the twostage faster R-CNN and one-stage YOLO-v3 methods to the problem. It is based on a single GPU-based architecture.
- (d) The work by Marcum<sup>1</sup> is a more comprehensive study, going into depth regarding challenges of the missile site detection and executing a thorough comparison with a visual Broad Area Search (BAS) carried out by image analysts (IA).

The model once fully trained and tested offline requires no manual verification of the results or corrections therein on the application.

#### **3.3 Deductions**

Based on the results and comparison following are the relevant deductions.

- (a) The RetinaNet algorithm as a DL model provides an effective target detection model in terms of both speed and accuracy compared to two-stage DL and traditional methods.
- (b) The red model is effective in detecting targets in highresolution imagery with a fair degree of scale invariance but is unable to detect targets of the coarse resolution GEE imagery type and reports improved precision compared to the blue model.
- (c) The blue model is effective in detecting the target in low-resolution GEE-based imagery compared to the red

Model	TPs	TNs	FPs	FNs	Recall	Precision	F1 Score
Red	7	5	4	3	.7	.64	.67
Blue	8	2	17	2	.8	.32	.46
Hybrid 1*	7	5	3	3	.7	.7	.70
Hybrid 2**	8	5	7	2	.8	.53	.64

\*: Image sequence no 6 is processed by the blue model and all others by the red model. This model provides a balance of precision and recall.

\*\*: Image sequence no 6 and 9 is processed by the blue model and all others by the red model. This model provides high recall at cost of lower precision.

Parameter	Marcum's work <sup>1</sup>	Feng Liu et al. <sup>3</sup>	Our Method
Algorithm/Method	Not specified. Only Classification	YOLO-v3, Faster R-CNN	Retina Net. Classification and localization
Architecture	CaffeNet,GoogLeNet,ResNet-50, ResNet-101	VGG-16,ResNet101	ResNet50
Training	The model requires training on a set of counter or negative examples. This enables a reduction in false positives. Time taken for training the model is 4.8 hrs	Based on pre-trained VGG-16 and ResNet101 architectures. Training time data not specified.	Training does not require negative example- based training as the focal loss feature circumvents this problem. Training time for our model is 30 minutes based on a training cycle of 25 epochs with 10 steps each with a training image set of 210 images (165 training and 45 validation)
Input image size	The Maximum permitted is 227x227	Not specified	830x1330
Spatial resolution (GSD)	1m	.156 m	As per Google Earth imagery (up to 43 cm)
Processing speed	<ul> <li>42.3 minutes. (BAS based on 88,640 sq km in the paper<sup>1</sup> using</li> <li>4 GPUs detected by 10 IAs simultaneously)</li> </ul>	3 hrs including human verification to search an area of 6000 sq km based on a search of a total of 40000 images. It has been implemented on a single system	-5.6 min * for unassisted fully automatic detection(by the algorithm itself) and hence nearly 8 times faster [(42.3/5.6)=7.55] compared to <sup>1</sup> -15.16 min** for unassisted fully automatic detection(by the algorithm itself) and hence nearly 12 times faster [(3x60)/15.16=11.87] compared to <sup>3</sup>
Accuracy and other metrics	Up to 99.4% based on a ResNet50 architecture based on a training dataset of 893,376 images post augmentation	Up to 91.02% precision was achieved using the Faster R-CNN based on a ResNet101 architecture. corresponding to a training dataset of 10443 images (post data augmentation)	Maximum Recall is 80 % and precision is 70%. The model is optimal considering training on a dataset size of 300 images post dataset augmentation(2988 less than <sup>1</sup> and 35 less than <sup>3</sup> )

### Table 3. Comparison with previous work.

Note: \*

Our algorithm takes 7s on average to search an area of dimension 8.9 km x 5.2 km on a single system.

88,640 sq km $\Box$ 1916 x (8.9 km x 5.2 km). Thus search time is =(1916x7)/60=223.53 min. on a single GPU-based system (as ours)

For parity, we apply the above in a similar setting on a 4 GPU system and substitute the 10 IAs with 10 parallel such systems. Hence estimated time=223.53/(4x10)=5.6 min.

\*\*:

Our algorithm takes 7 s on average to search an area of dimension 8.9 km x 5.2 km on a single system.

6000 sq km $\square$  130 x (8.9 km x 5.2 km). Thus search time is = (130x7)/60=15.16 min. on a single GPU-based system (as ours).

model which fails in this regard. But for the Google Earthbased test images, the blue model provides lower precision compared to the red model, which is suitably overcome via the hybrid scheme.

- (d) An effective target detection framework should comprise a combination of models to detect targets sourced from the imagery of varying spatial resolution. Fig 9 illustrates that a model trained on a limited number of high-resolution images can learn coarse-resolution target features as well. However, the contrary may not be true i.e. a model trained on coarse resolution images is unlikely to detect targets sourced from high-resolution imagery.
- (e) From Table 1 and Fig. 7 it is observed that increasing the training dataset indefinitely is not likely to result in commensurate improvement in mAP with the point of saturation being reached at a certain point of training (at approximately 270 images) as indicated by the nonlinear portion of the curve.
- (f) Table 2 illustrates the performance of the red and blue models. Considering the minimum image training aspect, the performance is satisfactory as the recall value of both models is greater than the individual mAPs. The Recall parameter is important in the case of military target detection wherein no target should remain undetected, even at the cost of lower precision. The Red model yields higher precision with fewer false positives while the Blue model provides a higher recall but lower precision. Thus increasing the training dataset by data augmentation has limitations and ultimately the variety of training images has to be increased for the model to learn a sufficient feature set of the missile target.
- (g) Table 2 also illustrates the performance of two hybrid models based on the Blue and Red models respectively by combining the advantages of both. The Hybrid 2 model is optimal for the military framework wherein the objective is to maximize recall. Even though the recall value in the hybrid 2 model is the same as that of the blue model however the overall precision has improved. The Hybrid 1 model may be used when a balance between recall and precision.

- (h) Overall the advantage of the DL-based methodology in form of an automated model for object detection is seen. Two of the authors have applied the concept to the crop identification problem using DL.<sup>22</sup> The advantages of speed based on area coverage are evident. From table 8 it is seen that the RetinaNet algorithm scores over Marcum et al<sup>1</sup> and FengLiu et al<sup>3</sup> in terms of the processing speed while using less training data. The advantage of the Retina Net algorithm compared to the Faster R-CNN and YOLOv3 used in<sup>3</sup> is evident. Advantages in terms of training and inference time of the one-stage detectors are accordingly validated in our work.
- (i) The work is demonstrative and once a sufficiently accurate model is obtained it can be used to substitute the image analyst whose task shifts from detection to that of validation. The valuable man hours saved will help in improving the overall efficiency within the image processing framework.

## 4. CONCLUSION

To the best of the knowledge of the authors, this is the only study on the application of a single-stage DL model to the problem of missile site detection. An automatic detection framework based on the RetinaNet algorithm has been studied and a comparison carried out with existing work in the domain. Automated detection or information extraction framework based on satellite imagery greatly simplifies the task of the image analyst enabling efficiency within the image analysis workflow. The advantage of speed in processing and validation of the minimum training dataset concept has been carried out. The scope for the development of an optimal hybrid framework based on the blue and red models applied to the detection problem has also been explored using the minimal dataset training model concept. This enables a reduction in the training time while also improving the mAP to a maximum limit. It may be noted that training based on data augmentation to a limit of 300 images has been carried out which is symbolic and may be increased or variation in the number of original training images be increased from 210. The approach to determining the minimum number of training images based on the graph is one of the possible methods among others. The combination of models used may be increased from 2 based on the feature set learned by the various models developed and tested therein. This scope will be defined by the resolution, quality, and nature of the unknown target dataset. The study has focused on the missile site detection problem however, the results obtained may be well applied to similar target datasets wherein a heterogeneous and complex feature set exists.

# ACKNOWLEDGMENT

The authors are extremely grateful to Col Satrh Harshvardhan who is an experienced Defence Analyst and without whose guidance this work would have been incomplete. The user perspective provided by him enabled a wholistic approach to be adopted as part of the work

#### REFERENCES

1. Marcum, Richard A.; Davis, Curt H.; Scott, Grant J. & Tyler, W. Nivin. Rapid broad area search and detection

of Chinese surface-to-air missile sites using deep convolutional neural networks. *J. Appl Remote Sens.*, 2017, **11**(4) 042614. doi: 10.1117/1.JRS.11.042614

- Cannaday II, A.B.; Davis, C.H.; Scott, G.J.; Ruprecht, B. & Anderson, D.T. Broad area search and detection of Surface-to-Air missile sites using spatial fusion of component object detections from deep neural networks. *EEE J. of Selected Topics in Appl. Earth Obs. and Remote Sens.*, 2020, 13, 4728-4737, doi: 10.1109/JSTARS.2020.3015662
- Feng, Liu; Zhu, Jihong; Wang; Wufan; Kuang & Minchi. Surface-to-air missile sites detection agent with remote sensing images. *Sci. China Inf. Sci.* 2021,64. doi: 10.1007/s11432-019-9920-2
- 4. Climate viewer maps. Available at http://climateviewer. org/history-and science/government/maps/surface-to-airmissile-sites-worldwide/
- 5. Lowe, D.G. Distinctive image features from scaleinvariant keypoints. *Int. J. of Comp. Vision*, 2004, **60**, 91–110.

doi: 10.1023/B:VISI.0000029664.99615.94

- Rublee, E.; Rabaud, V.; Konolige, K. & Bradski G. ORB: An efficient alternative to SIFT or SURF. Int. Conf. on Computer Vision, 2011, 2564-2571. doi: 10.1109/ICCV.2011.6126544.
- Lin, Y.C.; & Chen, W.D. Automatic aircraft detection in very-high-resolution satellite imagery using a YOLOv3based process. *J. of Appl. Remote Sens.*, 2021,15(1), 018502.

doi: 10.1117/1.JRS.15.018502

 Azam, Basim; Khan, Muhammad Jaleed; Bhatti, Farrukh Aziz; Rahman, M. Maud Abdur; Hussain, Syed Fawad; Hashmi, Ali Javed & Khurshid, Khurram. Aircraft detection in satellite imagery using deep learning-based object detectors.Microprocessors and Microsystems, 2022, 94.

doi: 10.1007/978-3-030-17795-9

- Dalal, N. & Triggs, B. Histograms of oriented gradients for human detection. 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), 2005, 1, 886-893. doi: 10.1109/CVPR.2005.177
- Girshick, R.; Donahue, J.; Darrell, T. & Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. IEEE conference on computer vision and pattern recognition, 2014, 580-587 doi: 10.1109/CVPR.2014.81.
- Redmon, J; Divvala, S; Girshick, R. & Farhadi, A. You only look once: Unified, Real-Time Object Detection. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, 779-788 doi: 10.1109/CVPR.2016.91.
- Lin T.Y.; Goyal, P.; Girshick, R.; He, K. & Dollár, P. Focal loss for dense object detection. IEEE transactions on pattern analysis and machine intelligence, 2020, 42(2), 318-327.

doi: 10.1109/TPAMI.2018.2858826.

- Carranza-García; Manuel; Jesús Torres-Mateo; Pedro Lara-Benítez & Jorge, García-Gutiérrez. On the performance of one-stage and two-stage object detectors in autonomous vehicles using camera data. *Remote Sens.*, 2021, 13(1), 89. doi: 10.3390/rs13010089
- Jiao, L. *et al.* A Survey of Deep Learning-Based Object Detection. *IEEE Access*, 2019, 7, 128837-128868. doi: 10.1109/ACCESS.2019.2939201.
- Liu, W. *et al.* SSD: Single Shot MultiBox Detector. In Leibe, B., Matas, J., Sebe, N., Welling, M. (eds) Computer Vision. ECCV 2016. Lecture Notes in Computer Science, 9905. Springer, Cham. doi: 10.1007/978-3-319-46448-0 2
- 16. https://github.com/fizyr/keras-retinanet doi: 10.5281/zenodo.3250670
- Everingham, Mark, Luc Van Gool, Christopher, K.I. Williams, John M. Winn and Zisserman, Andrew. The Pascal Visual Object Classes (VOC) Challenge. *Int. J. Comp. Vision*, 2009, 88, 303-338. doi: 10.1007/s11263-009-0275-4
- Ranyal, E. & Jain, K. Detection of rail fasteners from aerial images using deep convolution neural networks. In IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium (pp. 2217-2220). doi: 10.1109/IGARSS39084.2020.9323596
- Kanjir, Urška; Greidanus, Harm; Oštir, Krištof. Vessel detection and classification from spaceborne optical images: A literature survey. *Remote Sens. of Env.*, 2018, 207,1-26.

doi: 10.1016/j.rse.2017.12.033

- Gorelick, Noel; Hancher, Matt; Dixon, Mike; Ilyushchenko, Simon; Thau, David; Moore, Rebecca. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. of Env*, 2017, 202, ,18-27. doi: 10.1016/j.rse.2017.06.031
- He, K. ; Zhang, X.; Ren, S. & Sun, J. Deep residual learning for image recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016, pp. 770-778,

doi: 10.1109/CVPR.2016.90

 Pandey, Akshay & Jain, Kamal. An intelligent system for crop identification and classification from UAV images using conjugated dense convolutional neural network. *Comp. and Electr. in Agri.*,2021, **192**. doi: 10.1016/j.compag.2021.106543

# CONTRIBUTORS

**Mr** Ashwin Yadav holds dual master degrees in Electrical Engineering and Geomatics Engineering from IIT-Kanpur and IIT-Roorkee respectively. He is a practicing Communications Engineer with interest in signal processing and has two conference papers and a chapter publication in the field of target tracking in wireless sensor networks. His current interest spans image processing, object detection and change detection applied to remote sensing, thermal and UAV imagery.

As the primary author of the paper, he has carried out the literature review including study of traditional, machine learning and Deep Learning approaches based on which the methodology and application of the proposed method has been carried out. Subsequently execution of the proposed method in python programming language and testing has been carried out followed by analysis under guidance of the supervisor alongwith contribution of other co authors.

**Prof. Kamal Jain** is working in the field of Geomatics and its application since 1982. Presently Dr. Kamal Jain is a Professor (HAG) of Geomatics in Department of Civil Engineering, IIT Roorkee. He performs a broad spectrum of functions, not merely teaching but also at the forefront of cutting-edge research. He has research experience of nearly 39 years. He is working in education and research in the field of photogrammetry and remote sensing especially his efforts in field of UAV photogrammetry. He has published many papers in international peer-reviewed journals (>100) and guided more than 150 post-graduate dissertations and 38 PhD thesis. He has authored two books: One on Web GIS and another on UAV. Dr. Kamal Jain has applied reverse engineering and developed many new instruments. He has developed RTK drones, Offset GPS, etc. He has 13 patents and 45 copyrights.

He has been instrumental in providing the fundamental principles and approach for undertaking the work as well as continuous guidance and analysis at each stage.

**Mr Akshay Pandey** obtained his Bachelor's and Master's degree in Computer science & Engineering and is currently doing his PhD at IIT Roorkee. His research interests include the development of remote sensing and Web GIS-based techniques for agriculture applications. He is using a UAV (Drone), Spectroradiometer for his research. He is also having experience in Geospatial Tool, Web Development, Software Development, Networking, and Server Management. He was involved in fundamental structuring of the paper from

a technical point of view and analysis of the results alongwith the primary author Ashwin Yadav.

**Ms Eshta Ranyal** is currently a PhD scholar at the Indian institute of technology, Roorkee. She is a multidisciplinarian with Masters in computer application and engaged in the research of vision-based deep learning of multi-source remote sensing data in the field of civil engineering. Her current research focuses on artificial intelligenceassisted vision-based pavement health monitoring employing remote sensors integrated with aerial and ground platforms.

She was involved in the review of technical aspects of the paper including the fundamental algorithm employed and comparative analysis with previous work in the domain.

**Mr Joydeep Majumdar** is a practicing Reliability and Maintenance Engineer and researches on Statistical Data Analysis, Cyber Physical Systems & impact on future technologies especially in the area of military technologies. He holds masters in Reliability Engineering from IIT Kharagpur and is presently involved in providing Reliability Coverage to Single Shot Systems in DRDL, DRDO.

As co-author, Joydeep has co-founded the statistical reasoning and rigour to the data analysis required in the change detection problem along with the author. Further, he suggested the future road map of research in change detection culminating in object detection using machine learning.