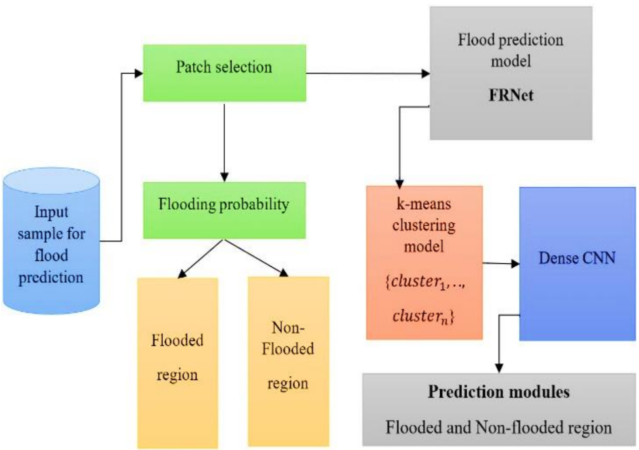
Global NEST Journal, Vol 27, No 5, 06702

Design of a dense flood-region prediction network for analyzing the flood over urban region

Madhumitha Ramamurthy[1], Sherubha P.[2], Sasirekha S. P.[3] and Rajendran Thavasimuthu[4*]
[1,2]Department of Information Technology, Karpagam College of Engineering, Coimbatore, Tamil Nadu, India
[3]Department of Computer Science and Engineering, Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India
[4]Department of Sustainable Engineering, Saveetha School of Engineering, Chennai, Tamil Nadu, India
Received: 24/08/2024, Accepted: 22/02/2025, Available online: 16/04/2025
*to whom all correspondence should be addressed: e-mail: madhuperu@gmail.com
https://doi.org/10.30955/gnj.06702

Graphical abstract



Abstract

Hydrological monitoring in urban regions is crucial to analyze hydrology and manage floods. Nevertheless, monitoring flood overflow in urban areas including the degree of flood overflow is frequently insufficient. Hydrological models' ability to provide early flood warnings is limited by this deficiency, making calibration difficult. This work developed a technique for assessing the urban floods depth using deep learning and image recognition methodologies to overcome this constraint. To recognize submerged things in images like a pedestrian's leg or a car's exhaust pipe, this approach uses the object identification model known as the Flood Region-prediction network (FRNet). The mean average precision for water depth identification in a flood images collection was 99%. According to the research, the reference object types submerged in the flood impacts how well FRNet detects flood depth; utilizing a car as the reference objects produced a high range of accuracy than employing a person. Moreover, the accuracy of recognition was considerably improved by image augmentation using learning technology. The presented method uses data as image or video from traffic cameras that are currently in use to extract on-site, continuous, real-time data of water depth. This technology provides an affordable and instantly deployable solution by removing the requirement to install additional water meters.

Keywords: Flood, image recognition, deep learning, prediction, accuracy

1. Introduction

Urban flooding poses growing issues for sustainable development globally. The impacts of climate change and increasing urbanization are to blame for the increase in flood damage frequency and severity. According to (Arshad *et al.* 2019), urban hydrology provides a foundation for estimating submerged areas and the damage caused by urban flooding. As cities have many impermeable surfaces, surface water flow speeds up during a storm while precipitation infiltration decreases (Bai *et al.* 2018). These circumstances provide a short period of time for alerting and evacuating individuals who may be affected by flash floods. In addition, the major vulnerability of humans and more costly infrastructure in urban regions make urban flash floods especially damaging (Basnyat *et al.* 2018).

Flooding is a significant natural disaster that poses severe risks to urban regions, impacting lives, infrastructure, and the economy. As climate change intensifies, the frequency and severity of floods are expected to increase, making urban flood prediction and monitoring more crucial than ever. Traditional hydrological models, while effective to an extent, often struggle to provide real-time, accurate assessments of flood conditions, particularly in dense urban environments where variables are numerous and complex. The ability to monitor flood conditions accurately is essential not only for timely evacuation and disaster response but also for minimizing economic losses and protecting vulnerable populations. Numerous hydrological models, including the Mike Floods and Storm Water Management Models, have been created to simulate urban floods; these models are used extensively worldwide. These models make it possible to anticipate places in cities that will be flooded which is crucial for early warnings of urban flooding (Bochkovskiy *et al.* 2020).

The lack of observed data on floodwater depths in urban areas, however, creates a significant gap in the calibration and validation of the model. According to Bulti and Abebe (2020), the key difficulties are the monitoring of narrow coverage areas and comparatively high equipment maintenance cost. Subway stations and traffic tunnels are just two examples of the many underground infrastructures and low-lying amenities contributing to dispersed and frequent waterlogging in urban areas. This means that employing a standard sensor-based system to monitor distributed flood overflow at these dispersed sites becomes expensive. According to (Chia *et al.* 2023), few worldwide comprehensive monitoring networks cover urban flood depths.

1.1. Problem statement

In the last ten years, numerous new technologies have been implemented to provide more flood depth data in urban areas. These include object detection in flood photography and remote sensing, deep learning (DL), crowdsourcing, or social media retrieval of data. The authors (Faramarzzadeh *et al.* 2023) developed a technique to distinguish "non-flood" and "flood" occurrences from the data set of over hundred thousand of online images by using artificial intelligence and content-based image retrieval technologies. By examining text from online news sources, the author created a DL classifier that identified flood incidents (Feng *et al.* 2021). Even though these methods have brought new perspectives to urban flood monitoring are needed to collect reliable, ongoing data on floodwater levels at particular sites. Higher-resolution images from the water level meters taken with CCTV devices were used to infer the data of water level with the identification methodologies to improve flood images' usefulness (Gauen *et al.* 2017).

Additionally, with a Raspberry Pi camera and image processing and text recognition methods, existing researchers have created low-power, low-cost cyber-physical system-based prototypes for detecting the increasing water range. This model could enable the broad implementation of a flash flood identification model that needs less human oversight (Guo *et al.* 2021; Hammond *et al.* 2015 & Ichiba *et al.* 2018). However, there are difficulties for implementing these systems in urban areas, mainly because frequently trafficked metropolitan areas need the installation of camera systems and water level meters on-site (Jiang *et al.* 2019; Jiang *et al.* 2020 & Kankanamge *et al.* 2020).

The main objective of this investigation is to introduce an efficient and precise approach for predicting flood regions by employing Dense Convolutional Neural Networks (Dense-CNN), which can detect flood-prone areas based on several factors like hydrological, meteorological, and geographical information. Real-time flood forecasts are intended to aid in risk reduction, emergency preparation, and rapid execution of evacuation protocols. In order to evaluate and retrieve the required characteristics from the source dataset, the research aims to construct and develop a Dense-CNN model. It will then be optimized by utilizing regularization approaches, modifying learning rates, and updating hyperparameters. One of the most frequent and destructive natural calamities is floods, which cause significant property loss, financial inconsistency, and loss of human lives. Still, no research provides precise techniques to predict floods since there are constant variations in weather conditions and climatic fluctuations. Current models frequently need complicated configurations and generate inaccurate outcomes. The ability of Dense-CNN to handle complicated, high-dimensional information and enhance identification accuracy by integrating learned characteristics and facilitating improved gradient flow helps in the ideal prediction of the flood. The authors suggested employing a Dense-convolutional neural network to build a reliable framework to predict flood-prone regions that results in improved accuracy, which in turn helps to support governments, affected people, and disaster rescue teams taking preventive measures in the flood situation. The main goal of this work is to evaluate data related to ecosystems and generate maps that depict regions prone to flood. The framework can be tuned and employed in several areas with varying environmental situations and terrain types.

1.2. Research contribution

Considering the above challenges, this research aims to create a novel technique that can continuously record urban flooding at a particular site during a flash flood. During floods, traffic images are used to determine the flood level using image recognition technology. This is done by selecting the relative places of the water surfaces to refer the objects in the water, like peoples and cars moving through the area. The article's main concept is the development of a novel technique for surface water depth image detection. The method presented in the research was more computationally effective and cost-effective than traditional methods which use sensors to monitor and statistical methodologies to acquire flood overflow depth. This makes it appropriate for providing flood overflow depth data through higher spatial extents. Additionally, by examining images from current road traffics surveillance camera, the developed FRNet technology can be able to present real-time data on flooded water levels. With the incidence of traffics surveillance cameras in urban regions worldwide, this methodology may help increase monitoring of urban floods and promptly deliver an early alert.

This research proposes a novel approach to flood prediction using deep learning and image recognition technologies. The proposed Flood Region-prediction network (FRNet) leverages advanced machine learning techniques to analyze images from existing traffic cameras, identifying and assessing the depth of floodwaters in real time. By incorporating methods such as k-means clustering and Dense CNN, the model can differentiate between flooded and non-flooded areas with high precision. This approach eliminates the need for additional infrastructure, providing an efficient, cost-effective solution for urban flood monitoring that can be rapidly deployed to enhance disaster preparedness and response efforts.

The work is structured as follows: Section 2 provides a broader analysis of various prevailing approaches. The methodology is explained in section 3. The numerical outcomes are provided in section 4, with the conclusion in section 5.

2. Related works

Around the world, flooding is a common and severe natural disaster that impacts people's lives, local ecosystems, infrastructure, and the economy. Floods can result in various property damage. Furthermore, the most vulnerable people in society bear a disproportionate share of the economic consequences of flood damage (Li *et al.* 2018). Satellite synthetic aperture radar (SAR) data are the most commonly utilized Earth observation (EO) data for operational flood monitoring because of their imaging capabilities which enable data collecting independent of light and weather conditions (Li *et al.* 2019). When employing SAR amplitude data for flood mapping, smooth water surfaces' specular reflection typically produces a black tone in the SAR data which helps to differentiate floodwater from other land surfaces (Mignot *et al.*2019 & Misra (2019)). Samples from metropolitan areas appear brighter due to increased backscatter caused by flooding.

The precise segmentation and classifications of flood water and the deployment of data sets for model training and comparisons were some of the problems in this area. Binary segmentation utilizing thresholding techniques like Otsu is a widely used technique. Nigussie and Altunkaynak *et al.* (2019) detailed their thorough examination of well-established pixel-based flood mapping techniques using SAR images of various flooded areas including global and enhanced thresholding, active contour modelling, and change detection. However, the effects of backscatter, geographic location, and image capture duration limit these approaches' potential to be broadly applied. Rule-based classification techniques used by (Park *et al.* 2021), represent another way of investigation for flood detection using SAR data. Using rule-based classification and Taguchi optimization approaches to map the extent of flooding, few studies have used texture information to map floods with SAR images (Rangari *et al.* 2021).

A texture analysis technique is presented to map the extent of flooding using a time series of SAR images. Web-based flood mapping applications outlined by (Rong *et al.* 2020) are also being introduced with Google Earth Engine's development. These models use a long stretch of pre- and post-flood SAR images to calculate each pixel's Z-score, standard deviation, and mean. In recent years, flood mapping has increasingly relied on machine learning techniques. Self-organized maps, support vector machines and Bayesian network fusion are techniques used to extract flooded areas from optical and SAR satellite images (Singh *et al.* 2018). However, the unavailability of large-scale labelled flood data sets has limited research in this area. Convolutional neural networks which are a representation of deep learning techniques have shown promising outcomes in assessing flood damage.

Additionally, the author has shown that these techniques have facilitated the creation of novel automated methods for extracting flood extent from SAR images. The performance of integrating a CNN with the data fusion model for creating mapping of compound flood was evaluated by (Wong *et al.* 2021) using multi-spectral land sat images and dual-polarized synthetic apertures radar images. These studies demonstrated the significant contribution that deep learning algorithms provide to improving flood classification. The researchers (Wang *et al.* 2013), in their recent publication of the extensive open-source Sen1Floods11 dataset, have increased research in DL algorithms for water types recognition in flood disaster. Similarly, the authors (Wang *et al.* 2018) tested the system on an untested set of the similar data set and used it for training and validation. Additionally, models such as UNet and Segnet are used for flood mappings and comparisons of various outcomes depending on various labelling methodologies (Yin *et al.* 2015).

As reported, fully convolutional network (FCN) was widely used for flood mappings. Previous research has demonstrated that rule-based and deep-learning approaches can both be used to detect floods. It is crucial to assess and compare the global performance of some of these DL models and whether they could improve automatedly created flood mappings (Yu *et al.* 2018). Improved detection quality also contributes to the timely and accurate mappings of flood and provides data for management services of disaster reliefs (Zou *et al.* 2019). **Table 1** shows the comparative analysis of reviewed current research works.

3. Proposed research methodology

The proposed model for flood prediction integrates multiple components to assess the probability of flooding in urban regions. The model begins with the input sample, which undergoes patch selection to identify relevant areas for analysis. These patches are evaluated for flooding probability, determining whether a region is flooded or non-flooded. Simultaneously, the model employs a k-means clustering technique to classify different regions into clusters. These clusters are then processed by a Dense-CNN to refine predictions. The Flood Region-prediction model, termed FRNet, consolidates this information to predict and distinguish between flooded and non-flooded regions effectively. The model's design aims to improve the accuracy of flood detection and provide actionable insights for disaster management.

3.1. Dataset

Water depth identification was applied to 1,177 images of flooded roadways with cars or pedestrians. The image resolutions ranged from 4613 x 2595 pixels to 142 x 107 pixels. Search terms like "urban waterlogging" and "urban floods" were used to find these images on Google (https://www.kaggle.com/datasets/zaraks/pascal-voc-2007). Traffic cameras provided some of the images for this collection. The selection of the flood images considered difficult and complex conditions like those with dirty water, nightfall, and splashes from moving vehicles.

Table 1. Comparative Analysis of Reviewed Current Research Works

Ref	Approach/Dataset	Application	Advantages	Disadvantages
Li *et al.* (2018)	General flood impact study	Global impact of floods on society, economy, and environment	Highlights the disproportionate effect on vulnerable populations	Broad and general, lacks specific technical solutions
Li *et al.* (2019)	Satellite SAR data for flood monitoring	Operational flood monitoring using EO data	Independent of light and weather conditions	Challenges in classification and segmentation of floodwater
Mignot *et al.* (2019)	SAR amplitude data, thresholding (e.g., Otsu)	Flood mapping via SAR data	Clear differentiation of floodwater from other surfaces	Limited by dataset availability, potential accuracy issues due to backscatter
Misra (2019)	SAR amplitude data	Urban flood detection using SAR	Effective in detecting smooth water surfaces	Same limitations as Mignot *et al.* (2019)
Nigussie & Altunkaynak (2019)	Pixel-based flood mapping with SAR images	Global flood mapping using various pixel-based techniques	Comprehensive evaluation of multiple techniques	Limited broad applicability due to backscatter, geographic and temporal constraints
Park *et al.* (2021)	Rule-based classification	Flood detection using SAR data	Simplicity and ease of implementation	Lack flexibility and adaptability
Rangari *et al.* (2021)	Rule-based classification, Taguchi optimization	Mapping flood extent using SAR images	Integration of texture information enhances mapping	Requires optimization and complex to implement
Rong *et al.* (2020)	Web-based flood mapping	Pre- and post-flood analysis using SAR time series	Access to real-time data, user-friendly	Dependent on available SAR imagery, potential limitations in real-time accuracy
Singh *et al.* (2018)	SOM, SVM, and Bayesian	Flood area extraction from optical and SAR images	Promising outcomes, potential for automation	Lack of large-scale labelled datasets limits broader application
Wong *et al.* (2021)	CNN with data fusion framework	Compound flood mapping using multi-spectral and SAR images	Significant improvement in flood classification	Requires advanced technical knowledge, dependent on image quality
Wang *et al.* (2013)	Sen1Floods11	Deep learning research for water type recognition	Facilitates deep learning research, availability of dataset	Limited to dataset-specific applications
Wang *et al.* (2018)	Testing on Sen1Floods11 dataset	Training and validation for deep learning models	Provides validation for deep learning approaches	Dataset-specific results, not generalize well
Yin *et al.* (2015)	UNet and SegNet	Comparison of flood mappings using various labelling methodologies	Provides baseline for comparison of different models	Not account for all variables in flood mapping
Yu *et al.* (2018)	FCN	Flood detection and mapping using FCN	Widely used and accepted in flood mapping	Requires significant computational resources
Zou *et al.* (2019)	Deep learning models	Disaster relief management and flood mapping	Improved detection quality, timely information	Computationally intensive, require large datasets

3.2. Detection algorithm

The FRNet pipeline comprises four primary phases as depicted in **Figure 1**. The process involves four steps: 1) training the network and extracting patches out of flood locations; 2) clustering candidates patches depending on their flood region morphology; 3) creating prognosis-oriented models particular to morphologies and general detection-related models; 4) combining the features of both models to arrive at an overall prediction.

Owing to the intricacy of structure and textural features, regions in flood areas are complex components. This includes normal, collagen and complex flood regions that lack information linked to detection. We train FRNet using the available classification dataset to discriminate between flooded and non-flooded regions to mitigate the noisy patches resulting from non-flooded regions. To facilitate the training of the prognosis prediction model, we use FRNet to score the flood areas and identify patches with a high likelihood of flood for additional examination. We train FRNet on the flooded dataset and the network architecture. The non-flooded patch regions framed in **Figure 1** contain complex flood regions and standard and other areas. In the meantime, patches that are surrounded in red are flooded patches that were taken from the area that FRNet determined to be complex.

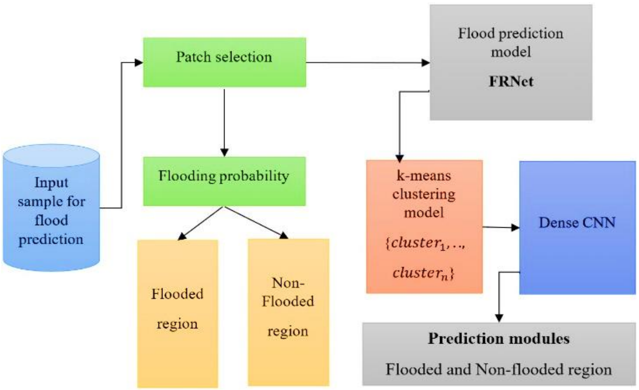


Figure 1. Pipeline of the Proposed Prediction Model

The novel FRNet can extract features from the entire flooded region and predict prognosis by training it using the patches sampled from the flooded area. It is seen that hundreds of these patches are highly pertinent to flood and can be obtained from a single image. Furthermore, it would enable to train more robust networks and include more imaging attributes. There are dense network blocks in the FRNet. Because every layer in a dense-CNN block was physically linked to every other layer, the data flow among the layers in the block is improved because all convolution layers' output features are passed on to every other layer. This layout can facilitate the extraction of image features and improve the flow of image information between layers. Formally, image features $E \in R^{s\times s\times 3}$ is the input of FRNet given a patch p, where S is the patch size as in Eq. (1):

$$h_1^m = \propto_m \left(W_1^m E^m + b_1^m \right) \tag{1}$$

The features of input, matrix weight, and bias of the initial layers of the m^{th} dense-CNN block was represented by the variables E^m, $W_1{}^m$, and $b_1{}^m$ and α_m was the non-linear activation function (ReLU) for the m^{th} dense CNN blocks. The concatenated characteristics of every preceding layer was the input of the dense CNN block layer which may improve the flow of image data. It is possible to compute the feature map output of the i^{th} layer in the m block of dense-CNN in the following way in Eq. (2):

$$h_1^m = \propto_m \left(W_i^m \left[E^m, h_1^m, \ldots, h_{i-1}^m \right] + b_i^m \right),\ 2 \le i \le N \tag{2}$$

Where, N denotes the number of layers in all the network blocks, and $h_{i-1}{}^m$ was the feature maps generated by the $i-1$ convolution layers of m block. Additionally, $W_i{}^m \in R^{N\times K\times i}$ with D_i being the total filters in i^{th} layer. After much experimentation, we find that FRNet performs best with diverse, dense blocks. **Table 2** displays the architecture of the FRNet. Each sample's general level feature is obtained by feeding FRNet with all patches sampled from the flooded area of the sample image. Subsequently, we utilize an average pooling method to fuse the features of the patches that belong to the same image. Given the heterogeneity of the flooded region and the importance of specific flood patches in distinguishing between them, prognosis prediction is crucial. Upon seeing that flooded morphological features have demonstrated clustering patterns, researchers cluster image patches rather than relying on the flooded annotation.

To represent feature morphological aspects, researchers employ patches that have dimensions of 128*128*3. Once the feature is extended by a row pixel, it has a highly high dimension. As such, it is imperative to decrease the feature's dimension. Some researchers used principal components analysis (PCA) to decrease imaging characteristics from 49152 to 50 dimensions. Next, to differentiate between several candidate patches, we employ $K-$ means clustering where n observations are divided into k groups using $K-$ means clustering and each group's sum of squares is minimized. Using K-means clustering, we partition the n patches into k sets $F = \{F_1, F_2, \ldots, F_k\}$ given a set of patches $(p_1, p_2, \ldots p_n)$, where a 50-dimensional feature vector represents each patch as in Eq. (3):

$$\underset{s}{\operatorname{argmin}} \sum_{i=1} \sum_{x \in F_i} \left\| x - \mu_i \right\|^2 \tag{3}$$

Formally, the goal is to find where μ_i was the point's mean in the i^{th} set, and x is the 50-dimensional characteristics. In this work, k is equivalent to 5. Based on the structural flood characteristics, this stage will categorize the patches into multiple groups. We display a few region clustering results. Various flooded features are grouped into distinct categories. For instance, interstitial infiltration is probably present in one set of patches while flooded regions are present in another set. We train FRNet on each cluster to obtain prognosis-associated features specific to morphology.

Table 2 displays the FRNet architecture. Models trained on various clusters extract distinct feature characteristics associated with prediction. In addition, we train FRNet to separate the flood region into flood and non-flood parts by implementing a feature fusion approach suggested to produce morphology-specific features for each area. Then, we collect patches from the malignant zone to train the FRNet and expected it for predicting prognosis at the basic level. Additionally, FRNet can estimate prognosis from the level of flooded morphology since patches of clusters with distinct flood characteristics power them. The patches of features are represented differently in the two models. However, they may complement one another. Next, we use the following equation to combine the flood level attributes of the two models as in Eq. (4):

$$final\ flood\ feature = flood_{feature} \oplus non-flood_{feature} \quad (4)$$

Where the variable $\oplus$ represents the concatenation operations, the flood feature was the general flood level features generated from FRNet and the $non{-}flood_{feature}$ feature was the morphology-specified features generated from FRNet. We finally use the Cox proportional hazard method to forecast the risk of flood survival based on final features.

Table 2. FRNet Architecture

Network	Layers	Channel	Kernel	Output
Dense_CNN_1	Conv_1	10	3*3	128*128
	Conv_2	10		
	Conv_3	20		
	Conv_4	30		
Intermediate layer_1	Conv_5	10	1*1	64*64
	Max_pool_1	10	2*2	
Dense_CNN_2	Conv_6	10	3*3	
	Conv_7	20		
	Conv_8	30		
Intermediate layer_2	Conv_9	10	1*1	32*32
	Max_pool_2	10	3*3	
Dense_CNN_3	Conv_10	10		
	Conv_11	20		
	Conv_12	30		
Prediction module	Max_pool_3	30	4*4	8*8
	Conv_13	16	3*3	8*8
	Max_pool_4	16	2*2	4*4
	Flatten	--	--	256*1
	FC_1	--	--	48*1
	FC_2	--	--	1*1

Algorithm 1:

Input: dataset sample, survival time, patch size and survival status

//**stage_1 initialization**

1. Analyze non-overlapping image region for extracting image patches from the flooded regions;

//extract the samples from the dataset https://www.kaggle.com/datasets/zaraks/pascal-voc-2007

2. **for all** flood region patches, **do**
3. **if** FRNet ($p > 0.5$) as flooded patch //set threshold limit
4. **else** non-flooded
5. **end for**
6. sampling flooded region patches from samples provided to FRNet; //examine the region-based patches from the samples

//**stage_2 initialization**

7. Train FRNet; //train the proposed FRNet network model
8. model = train FRNet ($\delta, patches$)
9. **for all** flood region patches **do**
10. PCA = component analysis ($feature_{patches}$);
11. region clustering = $k - means$ ($no.of\ clusters\ and\ PCA$) //clustering
12. **end for**
13. **for all** clusters, **do**
14. model = train FRNet ($\delta, patches$)
15. **end for**

//**stage_3 initialization**

16. Final feature = $flood_{feature}$ $\oplus$ non $-$ $flood_{feature}$; //feature representation
17. Final_prediction = train cox ($final\ flood\ feature, \delta, t$)

Output: risk score

3.3. Loss function analysis

This section helps to predict how long a flood death event which is the goal of survival analysis. A common occurrence in survival analysis is a specific flood region, meaning that the sample was still alive at the time of the flood and that their actual survival times may have exceeded the follow-up measurements. The restricted data cannot be modelled by conventional classification and regression loss techniques. The amount of data available for network development would be limited if samples were excluded from the training set under consideration. For model optimization, we can thus employ the Cox loss function which is commonly used in the survival analysis field. Minimizing negative log partial likelihood can be used to optimize the model parameters as in Eq. (5):

$$final\ flood\ feature = flood_{feature} \oplus \mathrm{non-flood}_{feature} \tag{5}$$

In this example, n denotes the total training samples, R_i was the samples set with no experience of the events before the sample i^{th}, and δ_i was the indicator, if the survival time was censored with one representing a Non-Flooded Samples and zero representing the flooded samples.

4. Numerical results and discussion

The research involved the random partitioning of the input data into three sets: a test set (10%), a training set (81%), and validation set (9%). Throughout the training process, the test set was not used. FRNet was trained using the deep learning approach in which the starting parameters (such as the CNN's initial weights) were taken from the VOC2007 dataset. The proposed learning is acknowledged as a useful technique in the training process that uses general features from more extensive training datasets. As a result, the model's training time is shortened. With over 500,000 labelled images split into 20 item groups, the VOC2007 data set provides a sizable and varied base to build initial weights for the CNN used in this investigation. Then, 200 epochs were set aside for the model training in this investigation. The first ninety epochs used a learning strategy with the learning rate $1e^{-3}$ and the remaining iterations were used for regular learning with the learning rate $1e^{-4}$.

Frozen learning produced only small modifications to other networks throughout training and not alter the network's starting parameters in the backbone section. On the other hand, regular learning requires changing the model's parameters for every network. By using pre-existing parameters of network from earlier datasets like VOC2007, frozen learning reduces the time needed for training. But, after only a few epochs, the model not converge because of the backbone network's static nature during frozen learning. Thus, ongoing learning was used moving forward to guarantee additional model convergences. A lowered learning rate is used in regular learning to lessen the chance of overfitting further. After 70 epochs, the stabilization in reduction of training loss is noted, likely caused by the network structure being unaltered during the frozen learning phase. After frozen learning was stopped, a subsequent decrease in the training loss persisted till the 150th iteration.

4.1. Performance metrics

The flood image identification accuracy was assessed in this study by mean average precision or MAP. According to [26], this metric is widely used in object identification and classification. The recall and precision indicators of the modelling outputs were used to calculate the map. According to Eqs. 6 and 7, recall is the ratio of correctly predicted positive observations to all actual positives, and precision is the ratio of correctly predicted positive observations to all positive ones where FP (false positives) is the total samples that were incorrectly identified, FN (false negatives) denotes the total unacknowledged outcomes, and TP (true positives) indicates the total samples that were correctly identified or categorized. The average precision, or AP, is the area under the recall-precision curves of the identified results for all the categories. Using the indicators above, the map can be computed as follows. There are a total of n categories. The Precision-Recall curve is represented by p(r), where r is the recall value from Eq. (6) to Eq. (9). The most significant value of the associated accuracy at a specific interval, say 0.1, when the recall value rises from 0 to 1 is represented by the y-axis of p(r).

$$Precision = \frac{TP}{TP+FP} \tag{6}$$

$$Recall = \frac{TP}{TP+FN} \tag{7}$$

$$mAP = \frac{\sum_1^n AP}{n} \tag{8}$$

$$AP = \int_0^1 p(r)dr \tag{9}$$

4.2. Result analysis

Table 3 shows the findings of the street image-based flood recognition. The reference objects, including cars and peoples, were encased in bounding boxes, as depicted in the image. The top of each box displays the detected category of water depth. This method uses street images or traffic camera surveillance films to extract the flood water depth through real-time image recognition. The method's dependence on the relative positions among the water surfaces and the referenced item makes it inappropriate for severe urban floods; it must be stressed. The accuracy of the model was likely to be compromised once the level of flood water rises over the reference object's height. **Table 3** displays the test set's flood recognition accuracy, recall, and AP. The model's performance in detecting the submerged depths of the reference items was deemed satisfactory, as evidenced by the 90% mean average precision (MAP) of its predictions. The highest accuracy of the flood depth detection from the imaging data set in this work was found in detecting the water level reaching the car's roofs and exhaust pipes. The AP for the predictions in car roof and car pipe categories was 98.5% and 98.6%. The categories "Personal" and "Car-handle" had the second-best performance percentages in identifying flood depth (92.5% and 90.1%, respectively).

Table 3. Performance Analysis of Flood Depth Prediction

Attributes	Precision	Recall	AP
1	96	95	98.1
2	87	97	98.2
3	92	96	99.4
4	95	97.5	97.4
5	95.7	96.5	98.6
6	95	98	99.2
7	92	81	93
Map			98.2

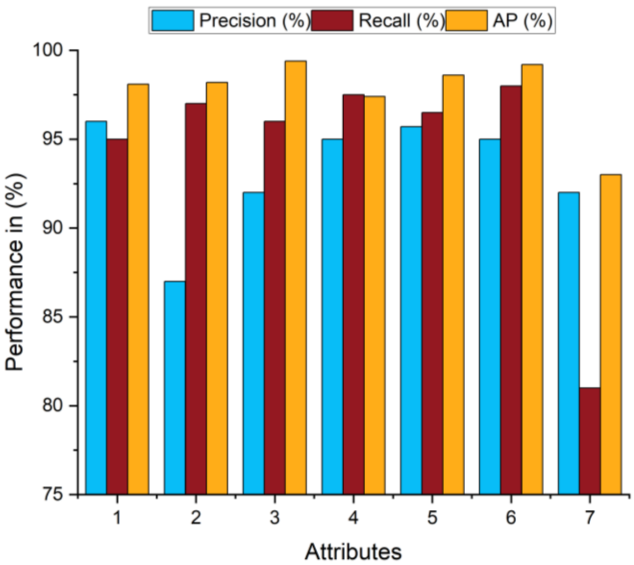


Figure 2. Graphical Plot of Flood Depth Recognition

Figure 2 displays the AP for each category and the number of classified items within each category. There were 2,095 objects in the entire collection. Twenty-one were set aside for testing, while 1,885 were utilized for model validation and training. Among the training and validation set, the "Per-leg" class contained the most significant objects total as 494. Within the test set, sixty-seven objects were classified as "Per-leg." Three hundred forty-two objects were in the validation and training sets, and thirty were in the test sets, making the Car-pipes category's object count smaller than that of the Per-leg category. Compared to the Per-leg category, which had an AP which was 90%, the Car-pipe category had an AP of 98.63%. Furthermore, the group classified as Per-waist had the lowest AP but the fewest objects overall. The amount of objects and the way their features were extracted also impacted the AP. Other investigations that used the VOC2007 dataset, in which the objects were categorized into 20 groups with different APs, corroborated this claim. In earlier research, more recognizable things (e.g., cars and trains) had comparatively sharper contours and more regular shapes. According to [25], on the other hand, things with irregular shapes, like humans, dogs, and cats, have lower APs. The study's findings are in line with this. The flood identification AP was higher if an automobile was utilized as the reference objects for the submerged depths of the flood than when the people’s body was utilized as a reference to be submerged. This was so that information in images may be extracted more readily for items with regular shapes and distinct contours, which makes them easier to identify.

4.3. Performance evaluation

Monitoring the variations in the maps and loss in the learning process utilizing these various learning approaches allowed for analyzing the regular learning on flood depth identification and frozen learning performances. **Table 4** displays the maps and the loss value at 50 iterations throughout the learning procedure. Regular learning had a greater mean a posteriori (MAP) at 200 epochs than frozen learning. Compared to frozen learning, normal learning resulted in less loss. This suggests that frozen learning performed less well than regular learning. Furthermore, the model's performances were further enhanced by utilizing a mixed-learning approach, which utilized frozen learning for the initial ninety epochs before switching to regular learning for the remaining epochs.

Table 4. Performance Evaluation of MAP and Loss Analysis

Epochs	Deep		Regular		Frozen	
	Map	Loss	Map	Loss	Map	Loss
50	94.9	33.4	89.4	31.3	88.3	33.6
100	93.5	33.4	82.3	28.6	89.9	23.8
150	93.7	33.2	83.5	26.4	93.1	18.3
200	93.3	32.4	85.5	24.6	93.3	17.8

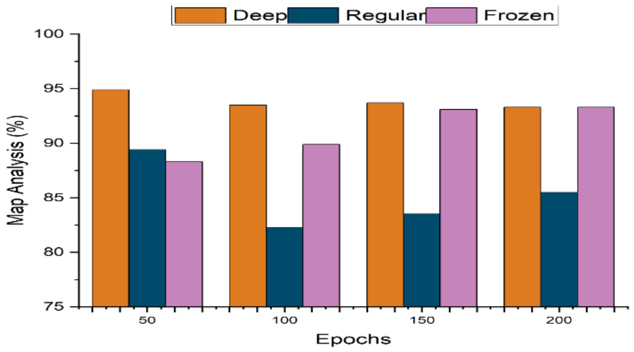


Figure 3. Graphical Plot of Proposed Model vs. Existing Model's Map Analysis

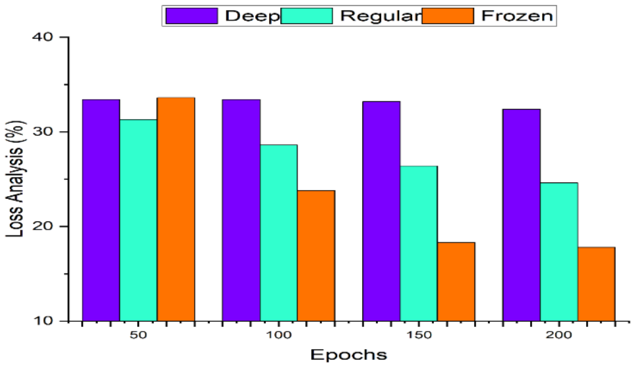


Figure 4. Graphical Plot of Loss Recognition

4.4. Training process

In the deep learning process, three technologies were used for model training and image processing to examine their effects on improving image identification performances for flood depths. These approaches are widely used in image recognition, including cosine annealing, label smoothing, and Mosaic. Four images are combined through cropping and mixing in the mosaic image augmentation technique to exploit the dataset. A regularization method called "label smoothing" adds noise to the labels to reduce overfitting. During the training phase, cosine annealing helps prevent local optima by allowing for bigger learning rate resets, typified by periodic drops and resets in learning rates.

Table 5. mAP Training-Based Average Precision

Scenario	Map (%)
1	87.5
2	84.1
3	90.9
4	90.5
5	91.9
6	97.5
7	96.2

As shown in **Table 5**, the application of several methodologies was examined by correlating the model MAP over multiple applications utilizing eight separate or hybrid methodologies. The model's mean approximate precision (MAP) was only 67.58% in scenario 1 when none of the previously indicated factors were applied. After employing smoothing of labels, Mosaic, and cosine annealing independently, the MAP values increases like 72.1%, 70.94%, & 70.48%. In scenario 2, where Mosaic was the only tool utilized, there was the most critical enhancement above condition 1. However, in situations 5 and 6, where cosine annealing and label smoothing were added to the Mosaic, the MAP increased to 67.5% and 72%, respectively. Once Mosaic was combined with cosine annealing and label smoothing, instead of being applied alone, the MAP dropped. Mosaic alone proved the most effective way to improve the maps for flood depths identification in this investigation. This matches earlier studies with YOLOv4, where it was demonstrated that mosaicking improved image identification accuracy.

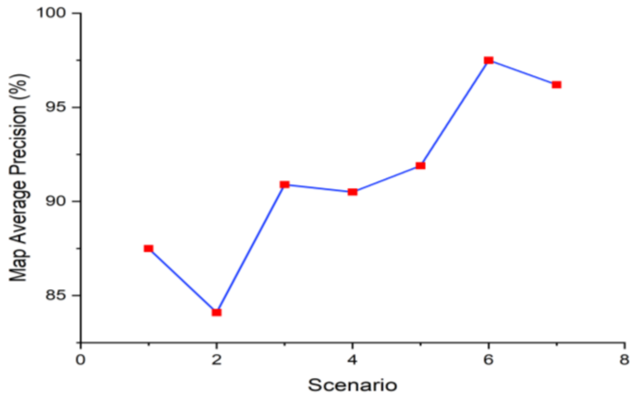


Figure 5. Graphical Plot of mAP-based Training Samples

Figure 6. Prediction Outcomes of an Image

In FRNet networks, dense CNN is the fundamental convolutional block. The activation function, normalization, and convolution are its three main parts. The accuracy and computing speed of deep learning models are significantly impacted by the activation function, which is crucial. This study measured the frames per second (FPS) and MAP, where a high FPS indicates a quick image recognition, to assess the effects of two activation operations, LeakyRelu and Mish.

Table 6. Map Analysis Based on FRNet

	Map										
Epochs	20	40	60	80	100	120	140	160	180	200	220
Dense CNN	85.5	82	83	80	88	93	92	90.3	85	94	90
Intermediate layers	75	72	72	77	73	82	82	81	81	82	84

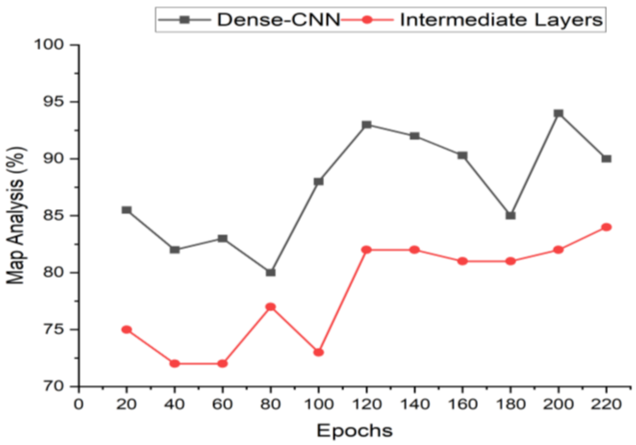


Figure 7. Graphical Plot of mAP-based FRNet Model

Table 6 presents the findings which is pictorially shown in **Figure 3** to **Figure 7**. When comparing the FRNet model with CNN to the one with LeakyReLU, the former usually produced greater mAP values. On the other hand, the model with Mish had a lower FPS (30.49) than LeakyRelu. This suggests that Mish performed image recognition tasks more slowly. These results are consistent with other studies that found Mish to have higher accuracy but a more significant drop in identification speed when compared to LeakyReLU. Notwithstanding these variations, it is important to note that the activation functions show remarkable prediction speeds, with an average of 32.8 and 28.6 ms per images for LeakyReLU and Mish.

The proposed research provides significant advantages in urban flood monitoring by utilizing deep learning and image recognition to provide real-time, accurate flood predictions without requiring additional infrastructure. The model's ability to process images from existing traffic cameras and identify flood depths with high precision through techniques like k-means clustering and Dense CNN makes it an efficient and cost-effective solution. However, the research has limitations, including potential challenges in generalizing the model across diverse urban environments with varying infrastructure and weather conditions. Additionally, the model's reliance on image data may be affected by poor visibility or obstructions, which could hinder its performance in certain scenarios.

5. Conclusion

This research developed a flood-based image identification technique that uses training techniques in integration with the CNN deep learning network to detect urban flood overflow. Vehicles and people submerged in flood waters were utilized as reference points in the image to determine the depth of the floodwaters. The model divided images with measurable flood water levels into four groups. The outcomes showed that the FRNet model can recognize flood depths with up to 99% accuracy and an astounding 30.49 frames per second (FPS) recognition speed. It was also discovered that the kind of reference object submerged impacted the accuracy of FRNet flood depth identification. When automobiles were used as reference items instead of individuals, recognition accuracy was greater. This is because vehicles have set places and regular shapes, which make identification easier. The study also investigated how different transfer learning strategies affected the ability to recognize flood depth. It was discovered that using the proposed approach, which included regular learning and frozen learning, produced superior outcomes compared to using either method separately. Additionally, improving the model's recognition accuracy using Mosaic technology for image augmentation was successful. FRNet model determines the depth of flood water from images with great accuracy and quick processing times. This feature makes it possible for the model for using video footage or images from current traffic cameras to determine flood levels during storms, making it easier to collect continuous, on-location observations of the water depth. The information obtained on flood depths using this method can be used for two things: first, it can help validate and adjust urban hydrological models; second, it can give timely information about flooding that is crucial for early warning systems in urban areas. In the future, the work intended to construct DL approaches that retrieve temporal as well as spatial data through Temporal Convolutional Networks (TCNs)/Convolutional long-short term memory (C-LSTMs). The suggested framework can process various data based on weather, satellites, and water levels in the rivers in order to predict floods earlier. Employing the relevant dataset in the DL strategy, further enhancement in research can facilitate precise detection, better precaution activities, and enhanced flexibility.

References

Arshad B., Ogie R., Barthalemy J., Pradhan B., Verstaevel N. and Perez P. (2019). Computers visions and IoT-based sensor in floods monitoring and mappings: A systematic review, *Sensors*, **19**(22), 5012.

Arun Mozhi Selvi Sundarapandi, Sundara Rajulu Navaneethakrishnan, Hemlathadhevi A, Surendran Rajendran*, "A Light weighted Dense and Tree structured simple recurrent unit (LDTSRU) for flood prediction using meteorological variables", *Global NEST Journal* **26**(8), 10*.* Available at: https://doi.org/10.30955/gnj.06242

Babu T, Raveena Selvanarayanan, Tamilvizhi Thanarajan and Surendran Rajendran (2024) "Integrated Early Flood Prediction using Sentinel-2 Imagery with VANET-MARL-based Deep Neural RNN", *Global NEST Journal,* **26**(10), 9. Available at: https://doi.org/10.30955/gnj.06554.

Bai Y., Zhao N., Zhang R. and Zeng X. (2018). Storm water managements of low impacts developments in urban area based on SWMM, *Water*, **11**(1), 33.

Basnyat B., Roy N. and Gangopadhyay A. (2018). A flash flood categorizations system using scenes-texts recognitions, *In 2018 IEEE International Conferences on Smart Computing (SMARTCOMP)*, 147–154.

Bochkovskiy A., Wang C.Y. and Liao H.Y.M. (2020). Yolov4: Optimal speed and accuracy of objects detections, *ArXiv Preprint ArXiv:2004.10934*.

Bulti D.T. and Abebe B.G. (2020). A review of flood modelling method for urban pluvial floods applications, *Modeling Earth System and Environments*, **6**, 1293–1302.

Chia M.Y, Koo C.H, Huang Y.F, Di Chan W. and Pang J.Y. (2023). Artificial intelligences generated synthetic dataset to remedy waters quality index estimations data scarcity, *Water Resources Managements*, **37**(15), 6183–6198.

Faramarzzadeh M., Ehsani M.R., Akbari M., Rahimi R., Moghaddam M., Behrangi A., Klöve B., Haghighi A.T. and Oussalah M. (2023). Applications of machine learning and remotes sensing for gaps-filling daily precipitation data of a sparsely gauged basins in East Africa, *Environmental Processes*, **10**(1), 8.

Feng B., Zhang Y. and Bourke R. (2021). Urbanization impact on floods risk based on urban growths data and coupled flood model, *Natural Hazards*, **106**(1), 613–627.

Gauen K., Dailey R., Laiman J., Zi Y., Asokan N., Lu Y.H., Thiruvathukal G.K., Shyu M.L. and Chen S.C. (2017). Comparisons of visual dataset for machine learning, *In 2017 IEEE International Conferences on Information Reuses and Integrations (IRI)*, 346–355.

Guo K., Guan M. and Yu D. (2021). Urban surfaces water flood modelling comprehensive reviews of current model and future challenge, *Hydrology and Earth Systems Science*, **25**(5), 2843–2860.

Hammond M.J., Chen A.S., Djordjević S., Butler D. and Mark O. (2015). Urban floods impact assessments: A state-of-the-art review, *Urban Water Journal*, **12**(1), 14–29.

Ichiba A., Gires A., Tchiguirinskaia I., Schertzer D., Bompard P. and Ten Veldhuis M.C. (2018). Scale effects challenge in urban hydrology highlighted with a distributed hydrological models, *Hydrology and Earth Systems Science*, **22**(1), 331–350.

Jiang J., Liu J., Cheng C., Huang J. and Xue A. (2019). Automatic estimations of urban waterlogging depth from video image based on ubiquitous referenced object, *Remote Sensing*, **11**(5), 587.

Jiang J., Qin C.Z., Yu J., Cheng C., Liu J. and Huang J. (2020). Obtaining urban waterlogging depth from video image using synthetic images data, *Remote Sensing*, **12**(6), 1014.

Kankanamge N., Yigitcanlar T., Goonetilleke A. and Kamruzzaman M. (2020). Determining disasters severity through social media analysis: Testing the methodology with South East Queensland Flood tweet, *International Journal of Disasters Risks Reductions*, **42**, 101360.

Li J., Zhang B., Mu C. and Chen L. (2018). Simulations of a sponge city's hydrological and environmental effect based on MIKE FLOOD, *Environmental Earth Science*, **77**, 1–16.

Li Y., Martinis S. and Wieland M. (2019). Urban floods mapping with an active self-learning convolution neural network based on TerraSAR-X intensity and interferometric coherences, *ISPRS Journal of Photogrammetry and Remote Sensing*, **152**, 178–191.

Mignot E., Li X. and Dewals B. (2019). Experimental modelling of urban flooding: A review, *Journal of Hydrology*, **568**, 334–342.

Misra D. (2019). Mish: A self-regularized non-monotonic neural activation function, *Arxiv Preprint Arxiv:1908.08681*.

Nigussie T.A. and Altunkaynak A. (2019). Modelling the effects of urbanizations on flood risks in Ayamama Watershed, Istanbul, Turkey, using the MIKE 21 FM model, *Natural Hazards*, **99**, 1031–1047.

Nirmal Kumar, M., Subramanian, P. and Surendran, R. (2025) “Multivariate time series weather forecasting model using integrated secondary decomposition and Self-Attentive spatio-temporal learning network”, *Global NEST Journal,* **27**(4), 11. Available at: https://doi.org/10.30955/gnj.06195.

Park S., Baek F., Sohn J. and Kim H. (2021). Computer visions–based estimations of flood depths in flooded vehicles image, *Journal of Computing in Civil Engineering*, **35**(2), 04020072.

Prajul, M., Subramanian, P. and Surendran, R. (2025) “Air Pollution Monitoring System using Stacked Attentional Vectormap Convolutional Bidirectional Network with Bobcat Optimization and IoT-Cloud”, *Global NEST Journal,* **27**(3), 12. Available at: https://doi.org/10.30955/gnj.06937.

Rangari V.A., Umamahesh N.V. and Bhatt C.M. (2019). Assessments of inundation risks in urban flood using HEC RAS 2D, *Modeling Earth System and Environments*, **5**(4), 1839–1851.

Rong Y., Zhang T., Zheng Y., Hu C., Peng L. and Feng P. (2020). Three-dimensional urban floods inundations simulations based on digital aerial photogrammetry, *Journal of Hydrology*, **584**, 124308.

Singh P., Sinha V.S.P., Vijhani A. and Pahuja N. (2018). Vulnerability assessments of urban road networks from urban floods, *International Journal of Disaster Risks Reductions*, **28**, 237–250.

Suresh Subramanian, Geetha Rani K, Maheswari Madhavan, Surendran Rajendran “An Automatic Data-Driven Long-term Rainfall Prediction using Humboldt Squid Optimized Convolutional Residual Attentive Gated Circulation Model in India”. *Global NEST Journal,* **26**(10), 11. Available at: https://doi.org/10.30955/gnj.06421.

Wang C.Y., Bochkovskiy A. and Liao H.Y.M. (2021). Scaled-yolov4: Scaling cross-stages partial networks, *In Proceeding of the IEEE/Cvf Conferences on Computer Visions and Pattern Recognitions*, 13029–130130.

Wang X., Yang M., Zhu S. and Lin Y. (2013). Regionlets for generic objects detections, *In Proceeding of the IEEE International Conferences on Computer Visions*, 17–24.

Wang Y., Chen A.S., Fu G., Djordjević S., Zhang C. and Savić D.A. (2018). An integrated framework for high-resolutions urban floods modelling considering multiple information source and urban feature, *Environmental Modelling & Software*, **107**, 85–95.

Yin J., Ye M., Yin Z. and Xu S. (2015). A review of advance in urban floods risks analysis over China, *Stochastic Environmental Research and Risks Assessments*, **29**, 1063–1070.

Yu H., Zhao Y., Fu Y. and Li L. (2018). Spatiotemporal variances assessments of urban rainstorms waterlogging affected by impervious surface expansions: A Guangzhou, China case study. *Sustainability*, **10**(10), 3761.

Zou Z., Shi Z., Guo Y. and Ye J. (2019). Objects detections in 20 years: A survey, *ArXiv Preprint ArXiv:1905.05055*.