SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I11P122

Volume 11 Issue 11, 272-285, November 2024


*Original Article*

# Beyond Static: A Satellite Image Change Detection Using Absolute Convolutional Prior Fusion (AC-PF) Approach

Aiswarya Jeevan[1], S Amala Shanthi[2]

[1,2]*Department of Electronics and Communication Engineering, Noorul Islam Centre for Higher Education, Tamil Nadu, India*

[1]*Corresponding Author : Aiswarya.Jeevan.321@outlook.com*

Received: 24 September 2024 Revised: 23 October 2024 Accepted: 21 November 2024 Published: 03 December 2024

***Abstract -*** *Urbanization is a dynamic process marked by rapid and intricate land use, infrastructure, and population distribution transformations. Monitoring these changes for sustainable urban development, resource management, and disaster readiness is imperative. Satellite image Change Detection (CD) has emerged as a potent tool for comprehensively and efficiently evaluating urban changes over time. Techniques for CD facilitate identifying, characterising, and quantifying modifications in land cover, land use, and natural phenomena. This capability is pivotal for environmental monitoring, resource management, and disaster response. Detecting changes in urban expansion, deforestation, agricultural practices, and natural disasters contributes to informed decision-making and sustainable development. Whether it is gradual urban expansion or sudden infrastructural developments, the ability to detect changes offers valuable insights into the patterns and drivers of urban growth. Integrating remote sensing technologies and advanced image processing techniques has remarkably enhanced the accuracy and efficiency of CD in urban environments. These methods enable the identification of land cover changes, such as converting green spaces to built-up areas or adjusting transportation networks. This paper introduces an effective CD model that incorporates a deep learning approach. The proposed architecture is directly inspired by the U-Net model, adapted into the AC-PF while considering the available training data. Finally, the Dice similarity score is computed for a specific image compared to the ground truth images and the corresponding input images.*

***Keywords*** *- Change detection, Convolutional neural network, Deep learning, Image processing, Remote sensing, Satellite imagery, Urban monitoring.*

## 1. Introduction

Remote sensing is vital for obtaining up-to-date information in various applications like land-cover categorization and inspection of agricultural changes. Despite its importance, traditional categorization approaches encounter challenges when processing satellite images. These approaches often make oversimplified hypotheses in their algorithms and neglect potential issues arising from sensor variations, atmospheric impacts, and the radiometric overlap of land-cover objects in interpreted images. As a result, inaccuracies may arise when analysing these images [1]. Remote sensing is a non-contact technique for gathering data about the Earth's surface. This is accomplished by recording the reflectance of different objects on the Earth. Upon the arrival of solar radiation at the Earth's surface, a portion of the energy is taken by the Earth's surface, atmospheric gases, clouds, water vapor, and dust. A portion of this energy is then reflected from both clouds and the Earth's surface. The fundamental principle of remote sensing, illustrated in Figure 1, involves reflecting solar energy from diverse objects on the Earth's surface. Some of this energy is received, while some is scattered. Satellite sensors record the reflected energy from different Earth objects, which is subsequently collected by a satellite receiver, processed, and finally analysed for various applications.

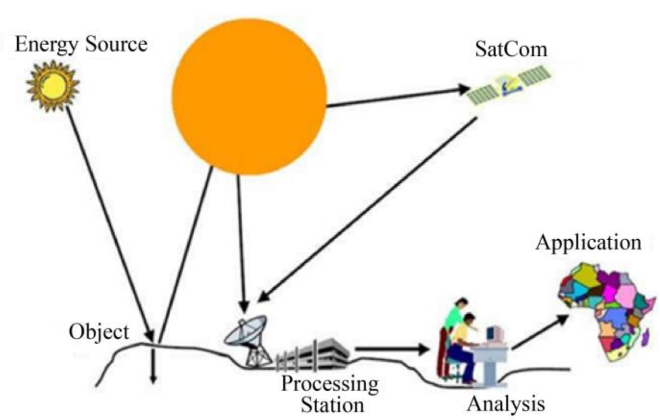



**Fig. 1 Basic principle of remote sensing**

**Fig. 2 Satellite images**

Images from satellites in orbit around the Earth document electromagnetic radiation released or bounced back from the Earth's surface across different wavelengths, encompassing visible light, infrared, and microwave frequencies. Processed satellite images serve diverse purposes and offer several advantages over traditional ground-based observations, such as providing bird's-eye views and enabling tracking changes over time. Governmental organizations, non-governmental organizations, academic institutions, and commercial enterprises widely utilize these images. Moreover, technological advancements have made High-Resolution (HR) satellite images readily accessible to the public through various online platforms. This increased accessibility has facilitated a better understanding of our planet by offering clear and recent global views of the Earth's surface. Figure 2 illustrates examples of satellite images. In Earth remote sensing CD, the process involves recognizing disparities in the Earth's surface. This detection process encompasses observing and evaluating differences to document the spectral and temporal progression.

The CD systematically analyses a pair of remote sensing images snapped over a similar geographical area but at distinct time points. The major objective is to recognize and highlight any alterations that may have arisen between the two attainment dates [2]. This technique, widely applied across various fields, integrates the examination of disparities within a specific area using remote sensing data or images obtained at distinct time intervals [3]. As urban areas continue to experience rapid growth and development, CD methodology, particularly based on remote sensing images, has proven to be highly effective in quantitatively extracting and analyzing these changes [4]. The detection of altered regions within images of similar scenes obtained at distinct times holds significant importance, given its diverse applications across various disciplines [5]. Figure 3 illustrates the general workflow for CD.

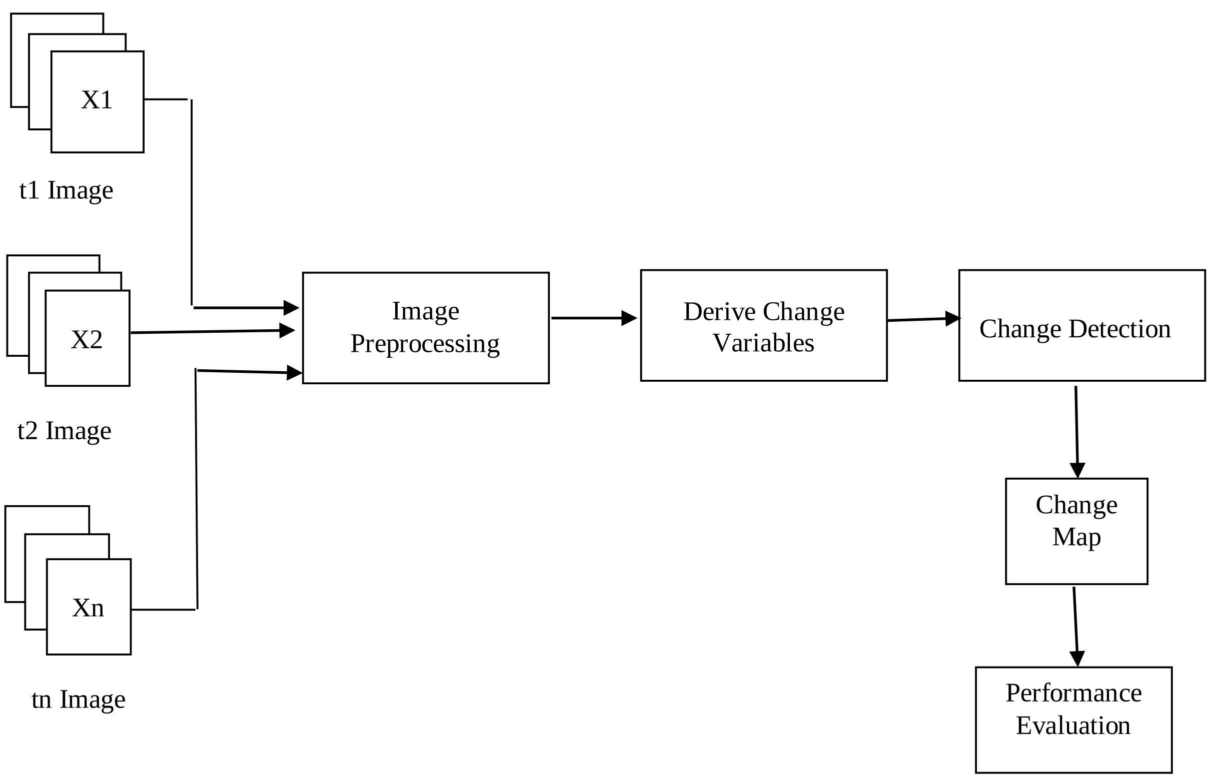



**Fig. 3 General workflow of change detection**

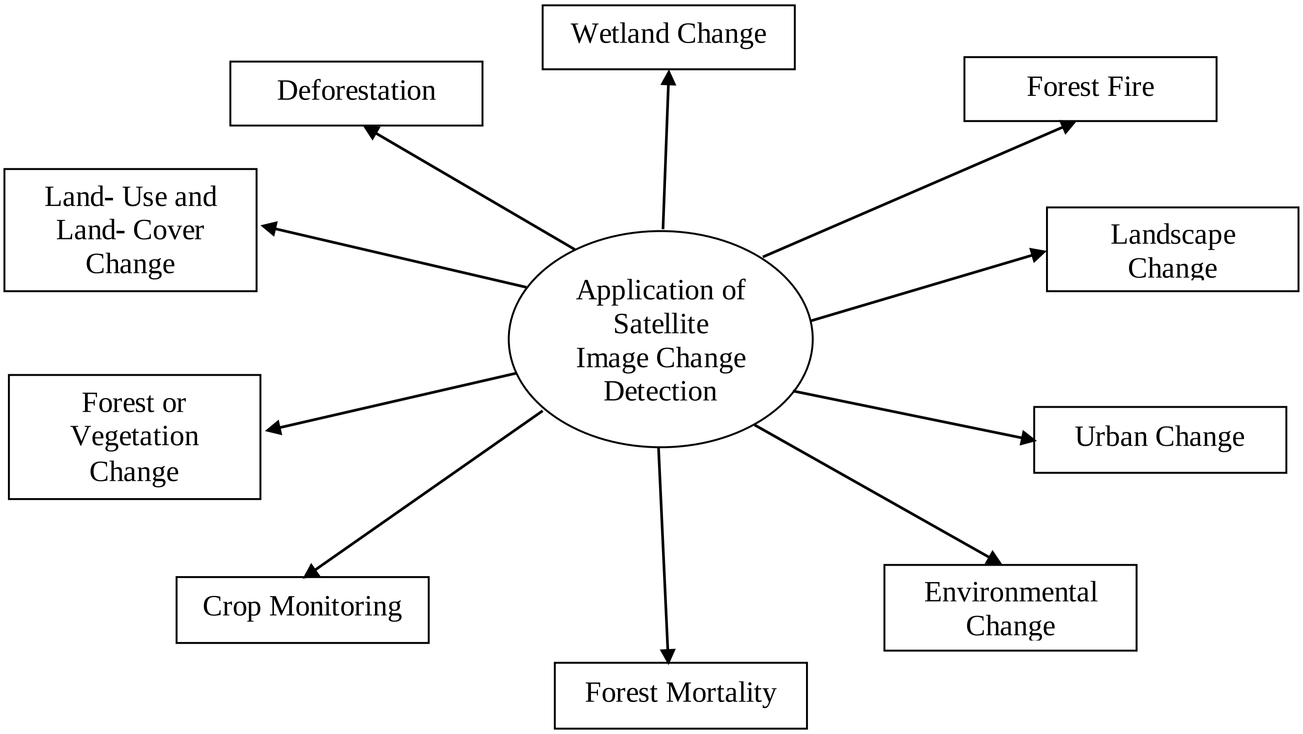



**Fig. 4 Applications of satellite image change detection**

Detection of changes is a significant use of data acquired through remote sensing from satellites orbiting the Earth [6]. The primary objective of employing remote sensor data for CD is to identify and comprehend modifications in features or phenomena of interest, such as deforestation or land use, across two or more sets of image dates [7]. Figure 4 illustrates the diverse range of applications for CD in satellite imagery.

The abundant information within satellite images is crucial in supplying topographical data. Techniques in satellite remote sensing regularly capture data and images, contributing to various environmental applications. These applications address diverse challenges, such as soil quality assessments, water resource studies, meteorological simulations, and environmental protection efforts. However, the substantial volume of satellite imagery data poses a challenge, requiring efficient acquisition and processing methods. The data received at remote sensing data centers is substantial and continually expanding. Consequently, there is a pressing need for a streamlined and effective mechanism to extract and interpret valuable information from the extensive satellite imagery dataset.

Using DL methodologies for CD in satellite images has brought about a paradigm shift in remote sensing and geospatial analysis. This approach, employing Convolutional Neural Networks (CNNs) and other sophisticated DL methodologies, has revolutionized the automated recognition and characterization of significant alterations on Earth's surface over time. The primary strength lies in the model's capacity to comprehend intricate spatial and spectral patterns inherent in satellite imagery, facilitating precise detection of changes such as urban expansion, deforestation, or natural disasters. Incorporating DL not only elevates the accuracy and efficiency of CD but also diminishes the need for manual interpretation. This renders it an indispensable tool for monitoring large-scale environmental transformations, providing valuable support to decision-makers in addressing environmental challenges. So, an effective satellite image CD model was proposed in this paper.

## 2. Literature Review

Ebrahim Ghaderpour and Tijana Vujadinovic [8] introduced a resilient jump detection technique that relies on anti-leakage least-squares spectral assessment and suitable temporal segmentation. This approach enables the concurrent exploration of trends and statistically notable spectral elements in each time series segment, allowing for the identification of potential jumps. The approach incorporates appropriate weights for the time series, enhancing its ability to discern jumps effectively. The simulation results indicated that the suggested approach surpasses existing Trends in recognizing jumps within the trend component of time series across different types.

Jie Chen et al. [9] developed an approach for CD in HR images, employing dual attentive, fully convolutional Siamese networks. The dual attention approach is integral in capturing extensive dependencies, improving the potential of the model to discern features more efficiently and thereby improving detection performance. Weighted double-margin contrastive loss is introduced to differentiate between altered and unaltered feature pairs. Experimental results on both a CD dataset and a building CD dataset illustrated the efficacy of the suggested method. Compared to various baseline methods, the new approach achieved significant improvements, demonstrating maximum enhancements of 2.9% and 4.2% in the F1 score for the respective datasets.

Xiuwei Zhang et al. [10] introduced a proficient network for detecting changes in satellite images. The innovative differential change network module enhances the extraction of change features. Multiple side output features are employed to accommodate the diverse scales of altered regions. Addressing the imbalance in the distribution of changed and unchanged data samples and facilitating challenging sample mining is achieved through the application of focal loss. The assessment of the suggested approach on two publicly accessible change detection datasets showcased its outstanding performance compared to current methodologies.

Caijun Ren et al. [11] developed a Generative Adversarial Network (GAN)-based technique for recognizing changes in satellite images. The GAN framework is specifically crafted to generate co-registered images representing similar landscapes but with variations in the changed regions. The defined objective functions include Lipschitz constraints to ensure that gradients do not vanish during training. A straightforward comparison strategy is employed to derive the final change map, utilizing the outputs of the optimized generator. The proposed method demonstrated effectiveness in generating authentic images and yielded compelling results in unsupervised CD by effectively addressing the challenges associated with unregistered pixels.

Ekaterina Kalinicheva et al. [12] presented an innovative method for recognizing and clustering satellite image time series changes. The process begins with creating bitemporal change masks for each set of consecutive images using neural network autoencoders. Following this, the recognized changes are associated with individual spatial objects. Spatial entities sharing identical geographical coordinates are consolidated into spatiotemporal evolution graphs. Subsequently, these graphs are clustered according to the nature of the change process using a model based on Gated Recurrent Unit (GRU) autoencoders. The effectiveness of this methodology was assessed utilizing real-world SITS data, yielding encouraging outcomes.

Xuan Hou et al. [13] developed a framework known as the High-Resolution Triplet Network (HRTNet) to address limitations in CD. The framework incorporates a dynamic inception phase designed to overcome challenges in this domain. Initially, a distinctive triplet input network was devised to learn attributes from bi-temporal images and capture temporal discrepancies between images across time. Then, a network was utilized to retrieve fine-grained attributes from HR images and preserve complex features with the least amount of information loss. In order to increase HRTNet's ability to express features and boost the representation of multi-scale data in the extracted features, the study included a dynamic inception component. Finally, an accurate change map was produced by examining the distances between pairs of attributes. Three widely used HR remote sensing image datasets were utilized to assess the effectiveness and reliability of HRTNet. The results of the systematic experiments showed that the suggested strategy outperformed the most advanced change detection techniques.

Hichem Sahbi et al. [14] presented an innovative active learning algorithm for CD in satellite images. The method employs a question-and-answer model, suggesting the most informative pairs of patches to an oracle. The decision criterion is adapted based on the oracle's feedback to align with the oracle's intent. The recommended display approach is probabilistic and is determined by reducing a constrained blend of objective functions. Experimental assessments carried out on the demanding task of satellite image CD showcased the reliability of the suggested method.

Xin Huang et al. [15] introduced an automated model for CD, with a primary focus on identifying recently developed building areas employing time-series multi-view ZY-3 HR satellite images. The central aim of this proposed approach is to assess the possibility of determining the timing of changes in these areas. The method comprises three key components and undergoes evaluation on two distinct image sets. The findings of the experiments affirm the effectiveness of the suggested methodology.

Xi et al. [16] developed an innovative spatiotemporal unit and a novel segmentation approach to retrieve and depict geographic objects within a coherent spatiotemporal framework. The study tackled four key technical challenges: ST-cube modeling, segmentation, scale, and evaluation. The suggested techniques' efficacy was validated using 18 Landsat-8 images through experimentation. The findings underscored the efficiency and dominance of ST-cubes in portraying geographic objects and examining spatiotemporal features when compared to current pixel- and object-based methodologies.

Kaiyu Li et al. [17] developed a fully convolutional siamese framework designed for CD. Three siamese networks are integrated with UNet++ to investigate how these structures influence the CD task when coupled with a robust backbone network featuring strong feature extraction capabilities. The experimental outcomes demonstrated significant enhancements across various metrics. Notably, the proposed approach outperforms other methods for CD, showcasing superior performance.

Ran Jing et al. [18] introduced an innovative approach for CD employing Very High-Resolution (VHR) satellite imagery. The recommended framework integrates multi-scale SLIC-CNN and SCAE attributes to effectively address the inherent challenges of CD using VHR satellite images. Combining multiscale, spectral, geometric, textural, and deep structural features, the suggested method leverages the self-learning SCAE framework as a feature extractor to boost the representation of ground objects in the acquired images. Through controlled experiments, it was observed that CD results become more fragmented without image segmentation, leading to increased false-positive and false-negative rates.

Yi Zhang et al. [19] developed an approach for CD in VHR satellite images, employing a boundary-aware approach coupled with a hybrid loss. The suggested approach adopts a coarse-to-fine framework, integrating a higher-level feature-guided coarse recognition and a refined residual recognition to detect changes in pairs of images. Furthermore, a hybrid loss is employed to oversee training at pixel, patch, and map levels, ensuring stability throughout the training process and accommodating scenarios. The experimental findings showcased the superior performance of the proposed model when compared to alternative approaches.

Ahram Song et al. [20] designed an innovative CD approach based on objects to identify alterations in VHR satellite imagery through the utilization of DL networks, eliminating the need for ground truth data. This methodology began by creating an initial pixel-based Change Detection (CD) map using diverse unsupervised CD techniques. Following this, the map underwent a reconstruction process to shape CD objects, which were further identified and refined based on the level of uncertainty. Experimental evaluations carried out on the Worldview-3 and KOMPSAT-3 datasets verified that the suggested approach surpassed conventional CD methods, showcasing enhanced performance.

Lukas Kondmann et al. [21] introduced an unsupervised approach for detecting changes in optical satellite images featuring medium and HR. This approach depends on spatial context by viewing a pixel as a linear combination of its remote neighbors. The model employs this concept to examine differences between the pixel and its predictions based on spatial context over successive intervals, aiding in Change Detection (CD). CD based on spatial context is combined with ensembling techniques applied to mutually exclusive neighborhoods to improve precision. The suggested method exhibited strong performance in change detection, especially when applied to medium-resolution Sentinel-2 and HR Planet Scope imagery across four diverse datasets.

Ramen Pal et al. [22] introduced an innovative approach to segmenting VHR multispectral images, utilizing a variable-length multi-objective NSGA-II algorithm. The algorithm produces a cluster of solutions that are near Pareto-optimal. The research utilises explicitly datasets from Pleiades-HR 1B and Landsat 5 TM sensors in the experimental phase. The study includes a thorough analysis to demonstrate the suggested method's supreme performance compared to various existing approaches.

In the realm of supervised learning for satellite image CD, the need for a substantial amount of annotated data is a significant challenge, as data acquisition and labeling efforts can be demanding. Additionally, there is a potential drawback to erroneously identifying reconstructed building areas, albeit such alterations are infrequent.

Incorporating an attention mechanism aims to improve focus on the foreground, especially in scenarios where positive ground truth pixels are scarce. Yet, this method faces difficulty accurately discerning between high-level and low-level attributes.

This leads to an unstable attention mechanism when positive and negative samples are evenly distributed. Approaches based on algebra cannot provide comprehensive metrics for change information, while image regression methods require the development of precise regression functions without yielding a comprehensive change matrix.

Image rationing methods introduce a limitation associated with scale changes based on a single date, potentially resulting in varying scores for the same ground-level change depending on the direction of change. Similarly, image differencing methods fall short in furnishing a detailed change matrix, necessitating the manual selection of thresholds for interpretation. The existing CD approaches with their limitations and benefits are tabulated in Table 1.

## 3. Materials and Methods

The process of satellite image CD involves several crucial stages, starting with acquiring a dataset capturing the same geographic area at different points in time. These images act as the input for the Absolute Convolutional Prior Fusion model, an advanced technique built upon the U-Net architecture. Preprocessing is crucial for enhancing the image's quality and setting it up for further analysis. Once the images undergo preprocessing, they are inputted into the Absolute Convolutional Prior Fusion model, leveraging DL to extract intricate spatial and spectral features automatically.

The model aims to amalgamate absolute differences between images, effectively enhancing its ability to discern subtle and substantial changes. Following the model's detection, the dice similarity score is calculated to quantify the similarity between predicted change regions and the ground truth labels. Ultimately, the areas where changes are identified are highlighted, offering a visual depiction of alterations in the landscape. Figure 5 illustrates the detailed block schematics of the suggested method.

**Table 1. Change detection methods with merits and demerits**

| Approaches | Advantages | Disadvantages |
|---|---|---|
| Image Differencing | • Simple Execution. | • This methodology does not give a point-by-point matrix and requires an acceptable range. |
| Image Regression | • Minimize the effects of atmospherical and environmental variations between reference images. | • This approach requires a precision regression function. |
| Change Vector Analysis | • Potential to handle more bands of spectrum. | • More complex. |
| Principal Component Analysis (PCA) | • The repetition of information lowers. | • It cannot give a total matrix to change data and requires an edge to identify the progression that happened in the territory. |
| Tasseled Cap | • Minimize the amount of gap between bands. | • It is inconvenient to interpret and probably will not offer an entire matrix of changes. |
| Post Classification Comparison | • Minimize the impact of atmospheric. | • More production time. |
| Spectral Mixture Model | • Steady and precise output. | • Complex implementation. |
| Visual Interpretation | • During analysis, human expertise and information are useful. | • It cannot give point-by-point data that has been changing, but it consumes more and more time to update the result. |
| GIS Approach | • It allows for mapping changes in the images of present and past data. | • The performance of results varies in the mathematical and classification processes. |
| Integrated GIS and RS Method | • It empowers the elucidation and investigation of information to be accessed. | • Detailed data from various sources changes the identification. |

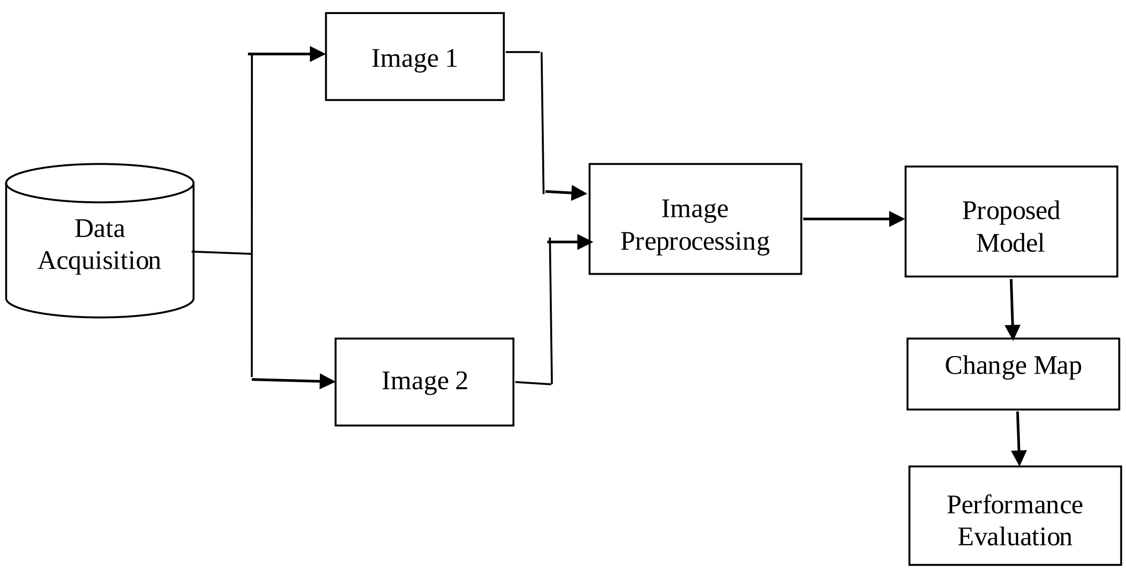



**Fig. 5 Block schematics of the proposed method**

### *3.1. Dataset Description*

The Onera Satellite Change Detection Dataset [23], which tackles the challenge of finding differences between satellite images taken on different dates, is used in this study. A total of 24 pairs of multispectral images gathered from Sentinel-2 satellites between 2015 and 2018 constitute the collection, which was assembled by the Office National d'Etudes et de Recherches Aerospatiales (Onera). The areas covered by the locations that were chosen are varied. Thirteen-band multispectral satellite image pairs with different channel resolutions are included in each location. The resolution information of various channels is tabulated in Table 2.

**Table 2. Resolution information of various channels**

| Band | Resolution (Meters) |
|---|---|
| Blue (B2) | 10 |
| Green (B3) | 10 |
| Red (B4) | 10 |
| Near- Infrared (B8) | 10 |
| Red Edge (B5) | 20 |
| Near Infrared NIR (B6, B7, B8A) | 20 |
| Short- Wave Infrared SWIR (B11, B12) | 20 |
| Coastal Aerosol (B1) | 60 |
| Cirrus Band (B10) | 60 |

**Fig. 6 Images of Beirut in 2015 and 2018, along with its Pixel-Level ground truth**

Ground truth data in pixel-level change masks, specifically focusing on urban changes like new buildings or roads, is provided for all 24 image pairs. In computer vision, 24 image pairs are regarded as a sparse dataset. The visual representation of sample images from this dataset is illustrated in Figure 6.

### *3.2. Image Preprocessing*

The distinctive design of the Absolute Convolutional Prior Fusion (AC-PF) model is tailored for processing pairs of satellite images (pre-change and post-change). An initial preprocessing step involves segmenting each input image into 96x96 chips to enhance training and diversify the dataset, generating a series of smaller image patches. This segmentation allows the model to analyze localized details. Following this, an augmentation step is implemented, which includes rotating each chip by 90, 180, and 270 degrees. This quadruples the dataset by introducing variations in perspectives, lighting, and spatial orientations, enriching the training samples. Dataset augmentation is crucial for improving the model's generalization capabilities across different scenarios, ultimately enhancing overall accuracy. Stacking operations are applied both before and after the chipping process to ensure the model receives a composite image as input and generates a single output image. This meticulous approach streamlines the input-output architecture, aligning with the goal of efficiently and accurately detecting changes in satellite imagery. The combined use of chipping, augmentation, and stacking techniques collectively contributes to the AC-PF model's robustness and effectiveness in satellite image CD tasks.

### *3.3. Convolutional Neural Network*

CNN stands out as a distinctive subset within Artificial Neural Networks (ANN), distinguished by its capacity to leverage spatial information inherent in the data. CNNs are integral to the field of DL, where "deep" denotes the network's depth, signifying the number of hidden layers [24]. Drawing inspiration from the concept of image processing filters and a sliding window, CNNs employ filters (kernels) and convolutional operations. In deep neural networks, the initial convolutional layers focus on extracting fundamental features like lines from the input image. Subsequent layers are dedicated to capturing more intricate features, progressing from basic elements to complex objects. This progression enables CNNs to construct increasingly sophisticated representations. Figure 7 illustrates the workflow of the CNN process. Loss functions, also known as error functions, play an integral part in the learning process of deep networks. They are instrumental in updating the network weights by utilizing derivatives computed from the loss function, which evaluates the discrepancy between the predicted output and the actual result. Additionally, loss functions guide optimizers in minimizing the error as closely to zero as possible throughout iterations and epochs.

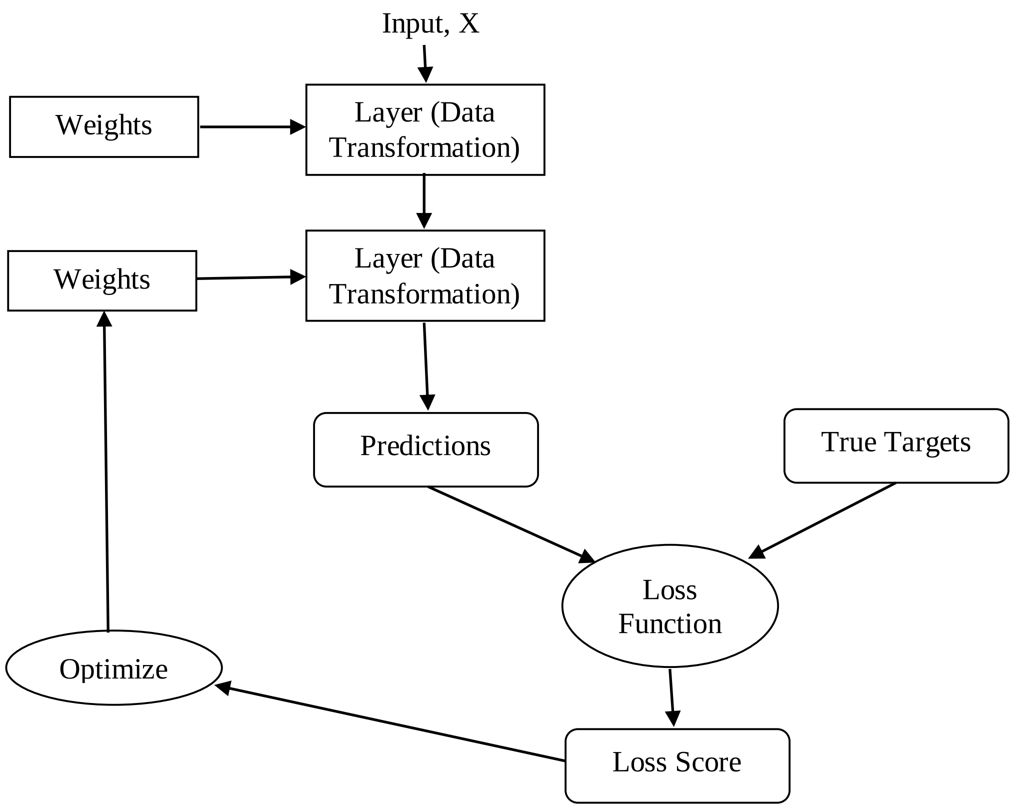


**Fig. 7 CNN workflow**

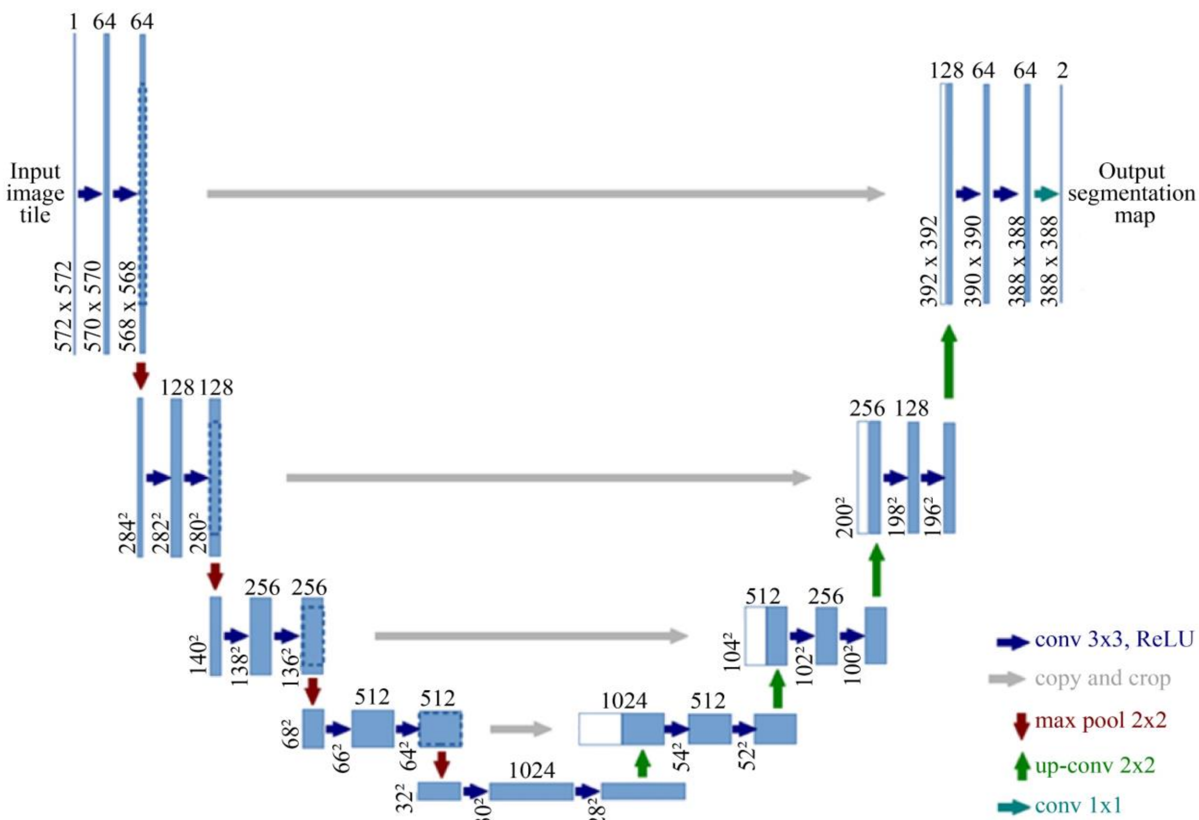


**Fig. 8 Basic U-Net architecture**

### 3.4. U- Net Model

The U-Net is a CNN framework designed explicitly to address semantic image segmentation tasks within computer vision. Its fundamental design draws inspiration from the encoder-decoder architecture and employs two paths: the encoder (contracting path) and the decoder. The encoder is responsible for extracting feature maps from input images through successive down-sampling using CNN layers [25], reducing resolution while enhancing feature extraction. In contrast, the decoder focuses on restoring spatial information lost during the encoding process.

The decoder utilizes convolution and deconvolution operations to up-sample the feature maps generated by the contracting path. Unlike autoencoders, U-Net incorporates concatenating skip connections, which is crucial in reconstructing spatial information and reinstating image resolution. This distinctive feature sets U-Net apart, allowing it to address semantic image segmentation tasks effectively.

The fundamental structure of the U-Net is built upon a regular CNN framework. It employs a series of consecutive (3×3) convolutions, accompanied by ReLU activation and max-pooling layers. This sequence is repeated several times to form distinct levels before reaching the bottleneck, which serves as the connection point between the encoder and the decoder. The receptive field is augmented in the encoder section, which is achieved by increasing the depth (number of channels). Simultaneously, the resolution is diminished due to the inclusion of stride convolutions and pooling layers. Figure 8 illustrates the basic model architecture of the U-Net framework.

The described architecture incorporates both convolutional and deconvolutional layers. Up-sampling layers play a crucial role in restoring resolution from the bottleneck by employing (2×2) up-convolutions. Corresponding levels in the encoder and decoder are aligned, with each decoder level featuring a (2×2) up-convolution layer, followed by a (3×3) convolutional layer, and ReLU activation. Unlike the encoder, the decoder reduces the number of channels while increasing the resolution.

Skip connections, established through concatenation, facilitate the transfer of spatial information between corresponding levels in the two paths. This process aids in reconstructing the spatial structure of the image. Feature map cropping is implemented to address the diminished contextual data at the borders. The operations of the U-Net model are visually depicted in Figure 9.

### 3.5. Proposed Change Detection Model Using Absolute Convolutional Prior Fusion (AC-PF)

The AC-PF model proposed here is a direct derivation from the U-Net model. This adaptation considers the available training data, resulting in the modified architecture depicted in Figure 10. In contrast to the U-Net model, the AC-PF comprises four max pooling and four up sampling layers instead of the original five. Additionally, the layers in the AC-PF are shallower compared to their counterparts in the U-Net model. The input for this network involves concatenating the two images within a pair that are being compared.

The convolutional and pooling layers capture contextual information in the model. The central bottleneck plays a crucial role in preserving this information. On the decoding path, up-sampling and concatenation operations are utilized to reconstruct segmentation maps at an HR, emphasizing the regions that have changed. Throughout the training process, loss functions, such as binary cross-entropy, guide the model to reduce the disparity between the predicted change maps and ground truth, ensuring effective learning.

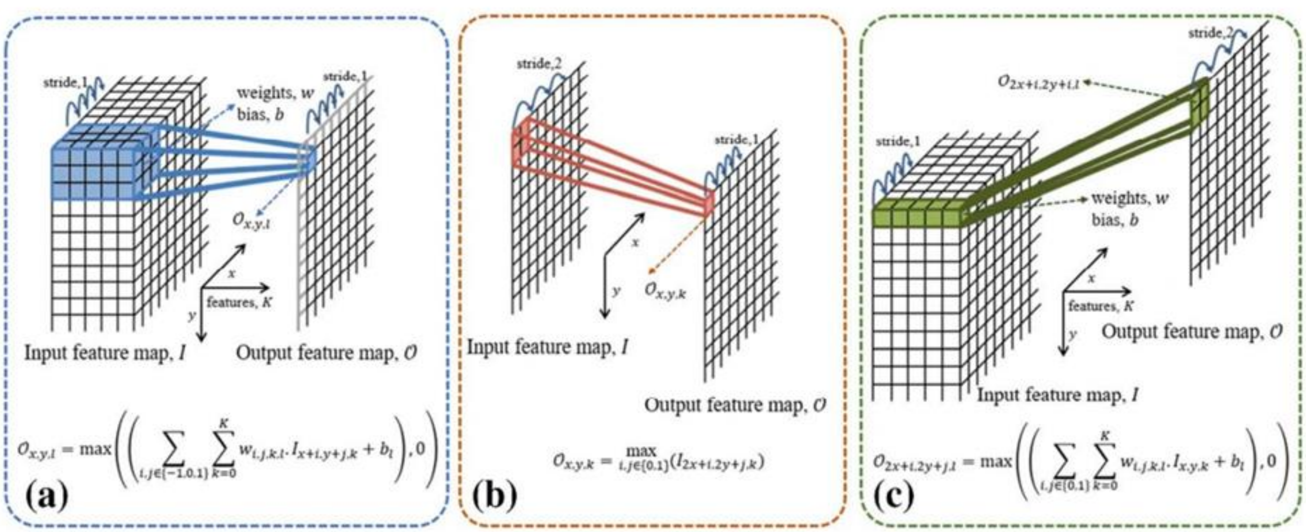



**Fig. 9 Summary of U-Net operations (a) 3×3 convolution + ReLU, (b) 2 × 2 max-pooling, and (c) 2 × 2 up-convolution operation.**

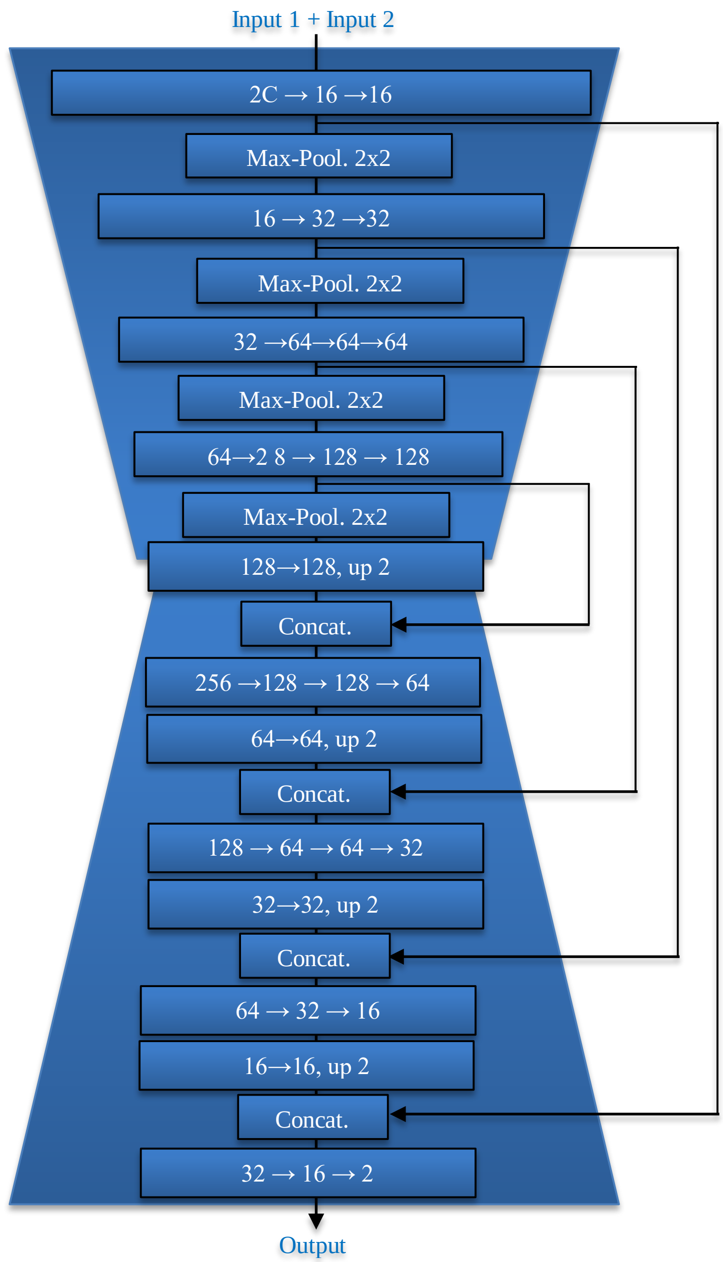



**Fig. 10 Proposed model architecture**

A pooling or down sampling layer is responsible for gradually decreasing the size of feature maps generated by the convolution layer. It is important to note that the pooling layer does not alter the depth of the convolution layer [26]. Max pooling, a specific pooling operation, picks the maximum element within the filter-covered region of the feature map. Consequently, the output from a max-pooling layer forms a feature map, highlighting the most prominent features from the preceding layer. In a singular stream U-Net model, max pooling layers play a pivotal role during the encoding phase by capturing hierarchical features and reducing spatial dimensions. As a down sampling operation, Max pooling entails choosing the maximum value from a group of neighbouring pixels within a specified region. The significance of max pooling layers lies in their ability to distill essential information while discarding redundant details progressively. Through successive reductions in spatial resolution, max pooling aids the model in developing a robust understanding of hierarchical features, encompassing local and global contexts. Moreover, max pooling contributes to the computational effectiveness of the approach by diminishing the number of parameters and computations in subsequent layers. This, in turn, facilitates quicker training and inference without compromising the model's capacity to identify pertinent features.

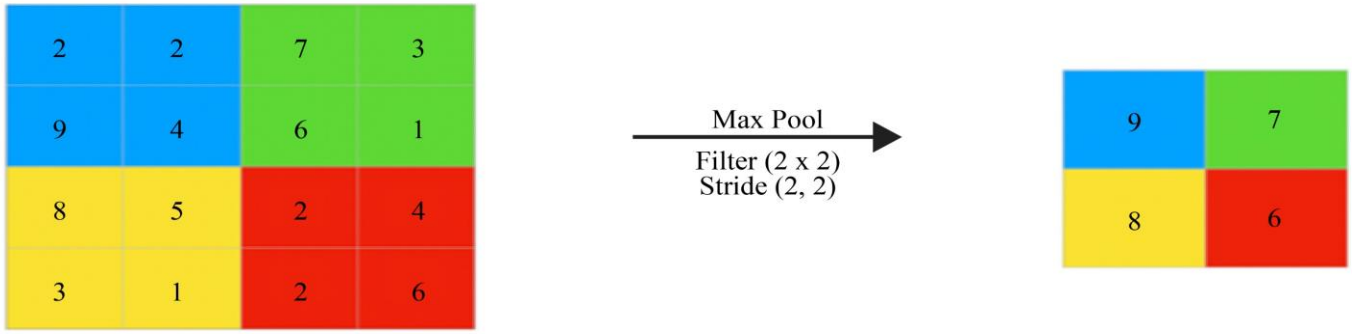



**Fig. 11 Visualization of maxpooling operation**

Concatenation is widely used in neural network architectures, particularly in the U-Net framework during the up-sampling phase. The architecture includes a contracting path for contextual information and an expanding path for precise localization. During down-sampling in the contracting path to extract features, it is crucial to maintain spatial information for accurate localization during up-sampling.

Concatenation is key in merging feature maps from corresponding layers in the contracting and decoding paths. Specifically, this involves taking feature maps from the relevant layer in the contracting path and concatenating them along the channel dimension with the feature maps from the current layer in the expansive path. This operation effectively combines detailed HR information with a semantically rich context.

The suggested model architecture processes an input of dimensions (96, 96, 8), representing an 8-channel image. It incorporates a contracting path with multiple layers of convolutional and max-pooling operations to capture hierarchical features. The bottleneck of the model retains the most abstract information, while the expansive path utilizes up sampling layers to reconstruct spatial resolution. Notably, the concatenation operation is employed in the expansive path to merge feature maps from the contracting path, facilitating the preservation of fine-grained details. To enhance regularization in the contracting path, dropout is applied.

### *3.6. Performance Evaluation*

The Dice similarity coefficient, commonly known as the Dice score or Dice coefficient, is a widely employed metric for assessing performance in image segmentation tasks. It quantifies the similarity between predicted and ground-truth segmentation masks.

This metric is especially valuable in scenarios involving imbalanced datasets, where the abundance of background pixels far exceeds that of foreground pixels. To compute the Dice similarity coefficient, the following formula is employed:

$$Dice\ Similarity\ Score = \frac{2\times(X\cap Y)}{|X|+|Y|} \quad (1)$$

## 4. Results and Discussion

### *4.1. Hardware and Software Setup*

The proposed work employs Google Colab and Microsoft Windows 10 to ensure a reliable computing environment. The system is powered by an Intel Core i7-6850K 3.60 GHz 12-core processor and features an NVIDIA GeForce GTX 1080 Ti GPU. The dataset is split into three parts: 10% is used for testing, 70% is for training, and 20% is for validation. Table 3 presents a tabulated overview of the diverse hyperparameters utilized in the study.

### *4.2. Experimental Results*

A common method for assessing model performance is through an accuracy plot, a visual representation that depicts the correctness of a model across various epochs or iterations during training. The horizontal axis depicts the number of training iterations or epochs in this graphical representation, while the vertical axis displays the corresponding accuracy metrics.

The trend in the accuracy plot is crucial, with an upward trajectory signalling the model's progressive improvement and convergence towards more accurate predictions. The accuracy plot for the suggested model is depicted in Figure 12.

A loss plot visually depicts the progress of a learning model throughout training and validation by illustrating the changes in the loss function over epochs or iterations. The loss function gauges the disparity between the predictions made by the model and the actual values to minimize this distinction throughout the training process.

This graphical representation is essential for assessing the model's learning from the training data and its generalisation ability to new, unseen data. The loss plot for the presented model can be observed in Figure 13.

**Table 3. Hyperparameters**

| | |
|---|---|
| **Activation Function** | ReLu, Sigmoid |
| **Optimizer** | Adam |
| **Loss** | Binary Crossentropy |
| **Batch Size** | 2 |
| **Number of Epochs** | 50 |

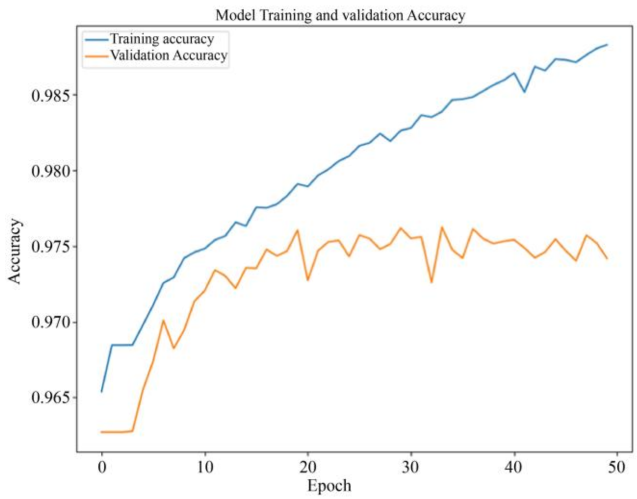



**Fig. 12 Accuracy plot of the proposed model**

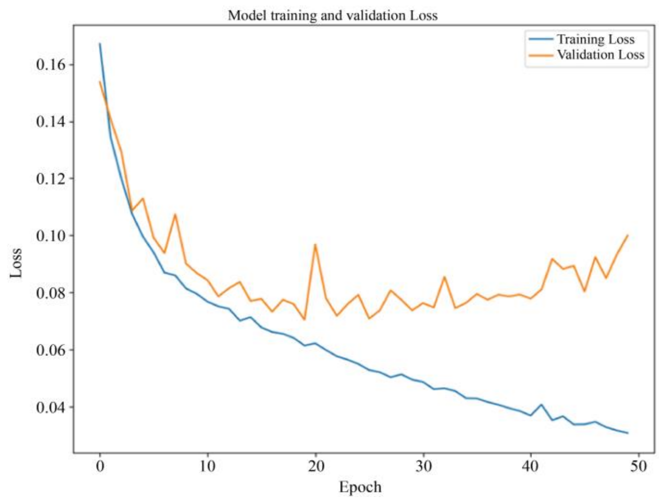
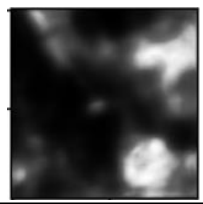
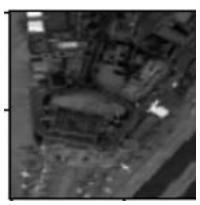
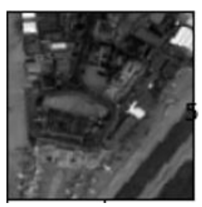
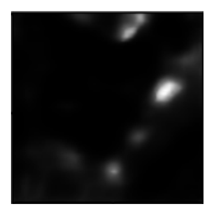
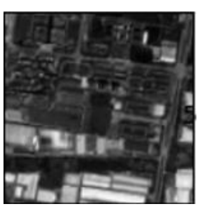
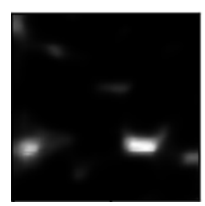
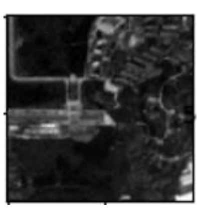
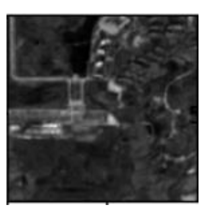
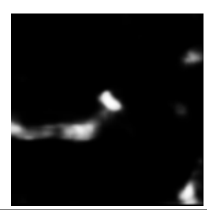
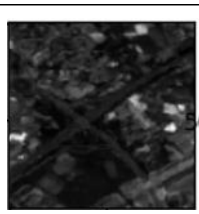
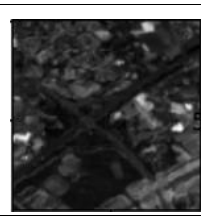
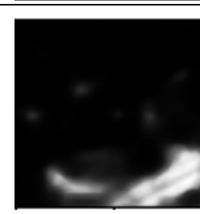



**Fig. 13 Loss plot of the proposed model**

**Table 4. Dice similarity score calculation**

| Initial Image | Changed Image | Ground Truth | Prediction | Dice Similarity Score |
|---|---|---|---|---|
|  |  |  |  | 0.6978368327274562 |
|  |  |  |  | 0.59949129493561 |
|  |  |  |  | 0.4754219845073826 |
|  |  |  |  | 0.6678255080203945 |
|  |  |  |  | 0.9037366465587999 |

The Dice similarity score is calculated for a specific image represented by the variable images compared to the ground truth images, and the corresponding input images are tabulated in Table 4.

## 5. Conclusion

CD using satellite imagery is crucial for monitoring urban developments, providing valuable insights across various sectors and decision-making processes. Systematically analyzing changes in urban landscapes over time yields critical information on urbanization, infrastructure growth, and environmental shifts. This technology aids in identifying trends, patterns, and anomalies, informing urban planning, resource management, and disaster response. The significance of satellite image CD lies in its ability to offer an objective view of urban transformations, surpassing the limitations of traditional ground-based monitoring. It enables the timely detection of alterations like land use changes, urban expansion, and environmental degradation, enhancing our understanding of the urban environment. This paper introduces a robust CD model using a DL approach based on the U-Net model. Adapting the U-Net architecture into the Absolute Convolutional- Prior Fusion (AC-PF) framework considers available training data, improving the model's effectiveness in detecting changes. The proposed model's efficacy is evaluated through the Dice similarity score, calculated by comparing 'images' with ground truth and corresponding input images in the final evaluation.

## Acknowledgments

I sincerely want to thank everyone who helped me finish my research report. I want to express my sincere gratitude to my supervisor, my family, my coworkers, and other researchers for their support and tolerance throughout the challenging parts of my work.

## References

[1] R.C. Luo, and M.G. Kay, "Multisensor Integration and Fusion in Intelligent Systems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, no. 5, pp. 901-931, 1989. [CrossRef] [Google Scholar] [Publisher Link]

[2] Y. Tzeng, S. Chiu, and K. Chen, "Automatic Change Detections from SAR Images Using Fractal Dimension," *IEEE International Symposium on Geoscience and Remote Sensing*, Denver, CO, USA, pp.759-762, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[3] Lisa Gottesfeld Brown, "A Survey of Image Registration Techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325-376, 1992. [CrossRef] [Google Scholar] [Publisher Link]

[4] B. Qiu et al., "Multi-Block PCA Method for Image Change Detection," *12th International Conference on Image Analysis and Processing*, Mantova, Italy, pp. 385-390, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[5] R.J. Radke et al., "Image Change Detection Algorithms: A Systematic Survey," *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp.294-307, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[6] Ashbindu Singh, "Digital Change Detection Techniques Using Remotely-Sensed Data," *International Journal of Remote Sensing*, vol. 10, no. 6, pp. 989-1003, 1989. [CrossRef] [Google Scholar] [Publisher Link]

[7] Jordi Inglada, and Grgoire Mercier, "A New Statistical Similarity Measure for Change Detection in Multitemporal SAR Images and Its Extension to Multiscale Change Analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 5, pp. 1432-1445, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[8] Ebrahim Ghaderpour, and Tijana Vujadinovic, "Change Detection within Remotely Sensed Satellite Image Time Series via Spectral Analysis," *Remote Sensing*, vol. 12, no. 23, pp. 1-27, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] Jie Chen et al., "DASNet: Dual Attentive Fully Convolutional Siamese Networks for Change Detection in High-Resolution Satellite Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 1194-1206, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Xiuwei Zhang et al., "DifUnet++: A Satellite Images Change Detection Network Based on Unet++ and Differential Pyramid," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Caijun Ren et al., "Unsupervised Change Detection in Satellite Images with Generative Adversarial Network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 12, pp. 10047-10061, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[12] Ekaterina Kalinicheva et al., "Unsupervised Change Detection Analysis in Satellite Image Time Series Using Deep Learning Combined With Graph-Based Approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 1450-1466, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] Xuan Hou et al., "High-Resolution Triplet Network with Dynamic Multiscale Feature for Change Detection on Satellite Images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 177, pp. 103-115, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Hichem Sahbi, Sebastien Deschamps, and Andrei Stoian, "Active Learning for Interactive Satellite Image Change Detection," *Arxiv*, pp. 1-9, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Xin Huang, Yinxia Cao, and Jiayi Li, "An Automatic Change Detection Method for Monitoring Newly Constructed Building Areas Using Time-Series Multi-View High-Resolution Optical Satellite Images," *Remote Sensing of Environment*, vol. 244, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[16] Wenqiang Xi et al., "A Spatiotemporal Cube Model for Analyzing Satellite Image Time Series: Application to Land-Cover Mapping and Change Detection," *Remote Sensing of Environment*, vol. 231, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[17] Kaiyu Li, Zhe Li, and Sheng Fang, "Siamese NestedUNet Networks for Change Detection of High Resolution Satellite Image," *Proceedings of the 2020 1st International Conference on Control, Robotics and Intelligent System*, pp. 42-48, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] Ran Jing, Zhaoning Gong, and Hongliang Guan, "Land Cover Change Detection with VHR Satellite Imagery Based on Multi-Scale SLIC-CNN and SCAE Features," *IEEE Access*, vol. 8, pp. 228070-228087, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] Yi Zhang et al., Coarse-to-Fine Satellite Images Change Detection Framework via Boundary-Aware Attentive Network," *Sensors*, vol. 20, no. 23, pp. 1-21, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] Ahram Song, Yongil Kim, and Youkyung Han, "Uncertainty Analysis for Object-Based Change Detection in Very High-Resolution Satellite Images Using Deep Learning Network," *Remote Sensing*, vol. 12, no. 15, pp. 1-26, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[21] Lukas Kondmann et al., "Spatial Context Awareness for Unsupervised Change Detection in Optical Satellite Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-15, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[22] Ramen Pal et al., "Very High-Resolution Satellite Image Segmentation Using Variable-Length Multi-Objective Genetic Clustering for Multi-Class Change Detection," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 10, pp. 9964-9976, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] Soumik Rakshit, Onera Satellite Change Detection Dataset, Kaggle, 2019. [Online]. Available: https://www.kaggle.com/datasets/soumikrakshit/onera-satellite-change-detection-dataset

[24] Jianxin Wu, "*Introduction to Convolutional Neural Networks*," National Key Lab for Novel Software Technology, Nanjing University. China, pp. 1-31, 2017. [Google Scholar] [Publisher Link]

[25] Xiao-Xia Yin et al., "U-Net-Based Medical Image Segmentation," *Journal of Healthcare Engineering*, vol. 2022, no. 1, pp.1-17, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[26] Hossein Gholamalinezhad, and Hossein Khosravi, "Pooling Methods in Deep Neural Networks, a Review," *Arxiv*, pp. 1-16, 2020. [CrossRef] [Google Scholar] [Publisher Link]