iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** V **Month of publication:** May 2020

DOI: http://doi.org/10.22214/ijraset.2020.5086

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

Survey on Change Detection Techniques for Multitemporal Aerial Video Sequences

A Meghana[1], Bharathi Gummanur[2], Likhitha Priya D[3], Arya G V[4], Krishnan Rangarajan[5]
[1, 2, 3, 4, 5]*Department of Computer Science, Dayananda Sagar College of Engineering, Bangalore-78*

Abstract: ***In this paper, we perform a survey of various techniques that can be used to perform change detection on a pair of images taken at different times. Each of these techniques perform analysis on multitemporal images and identify modifications, if any. The advantages and disadvantages of each technique is identified and scrutinized to evaluate the performance of each technique. A comparative analysis of the techniques is performed to determine the most suitable technique to be employed for different scenarios, like video surveillance, infrastructure monitoring and so on.***
Keywords: Deep Convolutional Neural Network, Optical Image, Conditional Random Field, Image Preprocessing, Object Identification, k-NN, Markov Model, Image stitching, Classification.

I. INTRODUCTION

Aerial images are captured through various platforms like Unmanned Aerial Vehicles. In the field of remote sensing and surveillance, analysis of these aerial images for the automation of the process of change detection is very necessary as the manual process of performing the same task is very tedious and time-consuming. The main focus of this paper is the discussion of different techniques used for detection and analysis of changes between two aerial video feeds taken with significant differences in time, seasonal and light conditions. In order to obtain the change between two video feeds, we begin by converting the videos into panorama images using image stitching techniques which have been described in this paper. This can be performed using two techniques - Identifying and detecting frames with max overlap and comparing them or by creating mosaics and comparing the mosaics. This is followed by comparison of the obtained panoramic image pairs to detect changes, if any. The popular techniques like Conditional multilayer mixed Markov model, intensity-based correlation methods, feature based comparison methods, Machine Learning based methods. that are used to perform change detection have been discussed in this paper. Conditional multilayer mixed Markov model is a probabilistic model has been employed to enable detection of appropriate changes in the pairs of aerial images. Global intensity along with correlation, performed locally and features, that are contrasting are integrated by the model. For optical photographs, differences in pixel levels is not sufficient to segregate the background from the modified region, according to methods like MLP and Parzen. The PCA (Principle Component Analysis) approach uses a scalar feature for the same. The k-Nearest Neighbor approach enables updation of the changed or unchanged pixels of a binary change classifier, like the FFNN pixel-based binary classifier. The DrLIM technique for learning a non-linear function which is coherent globally performs mapping of the data to the output. Siamese CNN has also been implemented due to its relevant advantages as produces feature vectors for each subsequent pixel of a pair of images rather than just one at a time. This is clearly more efficient than the various methods that must input one image at a time. Siamese CNN also extracts features directly from image pairs rather than using hand crafted feature vectors. The efficiency can be increased by using knn updating methods. In the section II we discuss the image stitching techniques used to convert a video into an image. The subsequent sections - Section III to Section X, discuss various techniques, like PCANetwork, Markov models, Deep learning, MFRC models and so on respectively, that can be employed to perform change detection on the image pairs. This is followed by discussion of techniques that can be employed to perform detection of the objects that have been identified as a source of the change in section XI.

II. IMAGE STITCHING

Combining multiple images with overlapping regions to obtain a panorama is called as image stitching. To realize 360°panoramic stitching, in [16], in a cylindrical method for stitching videos to obtain panoramic images which depends on template frame is proposed. First step is the collection of the template frame image. These template images are matched approximately and SURF points extraction are performed and then an efficient feature point filtering method is used for elimination of mismatching feature points. The feature points that match perfectly are projected onto the cylinder. Between two images, the transformation matrix is calculated earlier in order to fuse the video frames. Now, image fusion is performed. To meet the real time requirements, the algorithm is sped up by cylindrically projecting and fusing the images by parallel processing through CUDA. The video stitching was performed at 30 frames/second in real time.

©IJRASET: All Rights are Reserved

In image registration, the image is converted to the coordinate system of the template image and the coinciding parts of the images are aligned to fuse the images. This is the main aspect of image stitching. This includes three aspects: extraction of features, matching of features and determining the transformation matrix. To improve the speed and accuracy of the matching process, only the feature points that are similar are compared. The Euclidean distance serves as the similarity measure.

One of the images will serve as a template, and the feature points of this will be examined to obtain the closest feature points and the subsequent closest feature points in the other image. Then a ratio of similarity measure for the points are calculated. If ratio is within a specific threshold, then the feature points are correctly matched. SURF algorithm enables extraction of the feature points from the entire image. Feature points are extracted only from the overlapping parts of the images.

In order to calculate the modification between two image transformation of cylindrical projection is essential to fuse the images. In order to determine the shift between images, the source image is moved towards the target image, using set of matching feature points. After calculating the transformation, they can be fused. To enable fusion, cylindrical projection of the image is necessary. Through transformation matrix movement, perfect fusion between source and target images can be obtained by taking one of the images as the target image and the other image as the base image.

III. PCANETWORK

This method classifies images based on deep learning. This method uses the representative nearby features from every pixel using PCANetwork filters. These features are considered as convolutional filters. It reduces great amount of the speckle noise and will be able to produce change maps that have less noise.

Fuzzy c-means and Gabor wavelets (GW) are used and relevant pixels from the multi temporal images are selected which have higher chances of being changed or unchanged. The newly produced image patches that are centered at the relevant pixels are used and PCANetwork model is trained. Finally, the change map is generated by combining the PCANetwork classification result and pre-classification result.

This research work is described in [8], was mainly focused on thresholding clustering, active contours and Gabor features to identify the pixels from DI.

Gabor wavelet representations are used to work with the changed information. After implementing the fuzzy c-means method with the nearest neighbor (NN) approach, a better performance is given by the model. In every stage of the recent deep approaches, they comprise convolutional-filter-bank, feature-pooling layer and non-linear processing layer. Restricted-Boltzmann-machines (RBMs) are quite involved to learn about the filter-bank.

PCANetwork, proposed by Chan, is a deep-learning network where the filter-banks are taken from PCA filters that have showed good performance. In the PCANetwork change detection method for SAR images have contributions as follows:

A. PCANetwork is taken as the classification model and the PCA filters are used as filters and the traditional architecture of the CNN is retained.

B. A pre-classification scheme, inspired by Li's work, made some accurate labelled samples for PCANetwork using GW and FCM.

The proposed method is mainly composed of three steps:

1) *Step:* Log-ratio images are obtained. Gabor wavelets and FCM algorithm are used to select pixels that have high probability of being changed or unchanged.

2) *Step2:* The generated image patches that are centered at interested pixels are used to train PCANetwork model.

3) *Step3:* The trained PCANetwork model is used to classify the remaining pixels into either changed or unchanged. The remaining pixels are further separated into changed and unchanged classes using the trained PCANet model. The final change map is formed by using the results of PCANetwork classification and results of the pre-classification.

a) *Experimental Setup*: The proposed method is applied for the 3 data-sets which possess different features-Ottawa, San Francisco and Yellow River datasets.

b) *Result and Analysis:* On first dataset the D_MRFFCM produced 98.3% of PCC value, while the method proposed here gave 98.2% value. On second dataset the proposed method has given a best performance with 98.94% value. The results of Yellow River dataset had the influence of speckle noise which stopped from giving good performance. It showed a difficult situation to determine the efficiency of the proposed method.

©IJRASET: All Rights are Reserved

IV. MARKOV MODEL

Systems that change randomly can be modelled using a stochastic model known as Markov model.

A. *Multilayer Conditional Mixed Markov Model*

A probabilistic model has been implemented in [5] to detect appropriate changes in registered aerial image pairs taken in different seasonal conditions and with time difference of several years. A conditional mixed Markov model has been employed in this method which is a combination of a field of signals that are independent conditionally and mixed Markov model. Global intensity along with correlation, performed locally and features, that are contrasting are integrated by the model. Global energy optimization ensures smooth observation-consistent segmentation and simultaneous optimization of local feature selection.

Validation is performed on real aerial images. This signifies the advantages of the CMX model in comparison to the earlier techniques for the change detection problem. Improvements have been made by considering several features and integrating them using a probabilistic schema. The joint usage of multiple features and the proposed probabilistic feature integration schema contribute to the improvements made. For optical photographs, differences in pixel levels is not sufficient to segregate the background from the modified region. Meanwhile, a scalar approach has been employed by the PCA technique.

Processing of images with resolution of 1.5m/pix from regions of scanty population provided efficient results. Factors like misregistration and parallax at low resolution reduce the efficiency. This can be modified by choosing larger areas. However, a higher level of processing is necessary to deal with moving objects and densely populated areas. However, reduction in vegetation may modify the color or texture of the area resulting in false alarms. (Fig 1)

B. *Multilayer Markovian Model*

A Three-Layered Markov model for identification of change in multitemporal images is proposed in [6] as follows:

Two sets of features are considered: the Grey-Level Difference and the Modified HOG difference. The third layer is formed by combining the first two layers. The final result is generated taking both texture level and pixel level features into consideration. Pairwise interaction which enables retention of the sub-modularity condition for energy has been employed. Global energy is optimized by using max/min-cut flow algorithm.

The image pairs I1 and I2 of the same region but with different atmospheric conditions and a large time difference were converted to greyscale images G1 and G2 and then L2 normalization performed over the image pixels provides the normalized image.

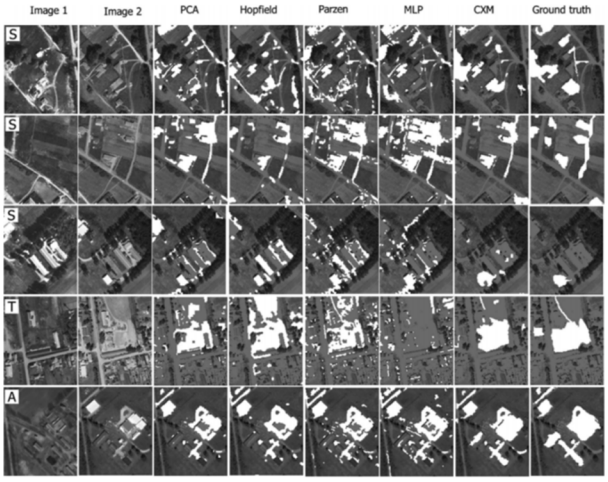


Fig 1. Qualitative comparison of the change detection results with the different test methods [5]

On the same dataset, this method has provided better results than the CMX method and other futuristic methods. The CMX method which also uses a supervised method is compared with this method qualitatively and quantitatively. The segmentation results of the CXM model were validated against Ground Truth. The number of missed changes (during segmentation, ignorance of changed pixels), falsely detected changes (detection of unchanged pixels as changed pixels) and total error (sum of number of instances of missed and falsely detected changes) were computed. A more homogeneous and smoothening effect is produced in this method than in CMX method. This method requires 10 to 15 seconds per image pairs. So, in comparison to CXM method, this method is more efficient computationally.

©IJRASET: All Rights are Reserved

V. DEEP LEARNING TECHNIQUE

Deep Learning is a part of machine learning family that is based on artificial neural networks. The algorithms based on this technique tries to resemble the structure and function of the human brain. In [3] the process of change detection of images is to analyze two pictures with the same scene which are taken at different times and determine the changes between them. Unsupervised detection in synthetic aperture radar images can be split into three steps:

A. Pre-processing, which includes co-registration, geometric correction and denoising
B. Generating a DI (difference image) among the multi temporal images. This is done by comparing the two in two co-registered images in pixels. This paper proposed a neighborhood-based ratio (NR) approach to generate DI.
C. Analyzing the DI. A decision threshold is applied to the DI's histogram in the thresholding method. Classical threshold methods like Otsu and the Killer and Illingworth are applied to actuate the threshold an un-supervised way.

Another method for segmenting images is the clustering method. The FCM, fuzzy c-means, algorithm can preserve the information better than the hard clustering.

The proposed model is compared with two other models which are,

1) Gaussian Model with KI-based method (GKI), which uses generalized Kittler and Illingworth minimum error thresholding algorithm on the test statistic images to determine the changes.
2) RFLICM which is the modified and reformulated function of the FLCIM (fuzzy local information c-means clustering algorithm).
a) *Proposed Model:* A deep architecture for change detection. In this model first pre-classification is performed for obtaining some examples. Second to pre-train a deep neural network RBMs are used and then fine-tune it. Third, to classify a DI a trained deep neural network is used. The neighborhood features of each pixel in the DI is taken as inputs and logistic output can be used to indicate the change in the pixel.
b) *Experimental Study:* Three datasets were adopted to test the algorithm proposed in this paper which were Ottawa dataset, Bern dataset, and Yellow River dataset
c) *Results:* For Ottawa dataset, in the GKI method there was noise due to the necessity to search an optimal threshold, and GKI and also in RFLICM had a better final map but there were missed detections. the proposed DNN algorithm gave good performance and has no requirements for the model or the data distribution. It has strong ability to learn complicated functions. Although DNN doesn't result on the lowest FP and FN, it can balance them better. For Bern Dataset There were noise on the final map generated by GKI and in the RFLICM model the sensitivity of the noise was decreased but many changed areas where detected as unchanged. The proposed DNN performed well when checked the final map and the table. For the Yellow River Dataset, here the changed areas are relatively small GKI did not gave good performance and the final map obtained by RFLICM gave false alarms because of existence of noise, but the proposed method DNN gave the best to complete the detection task which yielded 76.51% which is much higher than 59.98% by RFLICM and 30.26% by GKI
d) *Conclusion:* the deep learning network has an ability to learn features and represent them in an abstract way. It is not necessary to model the data and the results are robust to noise. Compared to the clustering method RFLICM and thresholding the method GKI, the proposed method exhibits good performance.

VI. A MCRF MODEL FOR OPTICAL IMAGES

A MRCF is used to label images at the pixel level with object classes. In [1] the paper proposed a method called Conditional Random Fields Ensemble Model that is subjected to detect the changes that take place in urban Areas. Urban Areas are rich in texture because they are composed of heterogeneous objects like buildings, man-made objects etc. In order to classify images multiple texture feature is applied. They are commonly grouped as structure based and gray level based. Any images that are fed as input must be preprocessed and for this reason object detection, image segmentation and image classification is done using CRF. Since contextual information is required Multiple Conditional Random Fields (MCRF) come in to picture.

MCRF ensemble model has three parts in it, they are feature configuration, classifier selection and classifier fusion. The probabilistic classifiers are CRFs. With different type of input features along with base CRF the model creates a pack of CRFs. Hence, feature configuration can be simply said as feature grouping that divides an entire feature into distinct features in which each feature is subjected to nonlinear transformation. This grouping is done to attain various sets of features, and each of them are independent. So, during learning process the categorical information can be easily obtained by classifiers.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

For training purpose, the parameter is estimated by standard maximum-likelihood approach, where the labelled images are given and it must choose the parameters such that logarithm of likelihood is maximized. Since the labelled images are provided it is a supervised learning technique. Feature extraction is one of the main steps in the whole process. It is done using existing feature operators and it is carried out by first dividing the input images into many non-overlapping 16x16 pixel blocks. To obtain the general stats of the blocks, five different kinds of multiscale texture features are applied. Instead of considering an entire image, data present in multiple local windows is used to calculate an attainable feature vector of each block.

Aerial Images and Satellite Images were used as dataset for testing images. Labelling every block as non-urban and urban was done manually to obtain ground truth. The performance of algorithm reduced with poor image quality but adequate results were obtained with normal quality images. To sum up, CRF model results one feature which is insufficient to distinguish between non-urban and urban areas. But MCRF's ensemble model integrates very favorable result.

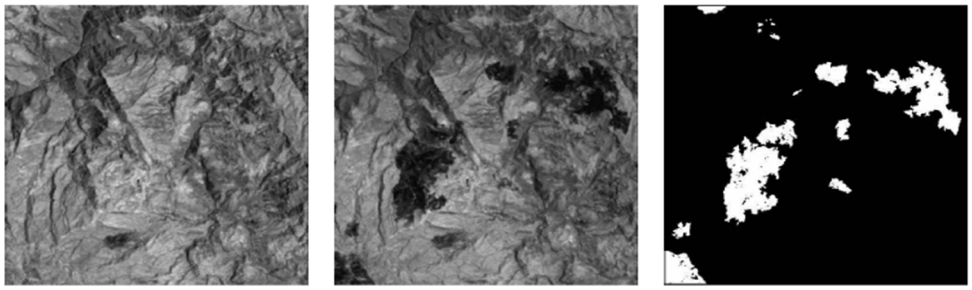
VII. A DEEP CONVOLUTIONAL COUPLING NETWORK

This is an unsupervised network that is used for change detection applications. In [7] a method is proposed on image detection based on heterogeneous images. Images captured from various sensors with different characteristics are called heterogeneous images, for example, radar sensors and optical sensors. Here it is concerned with optical images and SAR. Image preprocessing like co-registration and denoising are applied. To reduce the falsification for precisely detecting changes denoising is used. The multitemporal images contain objects at different scales, viewpoints, sizes and location, hence co-registration comes to picture. To compare two images, we need difference map.

Problems relating to heterogeneous images can be addressed by using classification-based method, in which Post Classification Comparison (PCC) is prominent one. The resultant classification maps are then compared to identify changes. Deep Neural Network is used for transformation of heterogeneous images into consistent feature space and along with this feature of raw images can also be extracted. The main focus is change detection using Symmetric Convolutional Coupling Network (SCNN). Two images of same size are given as input. Firstly, it converts two images into a feature space in which both have feature representation and then identifying difference based on such feature representation. It takes two image's pixel-wise representation, if there are unchanged pixels then the feature representation is kept unchanged. If there exists a change in pixel then it will demonstrate disparate feature representation which makes it easy to detect. Unsupervised pre-training is used which eases the backpropagation algorithm and also enables a useful representation of input images in order to form a compatible feature space. The noise present in the images are treated with Denoising Auto Encoder (DAE) that will reduce the noise layer by layer.

First datasets were one SAR image and one optical image, where SAR images was collected from Radarsat-2 and optical images was collected from Google Earth. Second dataset is a mixture of RGB optical images and SAR images. This also proposes unsupervised image segmentation algorithm for classifying two images. The change detection is represented in the form of binary map, where changed regions are denoted by pixels filled with white color and unchanged regions are denoted by pixels filled with black color. This approach of SCNN is completely unsupervised where it doesn't use any labelled pixels. Different values of λ (user defined weighting parameter) shows inconsistent performance commonly when the value λ is extremely low or high. For homogeneous images, good performance is obtained with small value of λ. But the value must be high for heterogeneous images.

A. Results



(a) Initial input image (b) Input image with change (c) Change detected output [7]

©IJRASET: All Rights are Reserved

VIII. UNSUPERVISED LEARNING

In [2] the paper proposes an unsupervised approach for detecting change in satellite images, where the change detection data is produced using PCA (Principle Component Analysis). Clustering plays an important role under unsupervised learning and k-means clustering algorithm is the most suitable technique described in this paper. This technique classifies the given data into k-different groups where the likeliness is calculated using Euclidean distance. The proposed model primarily aims at generating a change map that indicates the pixel level changes in the pair of images. The method assumes the change detection as a binary classification problem. The method follows the following steps: First is the production of the difference image, which is represented as Xd = |X2 − X1|. Secondly, the proposed method focuses on the production of h X h non coinciding image blocks derived from the difference image. The third step is giving rise to an eigen vector area utilizing PCA, done on h x h non coinciding image blocks. Subsequently, the next step is to create a feature vector area on the entirety of the difference image by reflecting coinciding h x h blocks surrounding every pixel onto the eigen vector area. The fifth step is the procedure of clustering the feature vector space into two clusters equivalent to wu(unchanged pixels) and wc(changed pixels) using k-means method where number of clusters is 2. The subsequent sixth step involves the production of the change map by designating every pixel of the difference image to 1 among all the clusters which is done corresponding to the min Euclidean distance among the pixel's feature and mean feature vector of the clusters. Considering a particular region, if one observes a change between a pair of images, it's quite assumed that the values of the difference image pixels in the particular area are larger than that of the pixels in the areas where there is no alteration. Utilizing this expectation, the cluster with pixels that possess lesser mean value in the difference image is designated as the unchanged class, and the other cluster is designated as changed class. Using the cluster map feature vectors for altered and non-altered pixels, the change map can be produced. For the pixel values, one indicates an altered pixel and zero indicates a non-altered pixel.

The proposed method utilizes block-based data scrutinizing to obtain the sub pixel level result change map. This technique is expected to have a better performance when there is noise incorporated in images when compared to approaches like EM and MRF based thresholding. All in all, the indicated method displays a good performance with respect to tackling the zero mean Gaussian and speckle noise, which is very preferable for change identification in optical and SAR images.

IX. A K-NEAREST NEIGHBOR APPROACH

In [11] the method proposes a solution based on knn to reiterate the altered or not altered pixel label produced by a binary change classifier, which considering this specific application, is an FFNN pixel dependent classifier that uses their spectral bands. The altered and non-altered labels are originated with respect to this binary classifier. The primary idea of the proposed method is assigning a label to the pixel based on the information obtained from the decision provided by that particular pixel's neighborhood. The required trials for this method are done on actual optical aerial images which are multi temporal in nature.

The indicated methodology comprises of two discrete portions of scrutiny: 1. Normal classifier like ffnn that gives score results, 2. knn post-processor which ranks again. This method has utilized a multiple layer ffnn in which inputs are sent to output layer using a feed forward mechanism. The mentioned ffnn comprehends using the training data, then consequently establishes relations among input and output which represents each change. The output layer comprises of a dedicated neuron distinguished as one indicating alteration and zero indicating non-alteration. The network subject to training is subsequently utilized on 2 particular images to give rise to a change map. The output topology implements a 1-out-of C target coding method for the result(O/P) neurons; considering the dimension of the resultant result vector to be equivalent to the enumeration of classes of C. The network is basically being made to comprehend to produce the probability of all individual classes for a given input. After the training of the network, a novel input vector is provided to it, and then identified with the class consisting of the result unit has the greatest activation. Considering every output score vector, (ov,oz), they utilized the rule:

Decision = {1, if o1 > o2 then change
0, otherwise (no change)}

The secondary stage knn classification is done for all individual pixels derived off of the ffnn output scores (o1,o2) and labels of the particular pixel's neighborhood. The ffnn label is not considered in this process even though the score data is preserved. The knn training ex denote the neighborhood hierarchy that is sketched surrounding the pixel position. The class of the middle one is then determined by the highest prominent class utilizing majority voting rule The method described demonstrates motivating results in this confounding topic, image change detection as the indicated approach gives a total performance detection improvement of about 13% and 4% of f-measure and g-mean values, respectively, considering the ffnn baseline system. However, it is required to emphasize that post-processing alteration depends on the correctness of the ffnn classification scores done before. Hence, subsequent objective of the proposed method is to develop a parallel cooperative and dependent scheme between the two classifiers.

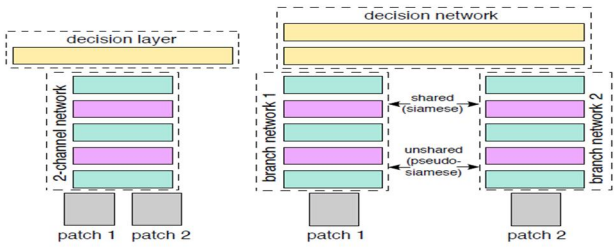
©IJRASET: All Rights are Reserved

X. NON-VARIANT MAPPING TO REDUCE DIMENSION

In [10] the proposed encompasses relating a category of large dimensional i/p points onto a low dimensional manifold so that "similar" points in the former are mapped to those in proximity on the manifold. The proposed approach implements DrLIM for comprehending a universally cogent non-linear function that relates the data uniformly to the output region. The method depends completely on neighborhood relations and has no requirement of any distance measurement in the input space. DrLIM is a solution for the following problems faced by a large number of existing dimensionality reduction techniques: Firstly, they don't create a method from input to manifold that can be deployed and used for novel points for which the relations to the training points is not known. Secondly, many methods require as a precondition, the existence of a meaningful distance metric in the input space. An additional drawback of current methods is that they lean towards clustering points in output space, often with such high density that they are considered obsolete. The method encompasses four important characteristics: First, It only requires neighborhood relationship among training examples. These relations be acquired from previous knowledge, or manual labelling, and not be dependent on any distance form of measurement. Second, it might comprehend functions that are not variant to complex nonlinear transformations of the inputs like variation in light and geometrical contortion. Third, the trained function can be utilized to relate novel samples not witnessed in training, with no previous intellect. Fourth, the mapping produced by the function is in a way "smooth" and cogent in the output space. The method utilizes a contrastive loss function, dissimilar to loss functions that perform over samples, that performs on pairs of samples and is used to push together and pull apart the neighbors and non-neighbors respectively. The proposed technique uses an energy-based model that utilizes the provided neighborhood relations to learn the mapping function. It uses a Siamese architecture (a pair of images as input) which is what the method proposed in our chosen base paper has utilized as well. DrLIM offers a solution to the possibility of distortion and domination of the outcome of dimensionality reduction by the variability of parameters such as lighting, registration, etc., by learning an invariant mapping to a less dimensional manifold using previous intellect. The system prevents collapsing and maintains equilibrium in output space because it uses using a separate loss function for alike and not alike pairs. It is better in comparison to LLE method as LLE may give degenerate results when the input examples are not locally too alike and finely registered.

XI. CNN

CNN is mainly used to analyze data and have a wide range of applications in image processing. Convolution is a mathematical operation which is basically a linear operation. Matrix multiplication is replaced by this linear operation in one of the layers and hence called convolutional neural networks. In [9] the main focus is on comparing two images patches. This helps in solving many Computer Vision problems and in image analysis. This can vary from low-altitude tasks such as motions, panorama constructions to high-altitude tasks such as recognition of objects, image retrieval of images, etc. It builds patch similarity function without using any features that are manually generated. Such functions are represented using Deep CNN. This paper describes three basic models - i) Pseudo Siamese, ii) Siamese and iii) 2-Channel. In Siamese, one patch is fed as input to each branch out of two branches and then a series of convolutional, ReLu and max-pooling layers are applied. The concatenated outputs of each branch are given to a top network. These branches of the Siamese network can be considered as descriptor computation modules and similarity function is represented as top network. 2-channel directly takes the patches for comparison rather than having a descriptor. It is more flexible and faster to train. For evaluating the model, they use evaluation protocol and prompts ROC curves by thresholding the distance between feature pairs. According to the experiments, 2-channel model showed better performance than the rest of them. It adequately outperformed the state-of-the-art. When compared to present state-of-the-art systems, Siamese based architecture also showed a better performance. To be precise, the two-stream network had best performance, which showed the significance of multi-resolution information. The Siamese architecture is shown in fig 2a and fig 2b



2-channel Siamese and Pseudo Siamese

Fig 2a: Network Architectures [9]

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

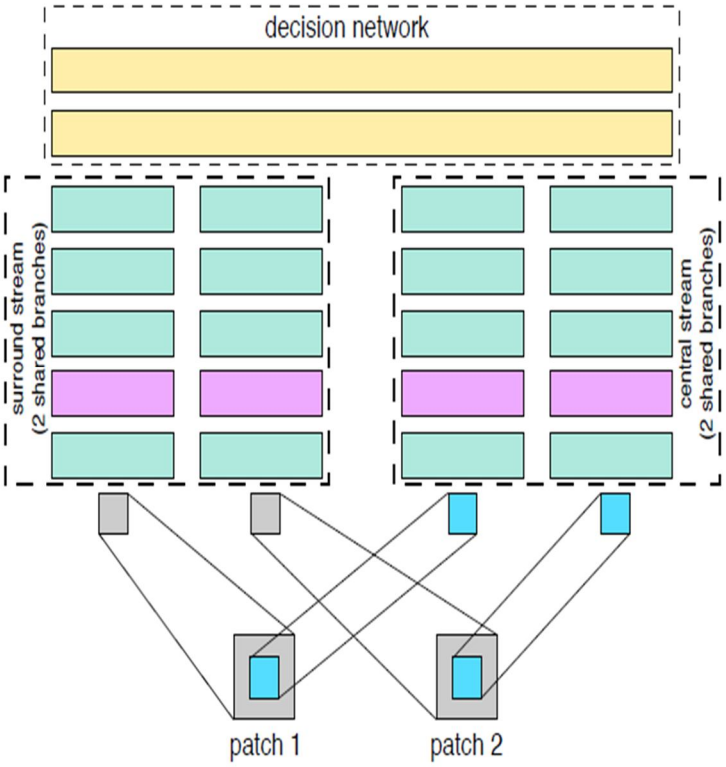


Fig 2b: Four branches resulted from processing of each stream that are fed as input to top decision layer [9]

XII. REPRESENTING RESULTS

Objects causing the changes are detected using object identification technique and CNN is implemented for this reason. CNN is considered because it helps ease in analyzing visual images.

In [4] the cifar-10 dataset is used to perform detection and classification of objects. It uses convolutional neural network (CNN) that works on TensorFlow supported keras. This CNN determines the time required to develop, train and test the model.

Deep neural networks (DNN) can be used to perform image recognition process, in depth. However, it requires a large amount of data. So, a preferred method for image recognition is the use of CNN. It is a type of neural network (NN) which is similar to the regular NN, except, it has an initial convolution layer. It facilitates parallelization of operations which enables object detection. It performs the following tasks:

A. Cropping portions of the image
B. Flipping the image over a level plane.
C. The alteration of saturation, the difference and the shading.

The model is trained for 25 epochs each using 60,000 images. This process takes about 722 to 760 seconds in TensorFlow CPU system. A training epoch of 25 yielded an accuracy of 96 percentage.

XIII. CONCLUSION

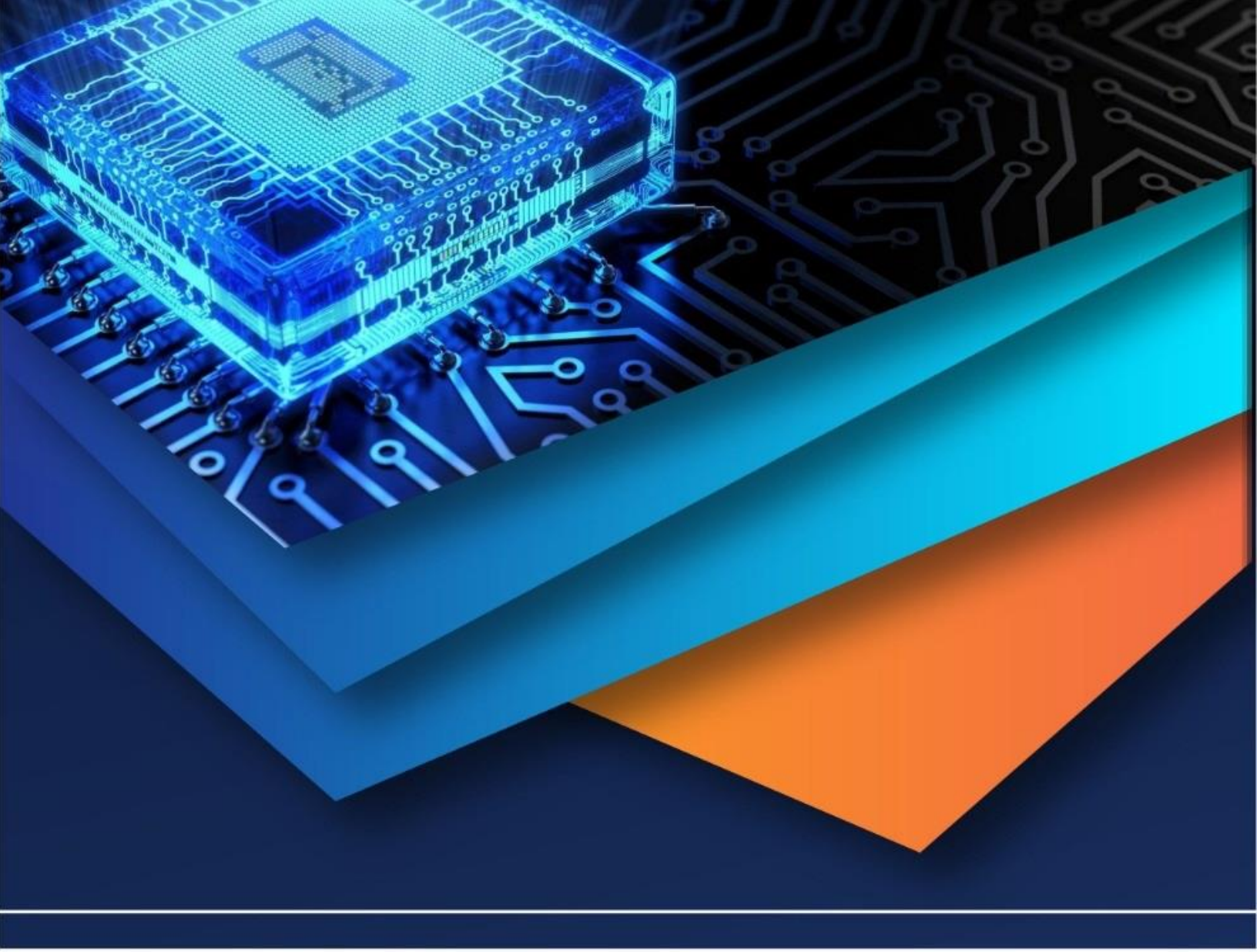
Having analyzed and perused various methods of implementing change detection, we have arrived at the conclusion that the most efficient and applicable method is to introduce panoramas of multitemporal video feeds of the same location into Siamese CNN, as input with knn methods to improve performance measures. We intend to use this in aerial surveillance application.

©IJRASET: All Rights are Reserved

REFERENCES

[1] P. Zhong and R. Wang, "A multiple conditional random fields ensemble model for urban area detection in remote sensing optical images," IEEE Trans. Geosci. Remote Sens., vol. 45, no. 12, pp. 3978–3988, Dec. 2007.
[2] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and k-means clustering," IEEE Geosci.Remote Sens. Lett., vol. 6, no. 4, pp. 772–776, Oct. 2009.
[3] J. Zhao, M. Gong, J. Liu, and L. Jiao, "Deep learning to classify difference image for image change detection," in Proc. IEEE Int. Joint Conf. Neural Netw., Jul. 2014, pp. 411–417.
[4] Mr.Sudharshan Duth P, Ms.Swathi Raj, "Object Recognition in Images using Convolutional Neural Network".
[5] C. Benedek and T. Szirányi, "Change detection in optical aerial images by a multilayer conditional mixed Markov model," IEEE Trans. Geosci. Remote Sens., vol. 47, no. 10, pp. 3416–3430, Oct. 2009.
[6] P. Singh, Z. Kato, and J. Zerubia, "A multilayer Markovian model for change detection in aerial image pairs with large time differences," in Proc. IEEE 22nd Int. Conf. Pattern Recognit., Aug. 2014, pp. 924–929.
[7] J. Liu, M. Gong, K. Qin, and P. Zhang, "A deep convolutional coupling network for change detection based on heterogeneous optical and radar images," IEEE Trans. Neural Netw. Learn. Syst., to be published.
[8] F. Gao, J. Dong, B. Li, and Q. Xu, "Automatic change detection in synthetic aperture radar images based on PCANet," IEEE Geosci. Remote Sens. Lett., vol. 13, no. 12, pp. 1792–1796, Dec. 2016.
[9] S. Zagoruyko and N. Komodakis, "Learning to compare image patches via convolutional neural networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2015, pp. 4353–4361.
[10] R. Hadsell, S. Chopra, and Y. LeCun, "Dimensionality reduction by learning an invariant mapping," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2. Sep. 2006, pp. 1735–1742.
[11] Touazi and D. Bouchaffra, "A k-nearest neighbor approach to improve change detection from remote sensing: Application to optical aerial images," in Proc. IEEE 15th Int. Conf. Intell. Syst. Design Appl., Dec. 2015, pp. 98–103.
[12] C. Benedek and T. Szirányi, "A mixed Markov model for change detection in aerial photos with large time differences," in Proc. IEEE 19th Int. Conf. Pattern Recognit., Sep. 2008, pp. 1–4.
[13] Yi Tan, Subhodev Das, Ali Chaudhry, "An Aerial Change Detection System Using Multiple Detector Fusion and Adaboost Classificatio". SRI Internatinal 201 Washington Road, Princeton, New Jersey,USA
[14] Y. Jia et al., "Caffe: Convolutional architecture for fast feature embedding," in Proc. 22nd Int. Conf Multimedia (ACM), 2014, pp. 675–678.
[15] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in Proc. IEEE Int. Conf. Comput. Vis., Feb. 2015, pp. 1026–1034.
[16] Jie Chen, Xiao YU, "Research on Cylindrical Panaromic Video Stitching and AR Prespective Observation Algorithm".

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)