International Journal On Engineering Technology and Sciences – IJETS™
ISSN(P): 2349-3968, ISSN (O): 2349-3976
Volume X, Issue III, March - 2023

Image change detection using a multi-scale feature progressivefusion network for remote sensing

I.Jeril Samson
Computer Science Engineering
Francis Xavier Engineering College
Tirunelveli-Tamil Nadu-India
jerilsamsoni.ug19.cs@francisxavier.ac.in

D.Jebaraj Samuvel
Computer Science Engineering
Francis Xavier Engineering College
Tirunelveli-Tamil Nadu-India
jebarajsamuveld.ug19.cs@francisxavier.ac.in

M.Guroosuntar
Computer Science Engineering
Francis Xavier Engineering College
Tirunelveli-Tamil Nadu-India
guroosuntarm.ug19.cs@francisxavier.ac.in

Dr.R.Ravi
Professor / Dept. of Computer Science and Engineering
Francis Xavier Engineering College
Tirunelveli - Tamil Nadu - India
dr.r.ravi@francisxavier.ac.in

J.Priskilla Angel Rani
Assistant Professor / Dept. of Computer Science and Engineering
Francis Xavier Engineering College
Tirunelveli - Tamil Nadu - India
priskillaangelrani@francisxavier.ac.in

Abstract

Deep learning-based change detection (CD) algorithms are currently a popular area in study. To gradually fuse semantic features in CD tasks, feature pyramid networks (FPNs) are frequently utilised. Existing FPN-based CD approaches, however, are unable to precisely pinpoint the boundaries of the change zone and cannot detect the entire change region. The Layer Feature Fusion Module (LFFM), Multi-Scale Feature Aggregation Module (MSFA), and Multi-Scale Feature Distribution Module of the new Multi-Scale Feature Progressive Fusion Network (MFPF-Net) are suggested as solutions to these issues (MSFD To be more precise, we combine

deeper semantic information with effective change region

representation by first concatenating the characteristics of each layer retrieved from

The bi-temporal images with their difference maps. In order to increase effective communication and complete feature map fusion in CD and to eliminate the influence of indirect information, the produced change maps of each layer are then directly aggregated.

After layering the aggregated feature maps once more using pooling and convolution operations, layers are fused from low to high using a feature

fusion strategy with a pyramid structure to obtain richer contextual information. As a result, each layer of the layered feature maps contains both its own unique semantic information and semantic features from other layers. To test the efficacy of the technique, we conducted extensive experiments on three benchmark datasets that are available to the public: CDD, LEVIR-CD, and WHU-CD. The experimental findings reveal that the method in this study performs better than other comparable methods.

1. Introduction

According to S. Raja Ratna, R. Ravi, and Beulah Shekhar's (2013) proposed the enhanced detection methods of three types of denial-of-service attacks. The methods are selective forwarding assault, pollution attack, and jamming attack [1]. Although FPN-like feature fusion models have achieved remarkable results in the field of computer vision, they still have some shortcomings. M. D. Amala Dhaya and R. Ravi (2021) introduced the approach, which eliminates nodes based on the backward trust score after detecting the presence of a botnet. Their suggested algorithm enhances botnet detection performance and lessens the incidence of money laundering [2]. S. Raja and R. Ravi (2015) presented a system that considerably increases network throughput while reducing jamming throughput, as well as identifying misbehaving nodes with higher detection rate and lower false positive [3].target feature information from the input of the model by introducing a dynamic convolution module and utilizes multi-layer supervision to train the network. In addition, the model focuses on learning both high- level and low-level feature

precision, and recall [4]. M. D. Amala Dhaya and R. Ravi (2021) introduced the approach, which eliminates nodes based on the backward trust score after detecting the presence of a botnet. Their suggested algorithm enhances botnet detection performance and lessens the incidence of money laundering [5]. On the other hand, in the final stage, P. Mano Paul and R. Ravi (2018) suggested applying feature probability to the clustered email, which results in a minimal detection time. Additionally, the CVRS system achieves high accuracy by confirming the reporter's feedback result and reducing the amount of false positives and negatives by calculating similarity detection on the clustered email [6]. A. Agnes, M. Bala Santhiya, V. K. Supriya Banu, and R. Ravi (2021) their idea refers to two frames. The computer vision technique known as Open CV helps with image processing and other motion prediction systems [7]. D. Priyadharshini and R. Ravi (2020) noted that there has been a late development in natural language processing. The deep learning research is still being conducted [8].

D.Priyadharshini, R.Malliga@pandeeswari, S. Shargunam, and R. Ravi (2020) introduces several image modification techniques, their use, and monitoring technologies [9]. D. Priyadharshini and

R. Ravi (2020) noted that there has been a late development in natural language processing. The deep learning research is still being conducted [10]. On the one hand, as mentioned in the paper33, in the pyramid feature fusion structure, the deep feature information is transferred to the shallow features layer by layer. Therefore, the proportion of feature information carried by deep features in the entire fusion continues to decrease. low-level features containing rich spatial information are introduced into the network, obtaining predicted change maps with boundary information. Finally, the predicted change maps may have ambiguous object boundary information or the predicted foreground information with too much weight .To solve the above problems. The model adaptively learns Some CD models10, 17, 34 adopt the attention mechanism to improve network performance. These models first extract the rich information in the feature map and then gradually integrate the contextual feature information. Although these models ensure rich information fusion, these models still have problems similar to FPN-like.).

In this article, a novel Multi -Scale Feature Progressive Fusion Network (MFPF-Net) is proposed for remote sensing image CD, which aims to fully fuse bi -temporal remote sensing images, exchange feature information, promote information propagation and achieve better detection results. In MFPF-Net, the features extracted from the backbone network are first fed into their idea refers to two frames. The computer vision technique known as OpenCV helps with image processing and other motion prediction systems which integrates the pre- and post-change feature information, thus improving the network's ability to identify the changed regions. The MFPF-Net also contains a Multi-Scale Feature Aggregation Module (MSFA), which adaptively assigns weights to the information in the features at different stages and allows communication between the different stages.

2. Related Works

The existing CD methods can be roughly classified into traditional methods and deep learning-based methods, and each will be briefly introduced in the following sections.

Traditionalmethods.

Traditional methods need to manually set parameters or thresholds when using remote sensing images for CD. Artificially designed features only achieve good results in specific scenes. The features designed manually by a priori knowledge are not representative and have poor generalization performance, which makes it impossible to achieve good results in CD of high-resolution remote sensing images, and it is not suitable for CD in complex scenes. Traditional CD methods can be divided into pixel-level CD methods

The pixel-level methods2, 3, 24, 25 are mainly suitable for remote sensing images with medium and low resolution. These methods mainly calculate the difference of the corresponding pixel values and obtain the change maps based on these differences by simply setting the threshold or clustering. Because the fusion of contextual information is ignored in the model, it may cause the model to extract deep features while ignoring shallow information. Simultaneously, with the development of a series of high-resolution optical sensors, high-resolution images contain a wealth of information. The object-level CD methods are proposed for high-resolution CD. The object-level methods4,5,24 divide the image into objects and then compare and analyse the objects in the bi-temporal image by extracting rich geometric information and

spectral information in the images various brain organizations and the normal of the outcomes portraying on the off chance that a gun was available or not. Quicker RCNN model used to identify the web-based entertainment picture information for guns and different weapons, utilizing a two-pass convolutional network. A work was made by to further develop weapon recognition rate in single energy X-beam pictures by utilizing pseudo shading, utilizing different variety channels on the information to attempt to recognize the weapons.

Deep‑learning‑based

Deep‑learning‑based methods. With the rapid development of deep learning technology in computer vision, remote sensing image CD has made great success in accuracy improvement with deep learning. In recent years, deep learning features with rich semantic information have been introduced to replace the low-level manual design features. Some methods use the deep convolution neural network (CNN) as the feature extractor, rather than using the descriptors that require a large number of domain knowledge designed by human beings. Due to its strong detection ability, CNN has successfully achieved great success in remote sensing image CD tasks. The CD methods10–15 use CNN structure to extract rich features from bi-temporal images and obtain the final change map, which has achieved good results. In proposed a synthetic aperture radar CD network, which can generate difference maps with good detection performance. Subsequently, CDNet37 uses an image pair as input, uses the SLAM system of multi-sensor fusion, and combined it with the density 3D reconstruction system to register the video sequence. Finally, the pixel-level structure change maps are obtained. In addition, this paper creates a new urban CD dataset. In this paper, two self-attention modules, BAM and PAM, are proposed. The attention weights of any two pixels at different times and positions are calculated by these two modules, and good results are achieved. DASNet10 captured many discriminative feature representations by missing the dual attention mechanism, which improved the recognition accuracy of the network. Which made innovations for insufficient context feature information fusion. HDFNet32 designed a dynamic fusion network considering the shortcomings of regional integrity detection and introduced a dynamic convolution model for adaptive learning. The network also achieved good performance introduced a CD method from course to fine, which is divided into the coarse detection stage and the fine detection stage. The detection of the two stages can obtain more abundant feature representations, and a mixed loss function is proposed to provide different levels of supervision. Although the above methods have achieved good performance, some features may introduce ambiguous context information for CD. How to obtain effective feature representation and fully integrate feature context information has become an urgent problem in CD.

3 Methodology

Research motivation.

At present, there are still some problems in remote sensing image CD that need to be dealt with: (1) High-resolution remote sensing images are rich in spectral and spatial information, but these information have not been fully utilized; (2) most SOTA CD methods are implemented by FPN-like feature fusion structure, in the process of feature fusion, the spatial structure details used to reconstruct the object boundary can only be obtained in the final fusion stage, which makes the change map predicted by these methods have low-quality object boundary or miss detection of small change regions.

The object of this article is to construct a novel remote sensing image CD network, MFPF-Net, to achieve better high-resolution detection performance. The MFPF-Net network can fully and effectively extract the bi-temporal feature information of high-resolution remote sensing images, and allow efficient information communication across multiple levels. It can detect the boundary information of the changing region more clearly, and effectively avoid the missing detection of small regions.

Overview of the proposed MFPF-Net.

The whole MFPF-Net network consists of the backbone network ResNet1840 and three modules LFFM, MSFA and MSFD. The bi-temporal images pairs are fed into a feature extraction network with two weight-shared ResNet18s, and the two images will output two groups of multi-scale feature maps, respectively. Then, the two feature maps with the same scale in both groups are sent together to the corresponding LFFM module for feature fusion. The fused multi-scale feature maps are fed into the MSFA module, which directly aggregates the multi-scale feature maps and then adaptively generates a set of weights to enhance the feature representation of the feature maps. The first Pred0 of the model is output after the MSFA module. After that, the aggregated feature maps updated by the MSFA are further processed by the MSFD module. MSFD uses global pooling at different scales and convolution operations with different convolution kernels to reallocate the aggregated feature maps to the corresponding layers. Finally, the layered features are fused in a top-down fusion method to obtain the second Pred1 of the network. In the training phase, the optimized network parameters are obtained by deep supervision of the model. The process of the MFPF-Net is shown in Algorithm 1. In this section, firstly, we present the overall framework of the proposed network. Then the three main novel modules are specifically discussed. Finally, we provide details of the loss function.

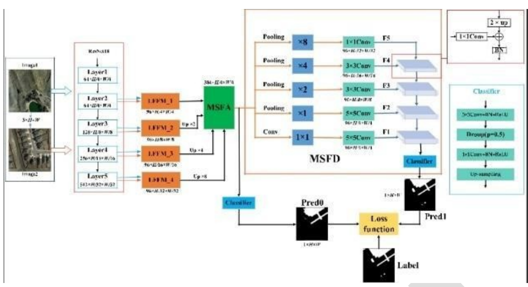


Figure 1. The structure of LFFM

Layer feature fusion module. Selection of Datasets

Present, the feature fusion methods for CD can be divided into two types: pre- and post-fusion shown as fig 1. The pre-fusion indicates that the images obtained after concatenating the bi-temporal image pairs or their difference maps are fed into the network for feature extraction to obtain the change maps, and then CD is performed. The pre-fused images do not adequately represent the high-level semantics of the features extracted from each of the bi-temporal images fed into the network, and the pre-fused images are more sensitive to noise compared to a single original image. The post-fusion is a feature extraction process for bi-temporal images separately using the same backbone network, and then the extracted bi-temporal features are fused and change inference is performed using the CD network. However, during the actual experiments, we found that such fusion does not enable the feature maps to have both high-level semantic information and low-level semantic information.

In this paper, the advantages and disadvantages of these two methods are fully considered. We consider that each layer of the feature map output by the feature extraction network has different semantic information. To make the final change maps better represent the change regions and boundaries, we designed the LFFM, and its structure is shown in Fig. 2. We apply the LFFM to perform the concatenation operation on two feature maps F1 and F2 of the same layer and their difference map. Specifically, the bi-temporal image pairs (Image1 and Image2) are fed into the backbone network ResNet18 to obtain (F11, F21, F31, F41) and (F12, F22, F32, F42), respectively, and then the feature maps corresponding to the two images at each layer and their difference maps are concatenated along the channel dimension.

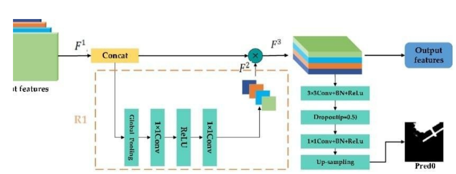


Figure 2. The structure of MSFA

4. Multi-scale feature aggregation

As mentioned in the previous section, FPN-based CD methods15, 25, 32 first extract the rich information in the feature map and then gradually fuse the contextual feature information. These models often produce incomplete

change maps due to gradual dilution of semantics during the progressive fusion shown as fig 2. To enable effective information transfer between the feature maps of different layers of the network, we propose to replace the fusion mechanism in FPNs by aggregating the feature maps of different layers. Specifically, inspired by we design the MSFA that adaptively predicts a set of weights based on the importance of different layer features. The purpose of this design is to effectively enhance the feature representation of the feature map. The structure of MSFA is shown in Fig. 3. The size of the four feature maps outputted by LFFM is H and W denote the height and width of the original image. The four feature maps output from the LFFMs are concatenated along the channel dimension to obtain F1 ∈ R384×H/4×W/4. Subsequently, F1 is compressed by global average pooling through the R1 branch, and then the adaptive weight coefficients F2 ∈ R384×1×1 at different levels are obtained by convolution and activation function, the calculation process is shown in the following equation:

$$F2 = f\, 1\times1[\mathrm{ReLU}\,(f\, 1\times1\ \mathrm{GAP}\ F1\ \ )]$$

Where the GAP denotes the global average pooling, the ReLU is the ReLU activation function, and the f 1×1 is the 1 × 1 convolution layer. From the channel dimension, for different layers, a rich feature representation is obtained by adaptive aggregating the weights of different layers. Then, F2 are multiplied with F1 to obtain a feature map F3 ∈ R384×H/4×W/4 with different weight coefficients for each channel.

4.1 Multi-scale feature distribution module.

Although a rich feature representation as well as better detection results can be obtained after processing by the MSFA module, the prediction results are still unsatisfactory by using this single-stage inference. Most of the existing fusion approaches for CD networks are based on a series of improvements of the FPN model. However, these methods tend to result in Inaccurate localization of change regions or poor change region boundaries, since the high-level features captured by deeper layers may be gradually diluted, and the low-level features learned from shallow layers are insufficient to detect precise change regions throughout the progressive feature fusion.

In this study, we propose to combine the aggregated feature map with the progressive structure of FPN again. Since the MSFA module performs feature aggregation for different stages, the feature map after aggregation can contain feature representations of different stage information in the backbone network. Progressive fusion of FPNs on top of this will significantly alleviate the limitations of the fusion approach with FPNs alone. Therefore, we design an MSFD module to allocate multi-layer features by multi-scale pooling. This enables the semantic and positional information in the feature maps to be fully accessible at each level, which contributes to the feature fusion in FPN and facilitates the model to detect the exact region of change. The feature maps after the MSFA module combine multi-scale feature representation, and the MSFD in this paper assigns multi-level features by multi-scale pooling of aggregated features. The model structure of MSFD is shown in Fig. 1, from which it can be seen that the feature map F1 is concerned with the details of objects in the image, and the ability to identify the boundary of the change region can be enhanced by extracting the detail information in the feature map. The prediction map Pred1 output from the MSFD structure contains rich semantic information and edge detail information, which helps to better detect change regions. Meanwhile, as shown in Fig. 2 of the experimental part of this paper, the feature map can effectively detect the edge information of the change region after MSFD processing, which proves the effectiveness of MSFD and the ability of the model to detect the edge information of the change region.

Specifically, the MSFD module first feeds the feature map outputs from MSFA module to the average pooling layers with pyramid down-sampling rates to convert the aggregated features to different scale spaces. As shown in, the down-sampling rates are {8, 4, 2, 1, 1} from top to bottom, and then these five layers are respectively passed through 1×1, 3×3, and 5×5 convolution operations to obtain five feature maps with different sizes and the same number of channels, which, from top to bottom, are sequentially fused by up-sampling features to obtain the prediction map Pred1. It should be noted that at the end of the progressive fusion, a pixel classifier consisting of a convolution layer and Dropout ($p = 0.5$) is used, and the structure of the main role is to change the number of channels of the features and, to some extent, serves to prevent overfitting. By this fusion, since the distributed feature maps at each fusion level simultaneously incorporate semantics and fine details, more discriminative and complementary representations can be well preserved along the progressive fusion path. The fusion effect is thus greatly enhanced for achieving superior performance.

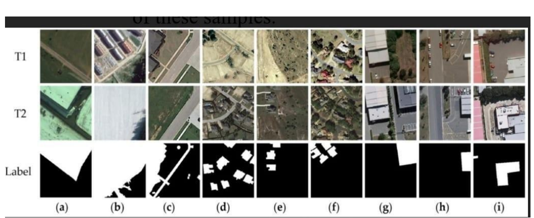
5. Experiments and analysis

In the experiments, we evaluate the effectiveness of the proposed MFPF-Net using three publicly available data-sets. We first introduce the three datasets used in this paper, followed by the evaluation metrics and detailed setup of the experiments. Finally, the experiment is analysed in detail.

Datasets

With the continuous development of remote sensing satellite technology, some high-quality remote sensing CD datasets have emerged in recent years. The publicly available remote sensing image CD datasets are useful for comparing the performance of different CD methods. We conduct experiments on three widely used CD benchmark datasets, including CDD16, LEVIR-CD17, and WHU-CD35.

The CDD dataset was acquired by Google Earth in 2018. It consists of seven pairs of 4725 × 2200 pixels seasonal variation images without appendages and four pairs of 1900 × 1000 pixels seasonal variation images with appendages. The authors of the paper divided the images into non-overlapping 256 × 256 pixels image pairs, and the changing objects include cars, large building structures, etc. The dataset contains 16,000 pairs of seasonal images, of which 10,000 pairs are the training set, and 3,000 pairs are the validation set and test set, respectively. shows some of these samples.



??Bitemporal??remote??sensing images from three open-source datasets CDD, WHU-CD and LEVIR-CD. The T1 are change before images, the T2 are change after images, the label represents the changed areas and the unchanged areas. (a–c) The images in the CDD dataset. (d–f) The images in the LEVIR-CD dataset. (g–i) The images in the WHU-CD dataset (created by 'Microsoft Office Visio 2013'

True value	predicted value	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

Table 1. The detailed explanation of TN, TP, FN, and FP.

6. Evaluation metrics and settings.

In this paper, to compare the difference between the label maps and the predicted change maps, we use four evaluation metrics precision (P), recall (R), overall accuracy (OA), and F1-score (F1) to evaluate the efficiency of the proposed method. In the CD task, the higher P denotes the more accuracy of detected changed pixels and the higher R represents the greater ability of the model to find more changed pixels. OA denotes the overall

International Journal On Engineering Technology and Sciences – IJETS™
ISSN(P): 2349-3968, ISSN (O): 2349-3976
Volume X, Issue III, March - 2023

accuracy. F1 is a metric for measuring the accuracy of the binary classification model. It considers the P and R of the classification model at the same time. The value of F1 ranges from 0 to 1, the higher the value, the better the performance of the model.

P = TP/TP+FP (12)

R = TP/ TP+FN (13)

F1 =2/ P−1+R−1 (14)

OA = TP+TN/TP+FP+TN+FN
(15)

In the above equation, as shown in Table 1, TP is the number of correctly detected changed pixels, TN rep-resents the number of correctly detected unchanged pixels, FP is the number of false alarm pixels, and FN is the number of lost unchanged pixels.

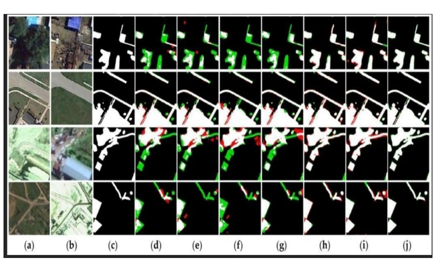


Fig 3. Output samples

Bi-temporal remote sensing images from CDD dataset. The (a) are change before images; the (b) are change after images; the (c) represents the changed areas and the unchanged areas; (d) are results obtained by CD Net; (e) are results obtained by FC-EF; (f) are results obtained by FC-Siam-Diff; (g) are results obtained by FC-Siam Cone; (h) are results obtained by DAS Net; (i) are results obtained by STA Net; (j) are results obtained by MFPF-Net (created by 'Microsoft Office Visio 2013 influence of irrelevant factors such as seasonal changes and illumination on the model to the maximum extent, and accurately detects the edge detail information of changing objects compared with the label.

Next, as shown in Table 2, we evaluate the model performance from the quantitative results of the evaluation indicators. Specifically, on the CDD data set, the R, P, F1, and OA of the method are 96.4%, 95.3%, 95.9%, and 99.0% respectively.

Method	R (%)	P (%)	F1(%)	OA(%)
CDNet	81.7	82.7	82.2	96.4
FC-EF	76.1	81.5	77.1	94.1
FC-Siam-Diff	83.6	85.8	83.7	95.8
FC-Siam-Conc	82.5	84.4	82.5	95.7
DASNet	93.0	92.0	92.5	98.1
STANet	**96.5**	87.0	91.5	97.9
MFPF-Net (ours)	96.4	**95.3**	**95.9**	**99.0**

Table 2. Comparison of CDD dataset results (the best performance is emphasized in bold).

Method	R (%)	P (%)	F1(%)	OA(%)
CDNet	89.1	74.6	81.2	97.1
FC-EF	85.6	76.5	80.8	97.9
FC-Siam-Diff	87.5	79.8	83.5	98.2
FC-Siam-Conc	83.9	81.6	82.7	98.2
DASNet	87.9	81.5	84.6	98.4
STANet	**89.9**	82.6	86.1	98.5
MFPF-Net (Ours)	89.6	**89.9**	**89.8**	**99.0**

Table 3. Comparison of LEVIR-CD dataset results (the best performance is emphasized in bold).

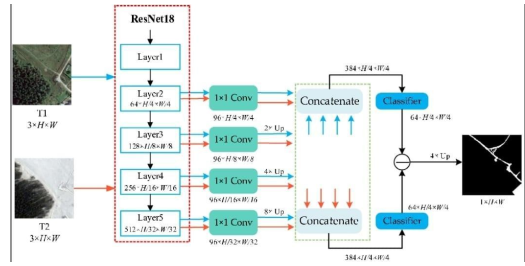


Fig 4.The structure of baseline (created by 'Microsoft Office Visio 2013'.

6. Ablation study

To demonstrate the effectiveness of our proposed method, we performed on three datasets CDD, LEVIR-CD, and WHU-CD a series of ablation experiments shown as figure 5.

Effectiveness of three innovation modules. First, we gradually added each proposed module to the Baseline and finally integrated all the modules, including LFFM, MSFA, and MSFD, together. The detailed structure of the Baseline is shown in. We conducted four ablation experiments on three datasets. There are four experiments: Baseline, Baseline + LFFM, Baseline + LFFM + MSFA, and Baseline + LFFM + MSFA + MSFD. Table 5 shows the results of these four experiments. It can be seen that without adding the three proposed innovative modules, the network performs poorly, with F1 of 89.8%, 82.9%, and 85.3% on the three datasets CDD, LEVIR-CD, and WHU-CD, respectively, which is a huge gap compared to other models that join the innovative mod-ules.

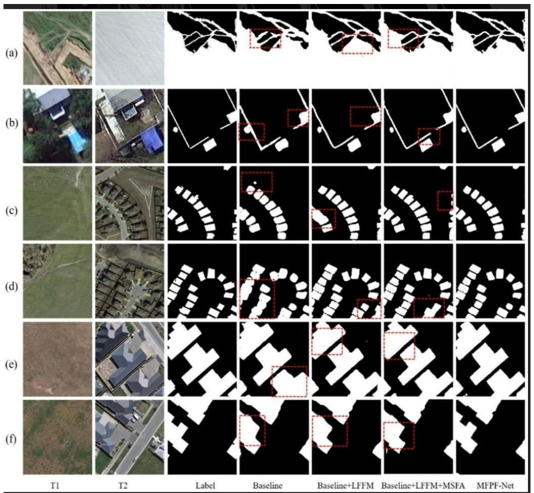


Figure 5. The ablation experimental images from three datasets

The ablation experimental images from three datasets. They should be listed as: (a,b) from CDD dataset; (c,d) from LEVIR-CD dataset; (e,f) WHU-CD .

7. Accuracy/efficiency trade-offs

We first evaluate the performance of the model on the CDD dataset according to the time complexity and spatial complexity. In this article, Table lists the time consumption and model parameters of all methods. T/E represents time/epoch, T/P represents time/parameters. In Table "time/parameters" represents the efficiency of the model. The lower the value, the better the trade-off between time complexity and space complexity. In addition, F1 and OA are selected to reflect the accuracy of the model, which can better explain the comprehensive performance of the model.

CD Net has the fewest model parameters, but the efficiency of the model is poor and its accuracy is also average. FC-Siam-Diff and FC-Siam-Cone achieve similar accuracy, but they are inefficient and time -consuming. FC-EF model has a small number of parameters and the least time consumption of test pictures, but it does not perform well in performance. The parameter quantity and time consumption of STA-Net are increased compared with those of FC series models, but the F1 and OA of STA Net have been greatly improved. Compared with STA Net, DAS Net model has further improved F1 and OA, but it also brings the problem of increasing the number of parameters. In addition, the parameters of MFPF-Net are large, but it only needs about 296 s to generate the change map of the whole test set, which is equivalent to only about 0.098 s to obtain every 256 × 256 change maps, which is acceptable for most CD tasks. At the same time, the proposed MFPF-Net model also achieves good performance on F1 and OA. In conclusion, the efficiency of MFPF-Net is competitive with several SOTA methods.

8. Conclusions

In this paper, a novel deep learning network for remote sensing image CD is proposed, named MFPF-Net. To fully fuse the feature maps of each layer of bi-temporal images by layer, a layer feature fusion module (LFFM) is designed. LFFM emphasizes the fusion of same-layer bi-temporal feature maps and their difference maps, which focuses on the change regions information while retaining some detailed information. We discuss the problems caused by the FPN-like feature fusion mechanism, based on which the MSFA feature fusion mechanism is pro-posed. This mechanism can perform feature aggregation for different stages while generating adaptive parameters to highlight feature information in changing regions. Finally, multi-level pooling operation is performed in the MSFD module and combined with FPN, where the feature maps of each layer have the semantics of the information of other layers, which makes the progressive fusion of multi-scale feature maps more effective. The efficient combination of our proposed three models reduces the information loss during feature extraction and enables the network to achieve SOTA performance. In this study, we analyse some problems in extracting features and then propose a deeply supervised CD network for high-resolution remote sensing images. The proposed network is improved for the problem of feature extraction. The network achieves superior results on three datasets, CDD, LEVIR-CD, and WHU-CD, and also proves the effectiveness and feasibility of the MFPF-Net.

Although MFPF-Net solves to some extent the problems of missed and false detection prevalent in remote sensing image CD, the number of model parameters is large and the MFPF-Net network is based on a deeply supervised strategy, which requires abundant model training time. Further exploration and research on model light-weighting and unsupervised can be carried out in future work.

References

[1] S. Raja Ratna, R. Ravi and Beulah Shekhar, “Mitigating Denial of Service Attacks in Wireless Networks”, International Journal of Advanced Research in Computer Engineering & Technology, vol. 2, no 5, pp. 1716-1719, 2013.

[2] R. Kabilan, R.Ravi, S.Suhirtha, M.Sankara Gomathi, and S.Sofia, “3D object recognition and detection using surf mapping”, International Journal of Emerging Technology and Innovative Engineering, vol. 5, no. 7, pp. 555-561, 2019.

[3] S Raja Ratna and R Ravi, “Securing jammed network using reliability behaviour value through neuron fuzzy analysis”, Indian Academy of Sciences, vol. 40, no. 4, pp. 1139–1153, 2015.

[4] Edwin Raja S and Ravi R, “A performance analysis of Software Defined Network based prevention on phishing attack in cyberspace using a deep machine learning it CANTINA approach(DMLCA)”, Computer Communications, vol. 152, pp.0-6, 2020.

[5] M. D. Amala Dhaya and R. Ravi, “Multi feature behaviour approximation model based efficient botnet detection to mitigate financial frauds”, Journal of Ambient Intelligence and Humanized

Computing, vol. 13, no. 7, pp.799-3806, 2021.

[6] P. Mano Paul and R. Ravi, “A Collaborative Reputation-Based Vector Space Model for Email Spam Filtering”, Journal of Computational and Theoretical Nanoscience, vol. 15, pp. 474–479, 2018.

[7] A. Agnes , M. Bala Santhiya , V. K. Supriya Banu, and R Ravi, “Automated Detection And Alert For Animal Intrusion In Agri Farm Fields”, International Journal of Advanced Research in Management, Architecture, Technology and Engineering, vol. 7, no.4, pp. 9-15, 2021.

[8] D. Priyadharshini, and R. Ravi, “Deep learning: a survey and techniques for language processing, image, speech and text”, Francis Xavier Journal of Science Engineering and Management, vol. 1, no. 1 , pp.11-14, 2020.

[9]D.Priyadharshini,R.Malliga@pandeeswari, S. Shargunam, and R. Ravi, “Image processing: a comprehensive survey and perspective on recent works”, Francis Xavier Journal of Science Engineering and Management, vol.1, no.1, pp.15-17, 2020

[10] D. Priyadharshini, R. malliga@pandeeswari, S. shargunam, and R. Ravi, “Data science: a comprehensive survey and perspective on recent works”, Francis Xavier Journal of Science Engineering and Management, vol.1, no. 1, pp.7-10, 2020.