INTERNATIONAL JOURNAL ON SMART SENSING AND INTELLIGENT SYSTEMS

Research Article | DOI: 10.2478/ijssis-2026-0039 Issue 1 | Vol. 19 (2026)

Enhancing Deep Learning-Based Human Detection in Flooded Areas Via Dynamic Lighting-Adaptive Image Enhancement

Pallavi Nehete[1,2], Anupkumar Bongale[3], Shrinivas Shrikande[4], Pallavi Mulmule[5] and Deepak Dharrao[1,*]

[1]Department of Computer Science and Engineering, Symbiosis Institute of Technology, Pune Campus, Symbiosis International (Deemed University), Lavale, Pune 412115, Maharashtra, India

[2]Department of Computer Science and Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune 411038, India

[3]Department of Artificial Intelligence and Machine Learning, Symbiosis Institute of Technology, Pune Campus, Symbiosis International (Deemed University), Lavale, Pune 412115, Maharashtra, India

[4]Department of Computer Science and Engineering, S. B. Patil College of Engineering, Indapur 413106, India

[5]Department of Electronics and Communication Engineering, School of Engineering and Technology, DES Pune University, Pune 411004, Maharashtra, India

*E-mail: deepak.dharrao@sitpune.edu.in

Received for publication August 04, 2025.

Abstract

Flood zones pose significant challenges for search and rescue (SAR) operations due to poor visibility, noise, reflections, and occlusions caused by adverse weather and low-light conditions. These factors degrade object detection performance. To address this, the proposed system integrates YOLOv11 with a Retinex-based image enhancement method. The Retinex theory separates illumination from reflectance, improving detail extraction in low-light images. Two publicly available datasets, the Roadway Flooding Image Dataset and the Human Detection in Floods Dataset covering varied flood scenarios, lighting, and occlusions, are used. Results show that Retinex enhancement significantly improves detection accuracy: YOLOv11 reached 95.72%, outperforming baselines by up to 9%–10% in recall and 5%–6% in F1-score. Enhanced models also showed 7%–10% robustness gains on occluded and reflection-heavy images. This demonstrates that the Retinex-enhanced YOLOv11 model provides reliable, real-time human detection in flood environments, offering valuable tools for life-saving SAR efforts.

Keywords

Retinex-based image enhancement, YOLOv11, illumination-invariant detection, flood surveillance, Low-Light Imaging

I. Introduction

Floods are one of the most destructive natural calamities, making it very difficult and time-consuming to find survivors at the time of emergencies. The floods that occur annually in India cause massive numbers of human casualties, displacement, and devastation of infrastructure and agriculture, and therefore, are more endangering to the affected communities [1]. The investigators of the risks of urban floods, for example, in the Poiser River Basin, in Mumbai, have noted the transformational nature of artificial intelligence (AI) and machine learning (ML) technologies in enhancing human detection and optimizing rescue efforts [2].

Having incorporated deep learning (DL) models, and specifically Convolutional Neural Networks (CNNs), autonomous drones and cars can now detect human survivors in flooded areas with high accuracy. The combination of AI, geographic information systems (GIS), and real-time data analytics has significantly improved the efficiency and speed of the localization of the victim, thus decreasing the time of the rescue response, and increasing the survival rates [3,4].

The recent development of ML and AI has automated the human object detection capacities to disaster response systems to detect objects faster and more accurately in chaotic flood environments. CNNs constitute the foundation of modern DL designs, and they are good at handling the complicated visual data in flood-impacted landscapes [5]. Adaptive learning enabled through AI enables such models to continuously refine their accuracy through time and in response to dynamic inputs to the environment. As an illustration, research proposing DL as a strategy to generate effective human and machine-marked ground-truth databases demonstrated a detection rate of over 96%, highlighting the strength of the model in flood-based human detection applications [6].

Such techniques as decision trees and CNNs have proven to be very efficient ML and satellite-based flood mapping techniques that can efficiently detect the extent of floods in short periods of time [7,8]. A comparative analysis of the ML strategies, such as support vector machines (SVM), multi-layer perceptrons (MLP), and Deep Convolutional Neural Networks (DCNN), showed that semi-supervised domain adaptation (SSDA) is more effective as a flood detector because it has high Area Under the Curve (AUC) and F1 scores and minimal labeled data are required [9]. These results suggest the increased opportunities of hybrid ML methods to enhance the detection strength in data-sparse and high-variability environments that are characteristic of floods.

The development of the object detection models, that is, YOLOv3 to YOLOv8, has improved the accuracy and effectiveness of flood detection and human detection activities to a great extent [10,11,12]. Architectures built on YOLO have been shown to be useful in detecting human beings, vehicles, and infrastructure in flood images that are taken by vehicles and aircraft, and achieve higher performance than traditional architectures due to their faster processing speed and higher accuracy. Practices like transfer learning, semi-supervised training, attention-based feature extraction, and others have also enabled the model to be more adaptable to the variety of flood settings [13,14]. More recent attempts using YOLOv8-seg models and analysis of object-based search and rescue (SAR) images have enhanced reliability of flood inundation mapping and change detection, providing a more spatially coherent representation of flood dynamics [15]. All these studies point to the great necessity of sophisticated vision-based AI systems to help with the fast and precise management of flood disasters [16].

Nevertheless, the ability of the human eye to sense under harsh visual circumstances, especially in low light, hiding, and blurring due to rough water and bad weather, is still an unsolved problem. Unstable light conditions and water reflections in the real-life flood rescue scenario tend to reduce the quality of the image, leading to a false alarm or overlooked survivor [17,18]. Recent studies have shown that Retinex-based illumination enhancement can enhance image brightness and color uniformity considerably and that noise in a dark setting can be reduced [19]. Together with the modern DL detection models, such improvements can significantly boost the reliability of detection in visually challenging flood situations [20].

The proposed system presents a comprehensive framework that integrates Retinex-based illumination enhancement with the advanced YOLOv11 object detection architecture [21,22] to improve human detection performance in flooded environments under dynamic lighting conditions. The suggested solution actively responds to changing lighting and visibility to enhance real-time human detection to flood-prone regions [23]. Being the most optimized in feature extraction, polished attention, and advanced real-time inference, YOLOv11 is deployed to guarantee strong detection performance in unfavorable environmental factors [24,25]. This framework overcomes the shortcomings of the earlier versions of YOLO when used with low-light and low-quality photographs because it uses a physics-based Retinex enhancement model and adaptive attention feature aggregation [26,27].

The main goal of this project is to test and improve the quality of detection of DL models in the unfavorable cases, namely, blur, occlusion, and unstable light [28,29]. Precision, recall, confusion matrix, mean average precision (mAP), and intersection over union (IoU) are evaluation metrics that are used to assess comprehensively the performance improvements [30,31]. This paper will provide a robust, high-precision framework of real-time human detection and rescue support with flood situations, which will make a significant contribution to the field of disaster management and humanitarian technology by integrating both illumination-adaptive preprocessing and the advanced architecture of YOLOv11 [32,33].

II. Methodology

The proposed study aims at offering a unified framework to improve human detection ability in flood regions through the combination of dynamic, lighting-adaptive image enhancement schemes with state-of-the-art YOLO-based object detectors. While the proposed model emphasizes illumination correction and contrast enhancement through Retinex-based preprocessing, it does not explicitly implement occlusion-sensitive training strategies within the detection pipeline. The proposed methodology as shown in Figure 1 is categorized into five primary phases such as dataset preparation, image enhancement, model training, evaluation, and comparative analysis.

The popularity of the YOLO models in real-time applications stems from their exceptional speed and accuracy. This ability is essential for teams that do disaster response and autonomous systems. Figure 2 is an illustration of the architecture of the YOLO model.

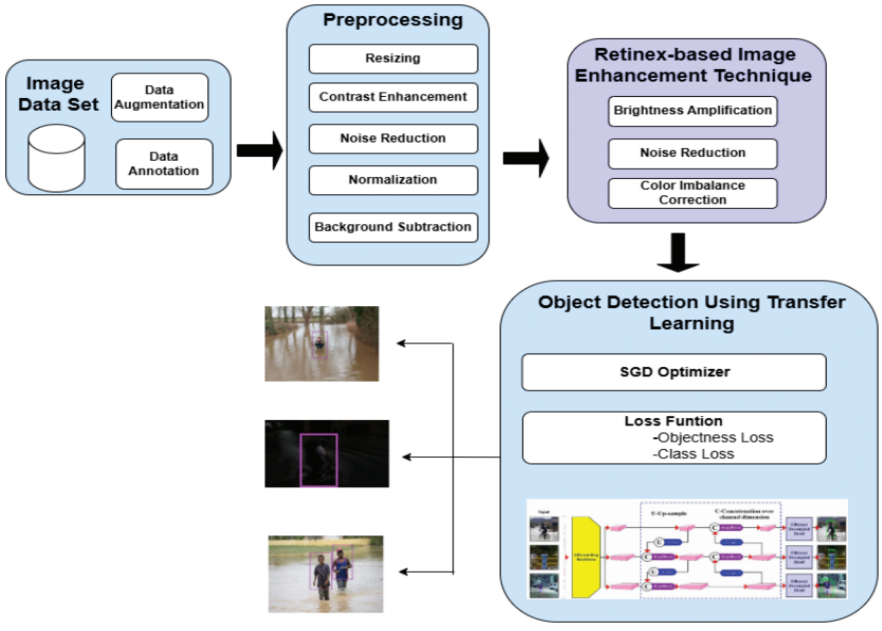

Figure 1: Proposed system architecture of enhanced object detection with Retinex and YOLOv11. SGD, stochastic gradient descent.

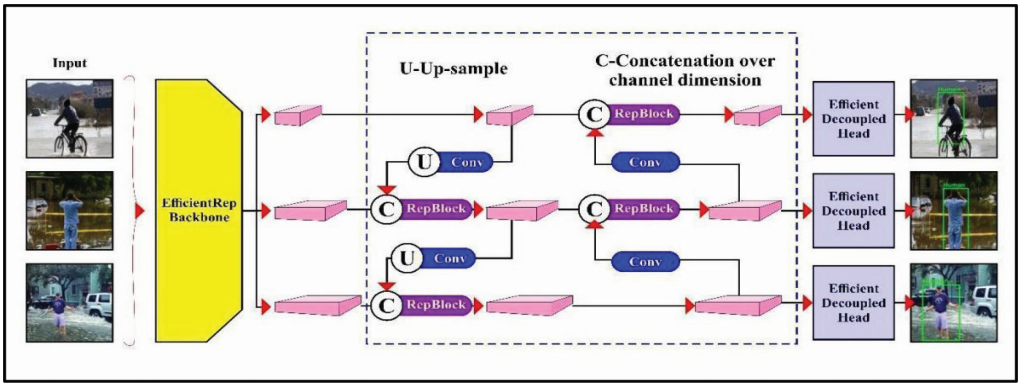

Figure 2: Schematic structure of YOLO model used in the study.

a. Dataset preparation

Two datasets were utilized to ensure diversity in lighting and environmental conditions:

a) Roadway flooding image dataset (Kaggle)

This publicly available dataset contains images depicting various flood scenarios on roadways, captured under different illuminations and weather conditions. This dataset consists of 441 annotated roadway flood images in varying lightning conditions. Examples of human detection under flooding conditions can be seen in Figure 3, which reflect the variety of situations comprising varying water levels, different views, and various weather conditions.

b) Human Detection in Flood Dataset (Roboflow)

This dataset is publicly available on Roboflow and consists of annotated images focusing specifically on humans in flood-affected areas, including challenging low-light and complex backgrounds. This dataset contains annotated flood scenarios with stranded humans, occluded cases, and varying illumination conditions. The dataset contains a total 1105 images further divided for training purposes. Splitting data: Each dataset is divided into 70% training, 15% validation, and 15% testing subsets.

Figure 4 shows sample images from the Roboflow database having diverse scenarios, including varying water levels, different angles of view, and different weather conditions.

Both datasets are accompanied by annotation in YOLO format for the proposed models. The annotations comprise exact positions of bounding boxes and class probabilities, which make the task of object detection fast. The labels are configured to be equivalent to “person,” which is the focus of the mission of the dataset, whose visual recognition is of individuals during floods. By means of training set, the detection algorithms will learn how to detect and mark people in flood environments, regardless of the light predominant in the environment. After the training the test set is used as a standard for the models to compare performance. This test determines the limit to which the models can use what they have learned, to apply it to unseen circumstances and the ability to identify different individual images in various flood

Figure 3: Sample images from roadway-flooding-image-dataset (Kaggel) [34] showing diverse scenarios, including varying water levels, different angles of view, and different weather conditions.

Figure 4: Sample images from human detection in flood dataset (Roboflow) [35].

circumstances. Furthermore, the parameters were optimized, and training performance of the models was also tested using the validation set as a reference. Using this constantly improving method boosts the model's robustness and reliability, enhancing the model's capacity to identify people in affected flood regions.

b. Image preprocessing

The proposed image preprocessing framework supports visibility of an image, removes noise, and enhances consistency of illumination prior to the model training. It combines various adaptive approaches, each dealing with a different factor of degradation into a single enhancement pipeline. It is illustrated in the following manner:

a) Image Resizing and Normalization:

All the images were downsized to a standard size of 640 × 640 pixels with the aspect ratio kept appropriate to the YOLO input format. The values of pixel intensities were scaled between [0, 1] to stabilize gradient changes and enhance convergence in training. This step makes sure that the model gets data of consistent scale, irrespective of the source differences in image size or resolution. Besides individual training, a mixed dataset (combining both sources) was also developed to test the model generalization in varying conditions of floods.

b) Adaptive median filtering (AMF) noise reduction:

Images captured in uncontrolled outdoor atmospheres often contain impulse noise or salt-and-pepper artifacts due to sensor errors, transmission interference, or water reflections. Such noise can severely distort object boundaries and can affect feature extraction by the detection model. To overcome this, AMF is applied on RGB channels individually. AMF also possesses dynamic window size as compared with a standard median filter, which possesses a fixed window size. This capability of adaptiveness guarantees high levels of noise reduction in corrupted regions and maintains edges and finer textures in uncorrupted regions. Mathematical equation for AMF is shown in Eq. (1) as follows:

$$Z_{\min} = \min\left(W_{xy}\right), Z_{\text{med}} = \text{median}\left(W_{xy}\right), Z_{\max} = \max\left(W_{xy}\right) \tag{1}$$

If $Z_{\min} < Z_{\text{med}} < Z_{\max}$, the median is a valid intensity; otherwise, S is incremented until the condition holds or a maximum limit S_{max} is reached. The output pixel is then as shown in Eq. (2):

$$I_c'\left(x,y\right) = \{I_c\left(x,y\right), if\ Z_{\min} < I_c\left(x,y\right) < Z_{\max}\ Z_{\text{med}},$$

otherwise (2)

Where $I_c\left(x,y\right)$ is the pixel value at position $\left(x,y\right)$ in color channel $c \in \{R,G,B\}$.

For a neighborhood window of size $W_{xy}\left(S\right)$ of size $S \times S$, compute the minimum, median, and maximum intensity values of the pixels within the window for further processing.

c) Retinex-Based Image Enhancement:

Flooded environments often present significant visual challenges, such as uneven illumination, low contrast, and reflections, from water surfaces that obscure key features of human figures. Object detection systems have significant challenges both when low-light situations prevail during floods and neighboring challenging conditions of Global Positioning System (GPS) receivers. Such problems are reduced visibility, enhanced noise, and changed accuracy of color, consequently greatly reducing detection accuracy. The negative implications of such issues decrease the precision with which bounding boxes are made and classifications decided and hence cause the precision, recall, and F1 score to decline. To reduce such issues, the proposed preprocessing integrates a Retinex-based lighting-adaptive enhancement technique that dynamically adjusts brightness and contrast. This ensures uniform illumination across image regions, making humans in shadowed or dark areas more visible for the detector and to enhance performance in low illumination settings. These techniques significantly enhanced the clarity of low-light images and helped the proposed model to perform object detection more effectively. The idea behind the retinex model is to enhance image quality by separating the reflectance and illumination components. By substantially increasing the clarity of dark images, these techniques enabled the proposed model to reliably detect objects.

The enhancement is based on the Retinex theory, which models an image as a combination of illumination and reflectance components, expressed in Eq. (3) as:

$$I\left(x,y\right) = R\left(x,y\right) \times L\left(x,y\right) \tag{3}$$

Where:
$R(x, y)$ is reflectance (enhanced image);
$I(x, y)$ is original image intensity;
$L(x, y)$ is illumination component.

The illumination component is modeled using a Gaussian kernel $G(x, y)$ as shown in Eq. (4):

$$L(x, y) = G(x, y) \times I(x, y) \tag{4}$$

Here, G(x,y) smooths the image to estimate illumination.

d) Multi-scale retinex with color restoration (MSRCR):

To ensure semantic uniformity and improve visual separability, a proposed approach integrates the MSRCR algorithm and the application of Gaussian surround functions to calculate spatial illumination. The point is that the efficiency of Retinex model is its capacity to distinguish and separate out reflectance and illumination, thus improving the quality of images. Mathematically, the relationship between the original image and the enhanced image is shown by using Eq. (5) as follows:

$$\text{RMSR}\left(x,y\right) = \sum_{n=1}^{n} W_n \left[logI\left(x,y\right) - log\left(G_n\left(x,y\right)\right.\right.$$

$$\left.\left. * I\left(x,y\right)\right)\right] \tag{5}$$

Where $G_n\left(x,y\right)$ denotes Gaussian kernels of varying standard deviations, and w_n are the corresponding scale weights. This formulation enhances both fine and global features, improving visibility in dark and bright regions simultaneously.

Finally, a dynamic adjustment is applied to adapt brightness and contrast based on average luminance using Eq. (6):

$$I_{\text{enh}}\left(x,y\right) = \alpha \cdot \text{RMSR}\left(x,y\right) + \beta \tag{6}$$

Where α and β are adaptively tuned parameters controlling contrast and brightness, respectively. The resulting image maintains natural colors, improved edge sharpness, and balanced brightness across the scene.

The peak signal-to-noise ratio (PSNR) is widely used to evaluate the quality of image enhancement

techniques. It measures the ratio between the maximum possible power of a signal (image) and the power of noise that affects its representation. In simple terms, it indicates how much the enhanced image differs from the original one.

PSNR is evaluated using the mathematical equation as shown in Eq. (7):

$$\text{PSNR} = 10 \times \log 10\left(\frac{\text{MAX}_I^2}{\text{MSE}}\right) \tag{7}$$

Where:

$$\text{MSE} = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[I(i,j) - K(i,j)]^2.$$

$I(i,j)$: pixel intensity of the original image;

$K(i,j)$: pixel intensity of the enhanced (processed) image;

M and N: dimensions (height and width) of the image;

MAX_I: maximum possible pixel value of the image (e.g., 255 for 8-bit images);

MSE: mean squared error between the original and enhanced image.

As shown in Figure 5 the PSNR value between the original and enhanced image is 10.21 dB, indicating significant improvement in brightness and visibility. This PSNR value suggests that the Retinex adaptive filtering technique significantly modifies image characteristics (contrast, color tone, illumination) to improve visibility in flood scenarios, even if it diverges in pixel intensity from the original.

Retinex reduces this noise by focusing on enhancing the reflectance (the true details of the scene) while smoothing out the illumination (lighting variations), where color correction balances the image's color distribution, restoring natural tones and improving the distinctness of objects. It corrects colors by minimizing the impact of uneven lighting (illumination) while preserving the natural colors and textures (reflectance) of the objects.

c. Object detection with the proposed YOLOv11 model

The object detection models used in this study, YOLOv11, are well known for their exceptional speed and accuracy, making them highly effective for real-time applications, particularly in disaster response scenarios. In the proposed framework, a transfer learning-based approach is employed to accurately identify human objects in flood-affected environments. The YOLOv11 models are trained on annotated flood imagery on the chosen datasets before being fine-tuned to address domain-specific issues like poor visibility, partial occlusions, and changing lighting situations that are prevalent during floods.

This process of fine-tuning allows the models to be more generalized to the different environmental conditions and therefore provides stable performance in detection even in complex scenes. These improvements not only increase the efficiency of the proposed system in terms of detection accuracy but

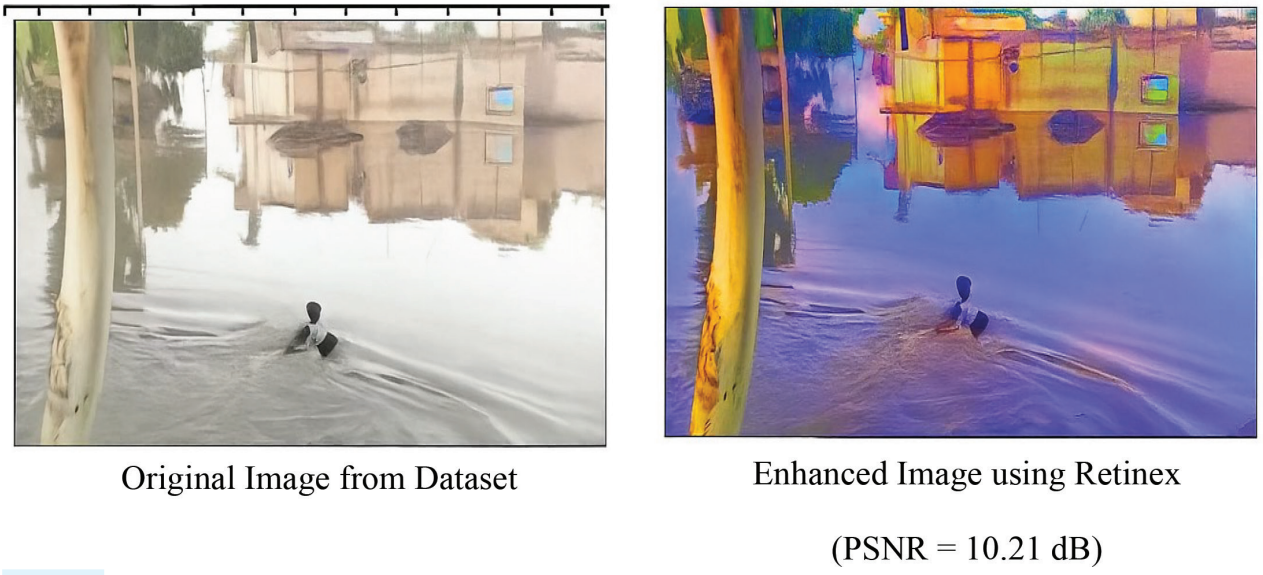

Figure 5: Original and enhanced flood image using Retinex adaptive filtering.

also make it more practical to implement in real-time, which is key to the emergency response applications and flood monitoring. The resilience of these models in varying light conditions and the unfavorable weather also enhances their capability as an effective instrument of human detection in disaster-prone areas.

Moreover, the proposed approach will incorporate the state-of the art image preprocessing with the help of a Retinex-based dynamic lighting-adaptive image enhancement (DLAIE) algorithm. It is a boost in the quality of images in low-light conditions, or when there is uneven light, allowing YOLOv11 to be more attentive to stand-alone, concealed, or half-seen humans. The fusion of both Retinex-based image refinement and fine-tuned YOLO architecture achieves a high level of detection accuracy and robustness of the system.

The parameters and performance measures of the proposed human detection models are optimized as shown in Table 1, with the efficiency and flexibility of YOLOv11 to address the flood conditions in real-world scenarios.

Composite loss function to optimize object detection consisting of localization loss (L_{loc}), confidence loss (L_{conf}), and classification loss (L_{cls}) as shown in Eq. (8):

$$L_{total} = \lambda_{loc} L_{loc} + \lambda_{conf} L_{conf} + \lambda_{cls} L_{cls} \tag{8}$$

Where λ_{loc}, λ_{conf}, and λ_{cls} are weights for balancing the loss terms.

Eq. (9) shows the localization loss, which is calculated using complete intersection over union (CIoU), as

$$L_{loc} = 1 - CIoU \tag{9}$$

where CIoU is defined in Eq. (10) as,

$$CIoU = IoU - \frac{\rho^2\left(b, b^{gt}\right)}{c^2} - \alpha.\upsilon \tag{10}$$

Table 1: Model parameter settings and training configuration

Parameter	Value	Description
Input image size	640 × 640	Standardized input resolution ensuring efficient processing and retention of visual detail.
Batch size	16	Optimized for GPU memory utilization and stable gradient updates during training.
Learning rate	0.01 (with cosine annealing)	Enables smooth convergence with adaptive reduction across epochs.
Optimizer	SGD with momentum = 0.937	Ensures stable learning and faster convergence by maintaining directional consistency.
Anchor boxes	Custom anchors for human detection	Tailored anchor dimensions optimized for human object scales in flood environments.
Epochs	50	Sufficient training iterations to achieve robust convergence and model generalization.
IoU threshold	0.5	Intersection-over-Union threshold for accurate bounding box matching and evaluation.
Loss function	Combined CIoU loss, objectness loss, and class loss	Balances spatial accuracy, confidence prediction, and class differentiation.
Data augmentation	Random flipping, rotation, scaling, brightness, and contrast variations	Increases robustness by simulating diverse lighting and environmental conditions typical of flood scenes.
Preprocessing enhancement	Retinex-based DLAIE	Improves image illumination and visibility for enhanced human detection performance.
Model variant used	YOLOv11	Advanced real-time architectures fine-tuned for detecting humans under variable lighting and occlusion conditions.

CIoU, complete intersection over union; DLAIE, dynamic lighting-adaptive image enhancement; GPU, graphics processing unit; IoU, intersection over union; SGD, stochastic gradient descent.

Where:

b, b^{gt} is predicted and ground-truth bounding boxes;

ρ^2 is squared Euclidean distance between the centers of bounding boxes;

c is the diagonal length of the smallest enclosing box;

υ is aspect ratio consistency.

The dataset includes YOLO annotations for YOLOv8, YOLOv9, and YOLOv11 as well as supports a Faster Region-based Convolutional Neural Network (R-CNN)-compatible annotation format. With these annotations the dataset provides elaborate bounding box locations and probabilities of each class, minimizing the effort required for an object to be detected easily. The annotations center around finding "person" as the category reflective of the central purpose of the dataset identifying persons in flood environments. The use of training set for algorithm development enables accurate detection and localization of persons on flooded landscapes, irrespective of the poor visibility. After the models are trained, it happens to be that the performance of the models is evaluated by using the test sets. When this process occurs, the performance of the models is evaluated on their ability to generalize learned information on unfamiliar situations to and successfully detect people in different flood situations. Besides, a validation set is also important to fine-tune each model's parameters and to determine how well each model performs at each training stage. Following this approach, the models become more robust and credible for practical usage, which improves individual detection in the flood zones.

III. Validation

This proposed work evaluates an approach to human detection in both normal light and varying light or low light conditions in flood situations. The objective of this work is to determine how well the model can accurately detect human objects in flood environments under normal illumination conditions as compared with low light. The evaluation centers around the comparison of certain measures that include precision, recall, and F1 score. Using a dataset containing flood scenarios that disrupted typical conditions, 3 measurements, precision, recall, and F1, score were analyzed closely in regard to how they contribute to the determination of effectiveness of object detection models in the complex environments being studied. In flooded contexts, an in-depth analysis of object detection abilities is necessary. In such conditions, the capability to see humans accurately and successfully can significantly improve SAR impacts, lower response times, and maybe even save lives.

The performance evaluation metrics used for the proposed work are calculated as shown in Eqs 11 to 14:

$$\text{Precision}(P) = \frac{\text{True Positives}(\text{TP})}{\text{True Positives}(\text{TP}) + \text{False Positives}(\text{FP})} \tag{11}$$

$$\text{Recall}(\text{R}) = \frac{\text{True Positives}(\text{TP})}{\text{True Positives}(\text{TP}) + \text{False Negatives}(\text{FN})} \tag{12}$$

$$F1\text{Score}(F1) = 2 \times \frac{P \times R}{P + R} \tag{13}$$

$$\text{Accuracy} = \frac{(\text{TP}) + (\text{TN})}{(\text{TP}) + (\text{TN}) + (\text{FP}) + (\text{FN})} \tag{14}$$

Using a confusion matrix, we managed to allocate predictions into true positives (TP), false positives (FP), false negatives (FN), or true negatives (TN) and obtain these values. Frames per second (FPS) is the metric for speed of computation providing insight into the system's real-time ability. The calculator for FPS is shown in Eq. (15) as

$$\text{FPS} = \frac{\text{Total Frames Processed}}{\text{Total Time Taken}(\text{in seconds})} \tag{15}$$

The model generalization is examined via the experiments with stratified test data that properly represents real variation in (from flood, lighting, and occlusion). Likewise, the model was good in its performance being stable in different cross-validation runs and had the potential to use learned features in unseen flood settings.

As shown in Figure 6A, the initial low-light image had poor contrast and color distortion and thus the human subject was hard to identify. Figure 6B (Retinex-corrected output), on the contrary, has significantly better illumination, lower noise, and restored color balance. The proposed individuals are easily enclosed within the bounding boxes, which proves the increased effectiveness of the YOLOv11 model in identifying and following individuals in the poor-quality lighting conditions.

(A) (B)

Figure 6: Human detection after low light image enhancement. (A) Original image. (B) Detected output.

In general, the results confirm the assumption that Retinex-based enhancement combined with YOLO detectors results in increased confidence and accuracy in human detection, which will be useful in decreasing the misclassification due to background clutter, water reflections, or debris. This validates the possibility of the suggested Retinex-modified detection pipeline to be used in real-time flood rescue and monitoring.

Table 2 represents the performance of the proposed model under daylight versus low light scenarios.

Table 2 and Figure 7 illustrate that the proposed Retinex-optimized YOLOv11 model can ensure high detectability in either regular or diverse light. Although there was a minor decrease in accuracy (95.72–94.83) and F1 score (96.39–92.28) in low light scenes, the precision and recall levels were at a steady high level. The proposed system focused on improving detection under dynamic lighting variations in Red Green Blue (RGB) imagery. While effective under low-contrast and uneven illumination conditions, RGB-based enhancement remains limited in total darkness or sensor-invisible conditions.

The strength of the suggested object detection strategy in various light conditions is emphasized by this validation. This kind of reliability may be of great benefit to rescue teams during actual rescue missions, say in areas hit by floods, where there are frequent changes in light and low visibility, which are usually a hindrance to the normal procedures.

The comparison of true labels (a) and the predicted result (b) obtained using the proposed Retinex-enhanced YOLOv11 model is depicted in Figure 8. The system is also capable of accurate identification of various human cases in the flooded scenes despite poor light, reflections, and occlusions. The fact that the ground truth and predicted box and high confidence scores (0.7–0.9) were close indicates the robustness, accuracy, and flexibility of the model. The Retinex-based enhancement enhances image visibility, which will assist the image to detect human beings accurately in real-time that will be needed in flood rescue missions.

IV. Results and Discussion

To ensure robustness and reliability, the proposed methodology was proved using an approach involving multi-data and multi-model analysis under different levels of floods as a way of testing its robustness and reliability. Two datasets were used: the Roadway Flooding Image Dataset (Kaggle) and the Human Detection in Floods Dataset, which had images of varying illumination, occlusion, and reflection, as are found in the real flood situation.

The performance analysis of the proposed YOLOv11 in comparison with YOLOv9, as presented in Table 3, was conducted under four major evaluation scenarios:

- Without enhancement (Baseline): Both models were evaluated directly on raw flood images to establish baseline detection performance.
- With Retinex enhancement: The models were tested on images preprocessed using the Retinex-

Table 2: Comparative analysis of proposed object detection model in varying lighting conditions

Approach	Lighting condition	Precision	Recall	F1 score	Accuracy
Proposed methodology (YOLOv11 with Retinex enhancement)	Normal light	95.68	97.15	96.39	95.72
	Varying/low-light	92.83	95.56	92.28	94.83

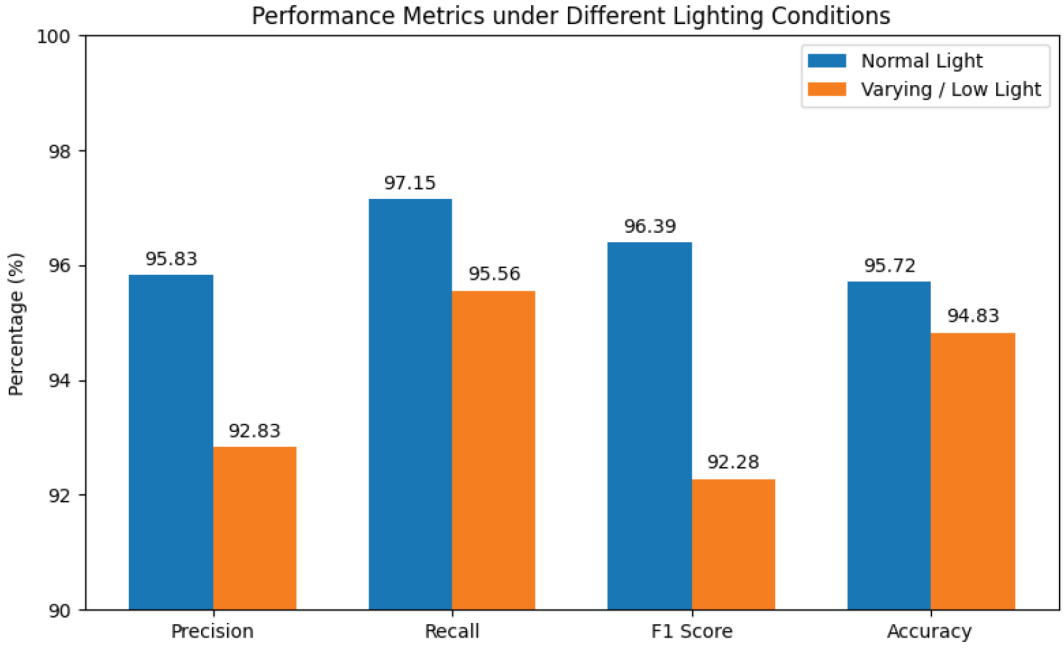

Figure 7: Performance analysis of proposed Retinex + YOLOv11 under normal light and varying light conditions.

based illumination correction and contrast normalization technique.

- Cross-dataset validation: The models were trained on one dataset and evaluated on a different dataset to assess their cross-domain generalization capability.

The measure was Precision, Recall, F1 Score, Accuracy, and FPS as the performance measures. The results validated the fact that the image enhancement by Retinex-based image enhancement significantly improved detection accuracy, particularly under low-light and reflective conditions.

Table 3 and Figure 9 draw conclusions about the effects of image enhancement with Retinex on the object detection performance of the proposed YOLOv11 in comparison with YOLOv9. The findings are a clear indication that the detection efficiency based on all the main evaluation measurements was greatly enhanced when illumination correction and contrast normalization was done utilizing the Retinex algorithm. In the case of YOLOv11, precision and recall have improved significantly (to 95.68 and 97.15, respectively), F1-score has improved (96.39%), and the overall accuracy has improved (to 95.72%). Although the performance improvements were also observed in YOLOv9, post-enhancement measures of YOLOv11 were always greater in all measures. The significant increase in recall of both models implies that the number of missed detections has been greatly reduced, especially during difficult flood conditions in terms of glare, shadow, and partial submersion. It is worth noting that YOLOv11 was the best in terms of overall post-enhancement performance, so it is possible to conclude that it has a high detection performance when it is combined with the Retinex-based illumination correction. These results confirm the usefulness of illumination normalization in

(A) (B)

Figure 8: True labeling and predicted output with proposed system. (A)True labels. (B) Predicted output.

Table 3: Object detection performance improvement analysis with before and after low light image enhancement

Model	Metrics	Before image enhancement (%)	After enhancement (Retinex-based)	Improvement (%)
YOLOv9	Precision	89.50	93.87	+4.37
	Recall	86.30	96.23	+9.93
	F1 Score	87.60	94.41	+6.81
	Accuracy	88.90	94.83	+5.93
YOLOv11	Precision	90.87	95.68	+4.81
	Recall	87.26	97.15	+9.89
	F1 Score	88.90	96.39	+7.49
	Accuracy	89.85	95.72	+5.87

boosting detection confidence, completeness, and strength in the object detection of complex flood scenarios in real-life.

These improvements underline the power of the proposed system in low-visibility or flood-like situations, when abnormalities in the lighting usually compromise the model performance. On a mid-range NVIDIA RTX 3060 graphics card, YOLOv11 was capable of real-time detections at 28 FPS, which is realistic in deploying to the edge, such as a drone or a mobile rescue unit during a disaster in the real world. Table 4 presents the comparative performance analysis of the proposed YOLOv11 model against state-of-the-art detectors, including Faster R-CNN, YOLOv8, and YOLOv9.

Table 4 conclusively proves that the statistical and practical usefulness of the suggested YOLOv11-based framework is more useful than the Faster R-CNN, YOLOv8, and YOLOv9. YOLOv11 has highest precision (95.68%), recall (97.15%), F1-score (96.39%), and accuracy (95.72%) and it has been determined to outperform all the models of its category in all the evaluation

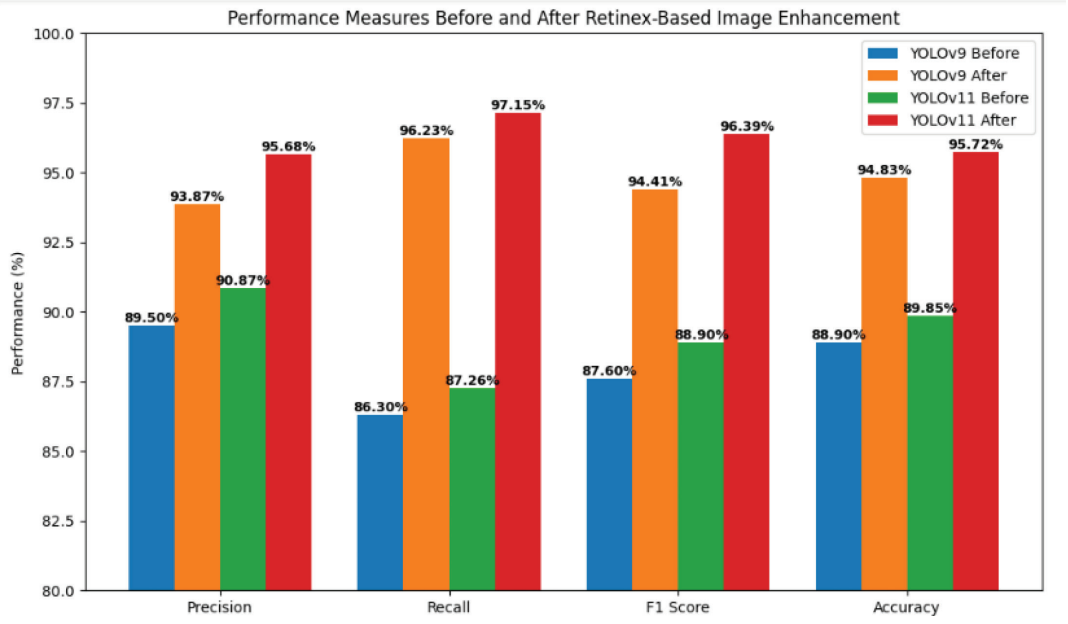

Figure 9: Performance measures before and after low light image enhancement techniques.

Table 4: Comparative performance analysis of the proposed method with state-of-the-art models

Approach	Faster R-CNN	YOLOv8	YOLOv9	YOLOv11
Precision (%)	85.43	92.54	93.87	95.68
Recall (%)	82.86	95.26	96.23	97.15
F1 score (%)	84.13	93.88	94.41	96.39
Accuracy (%)	85.55	93.97	94.83	95.72

metrics. The increased sensitivity in making a detection when the objects are in poor flood conditions, thereby minimizing the rate of FN, can be highlighted by its remarkably high recall. At the same time, the higher F1-score implies a well-optimized precision-recall trade-off, which means that the system detection behavior is stable and reliable. These steady performance improvements not only indicate that it is statistically, but also practically, robust, making YOLOv11 an effective solution to the problem of object detection in real-time, high-risk flood scenarios.

V. Conclusion

The overall results demonstrate the object detection model performances in various illuminated circumstances with particular emphasis on human detection within the floods, and in these conditions, visibility is usually grossly impaired. The models were performing well in normal lighting conditions, but performed poorly during low-light conditions because of poor visibility, more noise, and color distortions before improvement. To deal with these concerns, a Retinex-based image enhancement algorithm was put in place to enhance the consistency of brightness, reduce noise, and recover fidelity to color. The improvement brought significant results to all models with YOLOv11 being the most successful and suggested detection structure. YOLOv11 demonstrated better results than state-of-the-art detectors, including Faster R-CNN, YOLOv8, and YOLOv9.

Moreover, YOLOv11 also maintained real-time detection rates (28 FPS) with a mid-range NVIDIA RTX 3060, demonstrating its relevance to edge-based systems like an autonomous drone or a mobile rescue platform operating in dynamic conditions. The fact that the

proposed Retinex + YOLOv11 model was successful can be explained by the fact that the light enhancement algorithm is able to subtract illumination and reflectance images, boosting contrast and brightness without deforming the structural characteristics in an image.

Conversely, conventional enhancement methods like Histogram Equalization or Contrast Limited Adaptive Histogram Equalization (CLAHE) tend to over enhance noise and distort image details, thus producing unreliable detections. The proposed lighting-adaptive enhancement system aims to improve visual clarity prior to detection rather than explicitly modeling occlusion severity. While partial occlusions may occur in real-world flood scenarios, detailed occlusion-level analysis is considered beyond the scope of this study and is identified as a direction for future research. Additionally, in the future work, the attention-based optimization mechanisms and multi-sensor combination (e.g., RGB-thermal fusion) are going to be considered to improve performance and resilience in harsh environmental conditions. Thermal imaging can provide complementary temperature-based object signatures independent of visible light conditions. This multimodal fusion strategy is expected to significantly enhance detection robustness in extreme environments, particularly in nighttime flood rescue operations and low-visibility disaster scenarios.

References

[1] R. Samanta, G. Bhunia, P. Shit, H. Pourghasemi, Flood susceptibility mapping using geospatial frequency ratio technique: a case study of Subarnarekha River Basin, India, Modeling Earth Systems and Environment, 4 (2018) 395–408. https://doi.org/10.1007/s40808-018-0427-z.

[2] P. Zope, T. Eldho, V. Jothiprakash, Hydrological impacts of land use–land cover change and detention basins on urban flood hazard: a case study of Poisar River basin, Mumbai, India, Natural Hazards, 87 (2017) 1267–1283. https://doi.org/10.1007/s11069-017-2816-4.

[3] Arora, A. Arabameri, M. Pandey, M. Siddiqui, U. Shukla, D. Bui, V. Mishra, A. Bhardwaj, Optimization of state-of-the-art fuzzy-metaheuristic ANFIS-based machine learning models for flood susceptibility prediction mapping in the Middle Ganga Plain, India, The Science of the Total Environment, 750 (2020) 141565. https://doi.org/10.1016/j.scitotenv.2020.141565.

[4] T. Jiang, D. Kong, Video Based Human Head and Shoulders Detection Using Machine Learning, 2021 3rd International Conference on Industrial Artificial Intelligence (IAI), 1 (2021) 1–4. https://doi.org/10.1109/IAI53119.2021.9619282.

[5] J. Fan, J. Lee, I. Jung, Y. Lee, Improvement of Object Detection Based on Faster R-CNN and YOLO, 2021 36th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), 1 (2021) 1–4. https://doi.org/10.1109/ITC-CSCC52171.2021.9501480.

[6] M. Faruk, M. Rahman, Object Detection and Tracking: Deep Learning based Novel Tools to Generate Robust Human and Machine-Annotated Ground Truth Data for Training AI Models, 2022 4th International Conference on Sustainable Technologies for Industry 4.0 (STI), 1 (2022) 1–6. https://doi.org/10.1109/STI56238.2022.10103294.

[7] P. Lamovec, T. Veljanovski, M. Mikoš, K. Oštir, Detecting flooded areas with machine learning techniques: case study of the Selška Sora river flash flood in September 2007, Journal of Applied Remote Sensing, 7 (2013). https://doi.org/10.1117/1.JRS.7.073564.

[8] B. Basnyat, N. Roy, A. Gangopadhyay, Towards AI Conversing: FloodBot using Deep Learning Model Stacks, 2020 IEEE International Conference on Smart Computing (SMARTCOMP), 33 (2020) 33–40. https://doi.org/10.1109/SMARTCOMP50058.2020.00025.

[9] K. Islam, M. Shahabuddin, C. Kwan, J. Li, Flood Detection Using Multi-Modal and Multi-Temporal Images: A Comparative Study, Remote Sens., 12 (2020) 2455. https://doi.org/10.3390/rs12152455.

[10] P. Zachar, Z. Kurczynski, W. Ostrowski, Application of Machine Learning for Object Detection in Oblique Aerial Images, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (2022). https://doi.org/10.5194/isprs-archives-xliii-b2-2022-657-2022.

[11] U. Naik, V. Rajesh, R. R, Implementation of YOLOv4 Algorithm for Multiple Object Detection in Image and Video Dataset using Deep Learning and Artificial Intelligence for Urban Traffic Video Surveillance Application, 2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT), 1 (2021) 1–6. https://doi.org/10.1109/icecct52121.2021.9616625.

[12] Y. Chen,S. Dwivedi, M. Black, D. Tzionas, Detecting Human-Object Contact in Images,2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),17100 (2023) 17100–17110. https://doi.org/10.1109/CVPR52729.2023.01640

[13] A. Nallapareddy, B. Balakrishnan, Automatic Flood Detection in Multi-Temporal Sentinel-1 Synthetic Aperture Radar Imagery Using ANN Algorithms, Int. J. Comput. Commun. Control, 15 (2020). https://doi.org/10.15837/ijccc.2020.3.3616.

[14] J. Oeing, L. Neuendorf, L. Bittorf, W. Krieger, N. Kockmann, Machine and Deep Learning Approaches for Flooding Prevention in Distillation and Extraction Columns (2020). https://doi.org/10.22541/au.160660624.43119897/v1.

[15] A. Sarkale, K. Shah, A. Chaudhary, T. Nagarhalli, An Innovative Machine Learning Approach for Object Detection and Recognition, 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 1008 (2018) 1008–1010. https://doi.org/10.1109/ICICCT.2018.8473221.

[16] Wan, J., Shen, Y., Xue, F., Yan, X., Qin, Y., Yang, T., & Wang, Q. J. (2024). DSC-YOLOv8n: An advanced automatic detection algorithm for urban flood levels. *Journal of Hydrology*, *643*, 132028. DOI:10.1016/j.jhydrol.2024.132028.

[17] Lei, C., & Tian, Q. (2023). Low-Light Image Enhancement Algorithm Based on Deep Learning and Retinex Theory. Applied Sciences, 13(18), 10336, Mdpi, 10 (2023) 101068. https://doi.org/10.3390/app131810336.

[18] Zeng, Y. F., Chang, M. J., & Lin, G. F. (2024). A novel AI-based model for real-time flooding image recognition using super-resolution generative adversarial networks. *Journal of Hydrology*, 131475. DOI:10.1016/j.jhydrol.2024.131475

[19] Rasheed, M. T., Guo, G., Shi, D., Khan, H., & Cheng, X. (2022). An empirical study on retinex methods for low-light image enhancement. *Remote Sensing*, *14*(18), 4608. https://doi.org/10.3390/rs14184608. https://doi.org/10.3390/rs14184608

[20] Cheon, B. W., & Kim, N. H. (2023). Enhancement of Low-Light Images Using Illumination Estimate and Local Steering Kernel. *Applied Sciences*, *13*(20), 11394. https://doi.org/10.3390/app132011394. https://doi.org/10.3390/app132011394

[21] Gupta, H., Kotlyar, O., Andreasson, H., & Lilienthal, A. J. (2024). Robust Object Detection in Challenging Weather Conditions. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 7523-7532). DOI:10.1109/WACV57701.2024.00735

[22] Lu, J., Li, J., Chen, G., Zhao, L., Xiong, B., & Kuang, G. (2015). Improving pixel-based change detection accuracy using an object-based approach in multitemporal SAR flood images. IEEE journal of selected topics in applied earth observations and remote sensing, 8(7), 3486-3496. Digital Object Identifier 10.1109/JSTARS.2015.2416635

[23] Prabhu, B. B., Lakshmi, R., Ankitha, R., Prateeksha, M. S., & Priya, N. C. (2022). RescueNet: YOLO-based object detection model for detection and counting of flood survivors. Modeling Earth Systems and Environment, 8(4), 4509-4516. DOI:10.1007/s40808-022-01414-6

[24] CM, B. (2023). Robust human detection system in flood related images with data augmentation. Multimedia Tools and Applications, 82(7), 10661-10679. https://doi.org/10.1007/s11042-022-13760

[25] Nehete P. U., Dharrao D. S., Pise, P., & Bongale, A. (2024). Object Detection and Classification in Human Rescue Operations: Deep Learning Strategies for Flooded Environments. International Journal of Safety & Security Engineering, 14(2). DOI: https://doi.org/10.18280/ijsse.140226.

[26] Kaur, J., & Singh, W. (2024). A systematic review of object detection from images using deep learning. Multimedia Tools and Applications, 83(4), 12253-12338. https://doi.org/10.1007/s11042-023-15981-y.

[27] Jamali, M., Davidsson, P., Khoshkangini, R., Ljungqvist, M. G., & Mihailescu, R. C. (2025). Context in object detection: a systematic literature review. Artificial Intelligence Review, 58(6), 1-89. https://doi.org/10.48550/arXiv.2503.23249

[28] Xue, F., Cheng, Y., Shen, Y., Chen, J., & Wan, J. (2025). Urban Flood Inundation Area Detection using YOLOv8 Model. Water Resources Management, 1-18. DOI:10.1007/s11269-025-04211-9

[29] Hasnaoui, Y., Tachi, S. E., Bouguerra, H., & Yaseen, Z. M. (2025). Transfer learning-based deep learning models for flood and erosion detection in the coastal area of Algeria. Earth Science Informatics, 18(2), 380.

[30] Sun, J., Xu, C., Zhang, C., Zheng, Y., Wang, P., & Liu, H. (2025). Flood scenarios vehicle detection algorithm based on improved YOLOv9. Multimedia Systems, 31(1), 74. DOI:10.1007/s00530-024-01661-w

[31] Wu, J., Li, W., Du, H., Wan, Y., Yang, S., & Xiao, Y. (2025). An annotated satellite imagery dataset for automated river barrier object detection. Scientific Data, 12(1), 237. DOI:10.1038/s41597-025-04590-z.

[32] Ghosh, T., Paul, A., & Chaki, N. (2024, July). Flood detection in UAV images of urban areas using machine learning and deep learning techniques. In International Symposium on Applied Computing for Software and Smart Systems (pp. 225-239). Singapore: Springer Nature Singapore.

[33] Satpati, S., & Bhadra, J. (2025). Natural Disaster Recovery for Flood-Affected Areas Using YOLOv7 Object Detection and Instance Segmentation. In Geographies of the Indian Subcontinent: Cutting-Edge Methods in 21st Century Research (pp. 271-282). Cham: Springer Nature Switzerland. DOI:10.1007/978-3-031-74813-4_12.

[34] https://www.kaggle.com/datasets/saurabhshahane/roadway-flooding-image-dataset

[35] https://universe.roboflow.com/yolo-lkgro/human-detection-qrsic/dataset/3/download/yolov8.