INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Feature Based Image Registration Using ORB and CNN for Remote Sensing Images

Laukikkumar K Patel[1,2]*, Manish I Patel[3]

1 Ph.D. Scholar, Sankalchand Patel University, Visnagar, 384315, Gujarat, India
2 Lecturer, Department of Electronics and Communication Engineering, Government Polytechnic, Palanpur, Gujarat, India
3 Assistant Professor, Department of Electronics and Communication Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

OPEN ACCESS

Received: 18-07-2023

Accepted: 03-10-2023

Published: 13-11-2023

Citation: Patel LK, Patel MI (2023) Feature Based Image Registration Using ORB and CNN for Remote Sensing Images. Indian Journal of Science and Technology 16(42): 3803-3813. https://doi.org/10.17485/IJST/v16i42.1782

* **Corresponding author**.

laukiksky@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: To improve image registration by reducing the estimation error of the rotation transformation parameter under illumination change effect in remote sensing images using Oriented Fast and Rotated Brief (ORB) and Convolutional Neural Network (CNN). Also, to reduce computational complexity that can be increased due to use of CNN. **Methods:** The image registration process aligns two or more images geometrically and a novel feature based approaches for image registration is proposed here, where ORB and CNN are used to estimate rotation transformation parameter under illumination change effects in remote sensing images with novelty in generation of feature descriptor. The results of the proposed approach are compared with an approach which uses only ORB for both feature detection and descriptor generation. **Findings:** In the proposed approach, convolutional features from modified CNN are used in a novel manner with ORB descriptor to generate the final fusion descriptor. Also, this novel approach reduces the computational complexity by limiting the descriptor size that can be increased due to CNN. Here, three different combinations of CNN layers are provided for the generation of descriptor with ORB and this approach is also tested with transfer learning concept and other than remote sensing image, which shows improved results for taken cases. The results of novel approach show that the estimation of rotation transformation parameter and image registration is improved, and the estimation error is reduced to 0.1% to 0.9% for taken cases. **Novelty:** Novelty is provided in the generation of descriptor by fusion of CNN (modified visual geometry group (VGG19)) features with descriptor from ORB for reduction in estimation error, limiting descriptor size to reduce computational complexity, and improvement in image registration of remote sensing images. It is also tested for transfer learning case and other than remote sensing image, where improved results are also seen.

Keywords: Image registration; CNN; Remote sensing; VGG19; Oriented fast and rotated brief

1 Introduction

In the image registration process, the sensed image is aligned to the reference image to find the subtle changes in the two images. These images are of the same scene, but they can be taken from different viewpoints, or by different sensors, or at different times. There are various fields in which image registration is required, like in remote sensing for image mosaic, landscape planning, etc., in the medical field for monitoring tumor evaluation, specimen classification, etc., in computer vision for automatic change detection, target template matching, etc.

For image registration, Zitova et al.[1] classified image registration techniques as area based and feature based methods. Area based methods are applied on the pixel values and gray levels or colors that provide distinctive information, and feature based methods use features of the image, where distinctive information is provided by features of the image like corners, edges, lines, etc. In this paper, a feature based image registration approach is proposed.

In the process of image registration, generally, two images are taken. One is the reference image, and the other is the sensed image, which is processed to align with the reference image. There are four steps in the majority of image registration methods:[1] detection of features, matching of features, estimation of transform model, and image resampling and transformation. The processing steps of image registration are shown in Figure 1.

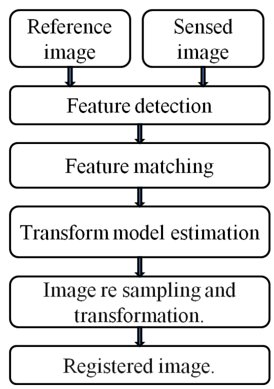

Fig 1. Steps of image registration

There are various challenges to image registration, but one of them is illumination change in multi-sensor, multi-spectral satellite images[2], which can affect the estimation of transformation parameters. It is important to have a better estimation of transformation parameters in the process of image registration. So, it is necessary to select appropriate feature detection and description methods for different applications to better estimate transformation parameters and improve image registration. This provided motivation to find different combinations of methods for feature detection and description to improve estimation of transformation parameters under illumination change effects in remote sensing images. A survey on various feature detection and description methods is found in paper[3], where a review of different methods for detection and description of handcrafted to learning based features is provided.

For feature detection and descriptor generation, various methods are available, like scale invariant feature transform (SIFT)[4], oriented fast and rotated brief (ORB)[5], etc. A comparative study of some methods is found in a paper[6], where a comparison of some algorithms of image registration is provided for rotation, scale change, etc., and it is found that the precision of SIFT is high for all cases, and it is also found that for applications of real time, ORB's performance is well and it is fastest for all taken cases. Also, Li et al.[7] compared four different methods for different light, different angle, change in scale, etc. conditions and found that ORB provided faster registration, a high rate of registration for taken cases, and used ORB for the actual task of detection of defect. In research paper[8], it is found that, due to the complexity of computation caused by having more feature points, instead of using a SIFT based descriptor, an ORB descriptor is used and proposed advanced registration method. From a review of research paper[9], it is found that ORB with phase congruence also provided good results for intensity variation and rotation distortion in remote sensing multimodal images. Karami et al. showed that compared to SIFT and SURF, ORB is faster and also provides good results for rotation angle proportional to 90 degrees[10]. In a research paper[11], it is shown that, under overexposure or underexposure conditions, improved ORB provides performance improvement of ORB. Also, Pei et al. showed that ORB with CNN provided improved results for aphid detection[12] and Rani et al. showed that, ORB provided good contribution in classification task[13].

Because of the improved results achieved by ORB related approaches, among various available methods, ORB is used in this paper for image registration process, but as shown in paper[9], for problems related to variation in intensity, ORB alone is not suitable, so to deal with this issue, it is required to use it with some other method.

In the last decade, machine learning-deep learning methods have also provided good results for applications related to image processing. Deep learning also provided good results in remote sensing[14]. Deep learning based approaches like deep neural network (DNN), CNN, etc. can be used for one or more steps of the image registration process. Like Gupta et al. used a segmentation network to extract semantic features for registration. In this proposed work for segmentation-based semantic features (SegSF) extraction, training of a LinkNet34 network is done for road segmentation for aerial images. These SegSF features consist of three components: (1) class label, (2) descriptor, and (3) keypoint location. They focused on registering multi-temporal high resolution nadir aerial images, which have large variations because of changing seasons, lighting conditions, etc. As per them, this proposed method has a limitation that its use is limited to images having visible roads[15]. By Zou et al. a deep learning network (self-supervised) called SAR-superpoint and transformation aggregation network (SSTA-Net) is proposed for SAR image registration. The proposed network has three parts, one for detection of feature points, another for matching, and third for unstable point removal. In this approach, SAR images are multi temporal[16]. Luo et al. used ResNet-50 for the registration of Unmanned Aerial Vehicle (UAV) images and found improved results compared to deep learning based algorithms and other algorithms taken for experimentation. But it found that there is a need of study to reduce computational complexity and also found that environmental changes like illumination changes, shadows, etc. on images from drones need to be studied[17].

Among various deep learning methods, CNN has achieved more attention in deep learning, and for remote sensing images, it also provided good results[18]. There are some approaches that used CNN for different steps of image registration, like Kuppala et al. used different CNNs in different stages of image registration for taken different approaches and discussed their effects[14]. Also, in some cases, researchers have used CNNs with classical methods like SIFT and SURF. Like by Ye et al., CNN is used with SIFT, and it provided improved results for the cases taken. In this approach, the VGG16 model is used, which is pretrained on a data set named ImageNet, and fine-tuning of the model is done to adjust the trained parameters using a custom data set. They have used the SIFT feature, which is a low-level feature, and the CNN feature, which is a high-level feature. In this case, for a specific keypoint, first, the SIFT descriptor is calculated. Then a patch of 64×64 pixels around that keypoint is given to fine-tuned VGG16 model to get the feature descriptor. For representing keypoint, these two feature vectors are transformed into one vector[19]. Patel et al. used VGG16 structure that is modified to generate feature descriptor from initial convolutional layers for each keypoint detected by SURF. Results of their proposed approach show that the feature descriptor generated by convolutional features from taken layers of CNN improved the correct match rate and hence improved the registration of satellite images that had illumination level change[20]. From a review of deep learning based papers, it is seen that lower layers of CNN detect low level features and higher layers detect high level features[19] and features from different layers can be used for improving the image registration process. So, features from some initial layers are used in a novel way for feature descriptor generation in our proposed approach.

From the review of traditional method based approaches, it is found that instead of using the same method for feature detection and description, a combination of different methods for feature detection and description can provide improved results for image registration for taken different conditions. Also, it is seen that ORB is invariant to rotation[5], but it is well suited for angles proportional to 90 degrees[10]. Also, it is shown that with intensity variation related problem, ORB is not suitable[9]. So, a proper combination of ORB with other method needs to be selected for better image registration for rotation parameter estimation under illumination change effect. From the review of deep learning based approaches, it is found that various deep learning based models or features from different models can be used for improvement in image registration. But it is seen that the use of these deep learning based approaches increases computational complexity, which needs to be reduced.

Hence, to overcome issues related to ORB and the use of deep learning based approaches, a novel approach with the use of CNN features with ORB is proposed here for improvement in the estimation of rotation parameter (small rotation like 10 degree) under illumination change effect in remote sensing images. The purpose of the proposed approach is to modify CNN (Visual Geometry Group (VGG19)) model for training and use a combination of features from initial different convolutional layers of modified VGG19 structure to generate a feature descriptor (in fusion with the descriptor from ORB) for each keypoint generated by ORB. Also, it is seen that, there is no single algorithm that can be used for registration of different types of image, so different algorithms need to be identified for registration of different types of images like remote sensing, computer vision etc. Here, this issue is also addressed to some extent by testing proposed approach with transfer learning case and for image other than remote sensing image. In both these cases improved results are also found.

The results from the proposed approach are compared with only one approach where feature detection and feature descriptor generation both are done using ORB only. Because, in the proposed approach, feature detection is done using ORB and novelty is provided in descriptor generation so actual comparison of proposed approach can be done with only approach which use ORB for detection and description both to show the improvement in results by novel descriptor generation process. The proposed approach shows improvement in transformation parameter (rotation) estimation and thus improvement in image registration due to the use of CNN features for images having illumination change. So, the main contributions of the paper are:

1. A novel approach to generate the final feature descriptor by fusion of descriptor from ORB and descriptor from different convolutional layers of modified VGG19 structure to limit the final descriptor size and reduce computational complexity, which could be increased due to the use of CNN.
2. Improvement in transformation parameter estimation (rotation), reduction in estimation error, and hence improvement in image registration.

2 Methodology

In the proposed approach, the original VGG19 structure is not used. It is modified for training purpose. This modified structure is trained with seven datasets of different types, like remote sensing images, computer vision images, etc., so that this modified structure can be used for different types of images and should not be trained separately for each image dataset, which saves time in training of the modified structure. In the proposed method, novelty is provided in the generation of the final descriptor, which not only improves rotation parameter estimation but also reduces computational complexity due to the limiting size of the descriptor which could be increased due to use of CNN. In this section, a brief review of ORB and VGG19 is provided, followed by modification and training of VGG19, the proposed method, and the datasets used.

2.1 Basics of ORB and CNN

2.1.1 ORB

In ORB [5], features from accelerated segment test (FAST) for keypoint detection and binary robust independent elementary features (BRIEF) for descriptor generator with some modifications are used. First, for keypoints' determination, FAST is used. Then, the top N points are found by applying Harris corner measure. FAST is a rotation variant, and computation of orientation is missing. Then patch's centroid (intensity weighted) is computed with located corner at the center, and orientation is found by the vector's direction towards the centroid from the corner point. The performance of the BRIEF descriptor degrades when there is in plane rotation. So, by using the patch's orientation, the rotation matrix is calculated in ORB and as per orientation, BRIEF descriptors are steered.

2.1.2 CNN

CNN consists of various different types of layers, like convolutional layer, fully connected layer, pooling layer, etc. Among various deep learning based CNNs, VGGNet [21] is one of them. In VGGNet, VGG16 and VGG19 architectures are generally used. Basically, VGG19 consists of more convolutional layers than VGG16. In VGG19, total convolutional layers are sixteen, and total fully connected layers are three. These sixteen convolutional layers in VGG19 are used for feature extraction, and they are segregated into five groups, with a max-pooling layer after each group. In the first and second groups, two convolutional layers are present and the rest of the groups consist of four convolutional layers. In convolutional layers, filters with small receptive fields 3×3 are used. In this structure, the convolutional stride of one pixel is used. In the max-pooling layer, for max-pooling, 2×2-pixel window is used and stride is kept to two. The number of channels (width of layers) of convolution layers starts from 64, and after each max-pooling layer, it increases by a factor of two and goes up to 512 channels. After the stack of convolutional layers, there are three fully connected layers with 4096 channels in the first and second layers and 1000 channels in the third layer.

2.2 CNN modification and training

The original VGG19 [21] structure is used for 1000 classes, but here, this structure is modified, where the number of channels in fully connected layers is changed and trained using seven datasets having different types of images like remote sensing images, computer vision images, so classes are seven at the last fully connected layer of the modified structure. As a modified VGG19 structure is trained with seven datasets together, the training time of the structure also reduces compared to training the structure for each dataset at a time separately. Here, each dataset contains one reference image and one sensed image. An overview of the original and modified VGG19 structures is shown in Figure 2. The modified VGG19 structure is trained using 64×64 patches of the reference image and augmented patches generated using image data augmentation during training. The training of the modified VGG19 structure is done in Google Colab with hyperparameters, learning rate=1e-4, epoch-10, batch size (training and validation each) = 5, etc. As training of model is done only for extraction of features from different layers, so no splitting of patches is applied during training and used same patches for training and testing both.

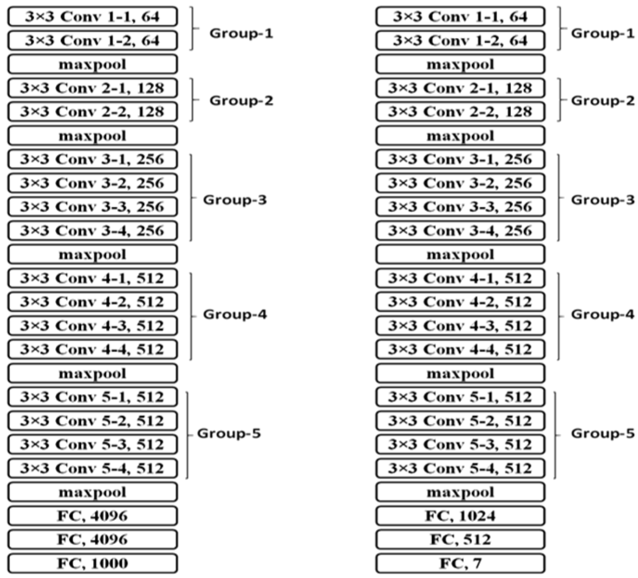

Fig 2. Overview of (a) Original VGG19 [21] an d (b) Modified VGG19 structure

2.3 Proposed Method

In the proposed approach, ORB is used for the detection of features from image. After the detection of features, a descriptor for each keypoint is generated by combining the feature descriptor generated by ORB and descriptors from multiple convolutional layers of the modified VGG19 structure with some modifications. In the process of descriptor generation, first descriptor from ORB is generated for each keypoint, which is of size 32. Then, from the image, 64×64 patch for each keypoint detected by ORB is taken and given to the modified VGG19 structure. In this structure, 16 convolutional layers are segregated into 5 groups as per the original structure. Output is taken from convolutional layers of three initial groups of modified VGG19 structure, and they are, 2nd layer (layer 1-2) that belongs to group-1, 4th layer (layer 2-2), which is in group-2, and 5th layer (layer 3-1) from group-3. Once descriptors from three different layers are generated, at a time, two layers' outputs are combined with a descriptor from ORB. In the proposed approach, three different layers' outputs are taken in combination of two layers at a time and combined with a descriptor from ORB. So, here, these combinations will be treated as, combination of 5th layer and 4th layer with ORB as (layer 5&4), combination of 4th layer and 2nd layer with ORB as (layer 4&2), combination of 5th layer and 2nd layer with ORB as (layer 5&2). As the size of a descriptor from different convolutional layers of the modified VGG19 structure is more (for the 2nd layer, the descriptor size is 262144, for the 4th layer, it is 131072 and for the 5th layer, it is 65536), the combined descriptor size will also be large. So in the feature matching stage, computation complexity will increase and it will take more time. So a novel approach to reduce (limit) the final descriptor size is proposed here, where the mean and variance of the descriptor generated from each taken convolutional layer are calculated, and these mean and variance values are appended with the descriptor from ORB instead of appending large size descriptors from layers of modified VGG19. Thus, the final descriptor consists of descriptor from ORB of size 32, the mean and variance values of one convolutional layer and mean and variance values of other convolutional layer. Hence, the size of the final descriptor will be 36 for each keypoint, which extensively reduces computations, also reduces computational complexity and saves time. Another novelty in this proposed approach is, most of the arrays used are of integer type, and the data type of the values of the descriptor used in matching is uint8. Also, the values are small, so to make them even higher, values of descriptor from ORB and mean and variance of descriptors from different layers of modified VGG19 structure are multiplied by 100 before combining. In the experiment, the precision of two digits after the decimal point is considered in the rotation parameter estimation.

Once the final descriptor for each keypoint is generated, the matching process is done using the brute-force matcher (https://docs.opencv.org/4.x/dc/dc3/tutorial_py_matcher.html). Then, found matches are sorted in ascending order of their distance, where matches with low distance come to the front. In the matching process, descriptor from the same combination of layers of the modified VGG19 structure for reference and sensed image are taken, like if layer 4&2 descriptor is taken for reference image, then for sensed image, layer 4&2 descriptor will be taken. After the completion of the matching process, the RANSAC[22] algorithm is used for outlier removal and transformation parameter estimation. Based on estimated parameters, the sensed image is registered. Here, in this experiment, comparison is done based on estimated transformation parameters for rotation. The block diagram and framework of the proposed method are shown in Figures 3 and 4 respectively.

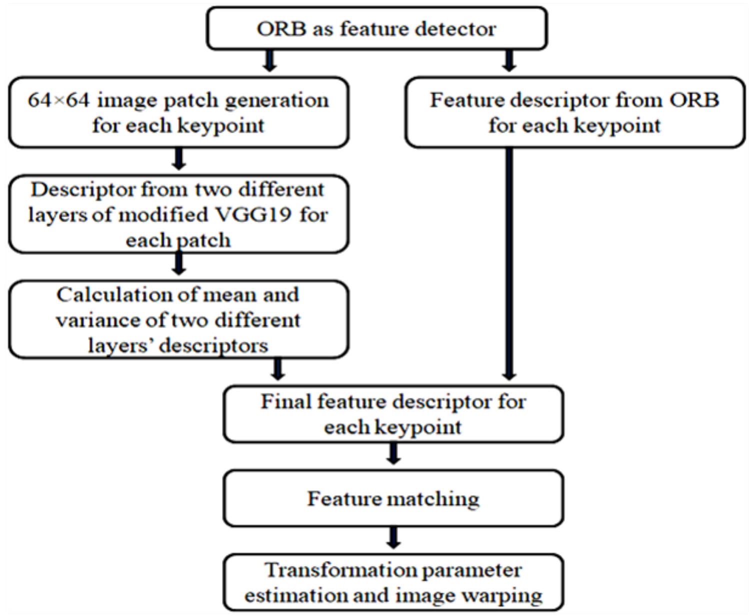

Fig 3. Block diagram of proposed approach with ORB and modified VGG19

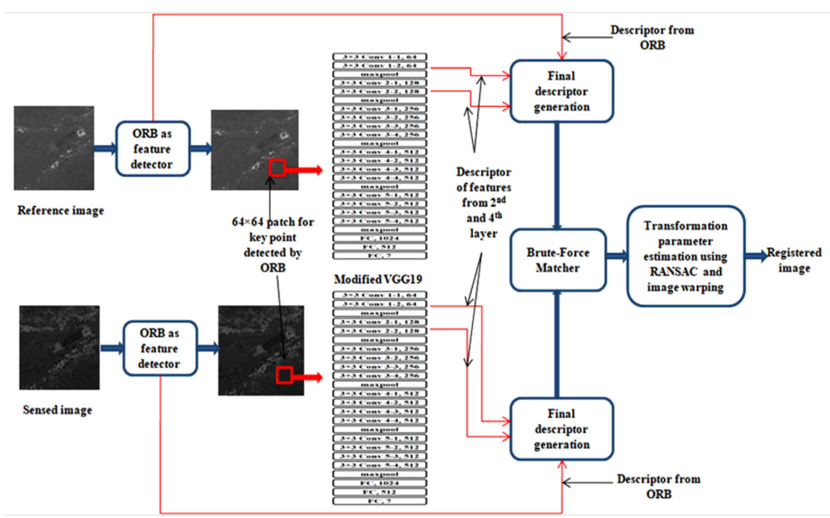

Fig 4. Framework of proposed approach with ORB and modified VGG19

2.4 Dataset

Here, in this experimentation, three remote sensing image datasets (Dataset-1 to Dataset-3) are used, which are also used in the training of the modified VGG19 structure. Also, the proposed method is tested for one r emote sensing dataset (Dataset-4) that is not used in the training process, which is also called transfer learning concept. Finally, the proposed method is also tested with one computer vision dataset (Dataset-5), which is also included in the training process to evaluate performance.

In order to reduce computational complexity at initial stage, reduced sized datasets are used in this experiment. So, size of Dataset-1 (From: https://bhuvan-app1.nrsc.gov.in/imagegallery/bhuvan.html) is reduced to 300×212, Dataset-2 (From: https://bhuvan-app1.nrsc.gov.in/imagegallery/bhuvan.html) is kept to size 300×300. Dataset-3 is used in paper [2], is taken here with some modification, and the size of it is kept to 300×300. Dataset-1 to Dataset-3 are remote sensing images and these are used in training of modified VGG19 structure. Dataset-4 (From: https://bhuvan-app1.nrsc.gov.in/imagegallery/bhuvan.html) is kept to size 300×300, which is also a remote sensing image but used as transfer learning case. Dataset-5 (From: www.vision.cs.rpi.edu/keypoints/) is taken as computer vision images and kept to size 200×300. Dataset-5 is also used in training of modified VGG19 structure. Reference and sensed images of all the datasets are shown in Figure 5.

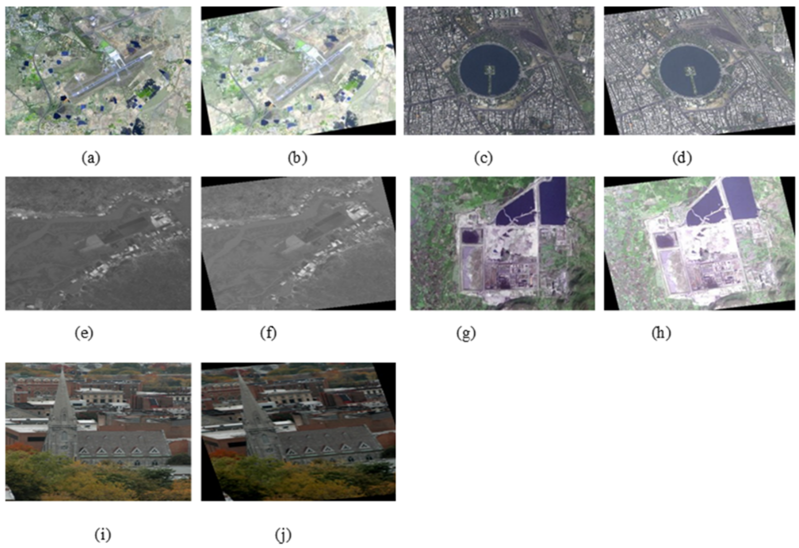

Fig 5. Dataset-1 ((a) reference image, (b) sensed image), Dataset-2 ((c) reference image, (d) sensed image), Dataset-3 ((e) reference image, (f) sensed image), Dataset-4 (g) reference image, (h) sensed image), Dataset-5 ((i) reference image, (j) sensed image)

3 Results and Discussion

In this experiment, the size of the datasets is reduced to minimize computational complexity at the initial stage, as CNN is used in feature descriptor generation process. Also, to reduce the size of the descriptor, the mean and variance of the feature descriptor from two different convolutional layers of the modified VGG19 structure are used and appended to the descriptor from ORB, which provides speed and ease in the calculation and descriptor matching process. This novelty in descriptor generation actually reduces computational complexity.

Here, a modified VGG19 structure is trained with different types of images, like remote sensing images, computer vision images, etc. In the training process, total seven datasets are taken, which include remote sensing and computer vision image datasets. Out of these seven datasets, four datasets are taken here, where rotation and illumination change effect is present in sensed images compared to reference image. In image registration, it is difficult to compare results if ground truth is not available. Hence, sensed images are generated manually, so that applied rotation parameters can be taken as ground truth. Here, sensed images are having an illumination change effect with 10-degree rotation compared to the reference image, which is considered as ground truth.

From the review of different papers, it is found that the combination of ORB with different other methods also provided improved results for image registration in different conditions, like the use of ORB with SIFT, phase congruence, etc. Further, it is found that ORB is used with CNN for detection related approaches like aphid detection, and ORB also provided improved results in classification related approaches. But it is a challenging task to find the combination of ORB and CNN for image registration for remote sensing images under illumination change effect. Hence, to take advantage of convolutional features with ORB, it is required to generate novel feature descriptor for improved results. As this proposed approach provides novelty in the descriptor generation method for features detected by ORB, the proposed approach is compared with the only approach where

ORB is used for both feature detection and descriptor generation so that the improvement can be compared in an effective way. Also, three different combinations of layers are provided in proposed method to check the effectiveness of the proposed novelty for different images taken. So, in this paper, the approach where ORB is used for both feature detection and descriptor generation is considered as Approach-1, and our proposed approach is considered as Approach-2 for further comparison purpose.

In this experiment, the feature matching process is done by brute-force matcher for both approaches, where L2 norm is used as distance measurement in our proposed approach (Approach-2) and for Approach-1, hamming distance is used instead of L2 norm as distance measurement because the ORB descriptor is a binary descriptor. The other parameter, crosscheck in brute-force matcher, is kept true, which provides only those matches where two features in both the sets match each other. Implementation of these experiments is done in pycharm with python 3.7, on intel core i7 processor with 2.80 GHz and 16 GB RAM.

After the completion of the matching process, the RANSAC algorithm is used for outlier removal and transformation parameter estimation. In the proposed approach, three different layers combinations are used, layer 5&4, layer 4&2 and layer 5&2. These three different layers combinations' results are compared with Approach-1. The results after using Approach-1 and Approach-2 in the image registration process, it is seen that Approach-2 provides better estimation than Approach-1.

First, the results of three remote sensing image datasets (Dataset-1 to Dataset-3) used in the training process and also involved in the image registration process are analyzed here. For Dataset-1, layer 5&2 estimates 9.32, layer 4&2 estimates 9.68 and layer 5&2 estimates 10.01, compared to estimation of Approach-1 that is 9.49, which shows that two combinations out of three provide better result compared to Approach-1.

In the case of Dataset-2, estimation of layer 5&4, layer 4&2, layer 5&2 are 9.72, 9.85, 9.99 respectively and all three combinations provided better results than Approach-1 estimation 9.56.

For, Dataset-3, layer 5&4, layer 4&2 and layer 5&2 estimated 10.73, 9.74, 9.94 respectively against estimation of Approach-1 that is 10.85, which shows improved results by all three combinations of Approach-2.

Dataset-4 is also a remote sensing image dataset, but it is not involved in the training process, so the performance of the proposed approach is tested for the transfer learning case here. For this case, layer 5&4, layer 4&2, layer 5&2 estimations are 10.09, 9.56, 9.87 respectively, in comparison with Approach-1 estimation 10.62. In the transfer learning case, improved estimation is also provided by the proposed approach.

Apart from remote sensing image datasets, the proposed approach is also tested for a computer vision image dataset, Dataset-5. In case of this dataset, estimations of layer 5&4, layer 4&2, layer 5&2 are 10.12, 9.96, 10.10 respectively, compared to Approach-1 estimation 9.53. In this case also estimated results are improved compared to Approach-1. Details of rotation parameter estimation for all five datasets are shown in Table 1.

Table 1. Rotation parameter estimation with Approach-1 and Approach-2

Dataset	Rotation applied	Approach-1 (Only ORB used for detection and description)	Proposed approach (Approach-2)		
			Layer 5&4	Layer 4&2	Layer 5&2
Dataset-1	10	9.49	9.32	9.68	10.01
Dataset-2	10	9.56	9.72	9.85	9.99
Dataset-3	10	10.85	10.73	9.74	9.94
Dataset-4	10	10.62	10.09	9.56	9.87
Dataset-5	10	9.53	10.12	9.96	10.10

Also, the comparison of the proposed approach with Approach-1 is done with respect to rotation parameter estimation error that is calculated by Equation (1).

$$Estimation\ error\ (\%) = \left(\frac{|Rotation\ applied\ -\ Rotation\ estimated|}{Rotation\ applied} \right) \times 100 \tag{1}$$

Here, rotation parameter estimation error is calculated with respect to applied rotation of 10 degrees in the experiment. Details of rotation parameter estimation error for all the five datasets is shown in Table 2.

Table 2. Rotation parameter estimation error (%) with Approach-1 and Approach-2

Dataset	Approach-1 (Only ORB used for detection and description)	Proposed approach (Approach-2)		
		Layer 5&4	Layer 4&2	Layer 5&2
Dataset-1	5.1	6.8	3.2	0.1
Dataset-2	4.4	2.8	1.5	0.1
Dataset-3	8.5	7.3	2.6	0.6
Dataset-4	6.2	0.9	4.4	1.3
Dataset-5	4.7	1.2	0.4	1

Results in Table 2, show that, except for one case of layer 5&4 for Dataset-1, estimation error reduces significantly for other cases. For Dataset-1, Approach-1 provides an error of 5.1% that is reduced to 3.2% and 0.1% by layer 4&2 and layer 5&2 respectively. In the case of Dataset-2, estimation error is reduced to 2.8%, 1.5% and 0.1% by layer 5&4, layer 4&2 and layer 5&2 respectively, compared to 4.4% of Approach-1. For Dataset-3, layer 5&4, layer 4&2 and layer 5&2 reduces estimation error to 7.3%, 2.6% and 0.6% respectively compared to Approach-1 with 8.5% error.

In the case of Datast-4, estimation error is reduced to 0.9%, 4.4% and 1.3% by layer 5&4, layer 4&2 and layer 5&2 respectively, compared to Approach-1 error 6.2%.

For, Datast-5, estimation error is 1.2%, 0.4% and 1% for layer 5&4, layer 4&2 and layer 5&2 respectively, compared to 4.7% error with Approach-1.

After the completion of the rotation parameter estimation process, the sensed image is registered based on the estimated values. The registered images for all datasets with Approach-1 and one of the combinations of Approach-2 are shown in Figure 6.

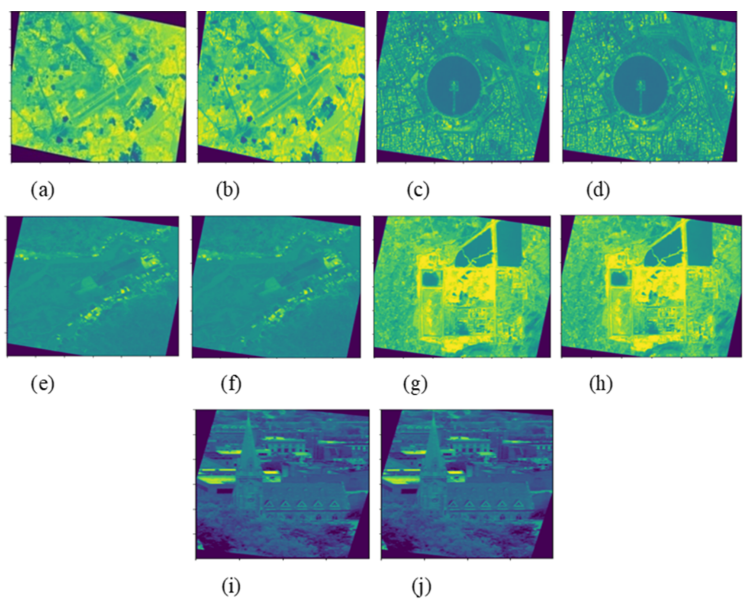

Fig 6. Registered images: Dataset-1((a) Approach-1, (b) Approach-2 (layer5&2)), Dataset-2 ((c) Approach-1, (d) Approach-2 (layer 5&2)), Dataset-3((e) Approach-1, (f) Approach-2 (layer 5&2)), Dataset-4 ((g) Approach-1, (h) Approach-2 (layer 5&4)), Dataset-5((i) Approach-1, (j) Approach-2 (layer 4&2))

Hence, from the analysis of estimated rotation parameter and estimation error for the proposed approach and approach where only ORB is used for detection description purpose, it is seen that, for remote sensing image datasets (Dataset-1 to Dataset-3), those are used in training purpose of modified VGG19 structure, proposed approach with all three combinations of layers provides improvised results in major cases, but layer 5&2 combination provides much improved results than Approach-1. For Dataset-4, considered as a transfer learning case, all combinations of the proposed approach provide improved results, but among all, layer 5&4 combination provides a much improved result. In the case of Dataset-5, the proposed approach with all combinations provides good results than Approach-1, but layer 4&2 provides a much improved results. From the overall analysis, it is found that the proposed approach with layer 5&2 provides improved results in terms of rotation parameter estimation compared to Approach-1.

As there is illumination change in reference and sensed image, and datasets are of different types, the content in various datasets can be different. Due to this, there may not be one single algorithm that can provide improved results in image registration for all kinds of datasets. Hence, it is required to apply different approaches for different conditions and different types of images. But in the case of the proposed approach with three different combinations of layers, all the combinations provided improved results, and one of them with superior results for different datasets for major cases compared to the approach taken in comparison. Hence, in this proposed approach, three different layer combinations are used to address the above issue to a reasonable extent.

4 Conclusion

Basically, image registration is the process of aligning two or more images in the same coordinate system. From the review of some research papers, it is seen that learning based methods can improve image registration but increases computational complexity, so in the proposed approach, learned features from different layers of modified VGG19 are combined with ORB descriptor in a novel manner to improve estimation of rotation parameter under illumination change effect in images and limit descriptor size to reduce computational complexity. The proposed approach is compared with the original ORB method which is used for both feature detection and descriptor generation in the image registration process. It is found that the proposed approach provides improved estimation towards ground truth for rotation parameter estimation and reduces estimation error to 0.1% to 0.9% compared to the estimation error of 4.4% to 8.5% of the approach taken in comparison under illumination change effect in reference and sensed images for taken different cases in the experiment. This shows that the use of convolutional features (learning based method) with the classical method ORB can improve image registration, and this proposed novel approach reduces the computational complexity even though CNN is used. The proposed approach also resolves the issue of selection of different methods for registration of different types of images to some extent for taken case in experiment. In the future, estimation of other transformation parameters instead of rotation can be tested with the proposed approach and can try different layers combinations or incorporate different CNN models for improvement in results.

References

1) Zitova B, Flusser J. Image registration methods: a survey. *Image and vision computing*. 2003;21(11):977–1000. Available from: https://doi.org/10.1016/S0262-8856(03)00137-9.
2) Patel MI, Thakar VK, Shah SK. Image Registration of Satellite Images with Varying Illumination Level Using HOG Descriptor Based SURF. *Procedia Computer Science*. 2016;93:382–388. Available from: https://doi.org/10.1016/j.procs.2016.07.224.
3) Joshi K, Patel MI. Recent advances in local feature detector and descriptor: a literature survey. *International Journal of Multimedia Information Retrieval*. 2020;9(4):231–247. Available from: https://doi.org/10.1007/s13735-020-00200-3.
4) Lowe DG. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*. 2004;60:91–110. Available from: https://doi.org/10.1023/B:VISI.0000029664.99615.94.
5) Rublee E, Rabaud V, Konolige K, Bradski G. ORB: An efficient alternative to SIFT or SURF. In: 2011 International Conference on Computer Vision, 06-13 November 2011, Barcelona, Spain. IEEE. 2012;p. 2564–2571. Available from: https://ieeexplore.ieee.org/document/6126544.
6) Ihmeida M, Wei H. Image Registration Techniques and Applications: Comparative Study on Remote Sensing Imagery. In: 2021 14th International Conference on Developments in eSystems Engineering (DeSE), 07-10 December 2021, Sharjah, United Arab Emirates. IEEE. 2022;p. 142–148. Available from: https://ieeexplore.ieee.org/document/9719538.
7) Li H, Yu Y, Du W. Research on rapid image registration based on feature point detection. In: Eighth Symposium on Novel Photoelectronic Detection Technology and Applications, Kunming, China;vol. 12169. SPIE. 2022;p. 964–969. Available from: https://doi.org/10.1117/12.2623954.
8) Shang MS. An image registration method based on SIFT and ORB. In: 2nd Conference on High Performance Computing and Communication Engineering (HPCCE 2022), Harbin, China;vol. 12605. SPIE. 2023;p. 366–373. Available from: https://doi.org/10.1117/12.2673245.
9) Alnagdawi MA, Hashim SZM. ORB-PC feature-based image registration. In: 2019 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), 17-19 September 2019, Kuala Lumpur, Malaysia. IEEE. 2020;p. 111–115. Available from: https://ieeexplore.ieee.org/document/8977793.
10) Karami E, Prasad S, Shehata M. Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images. 2017. Available from: https://doi.org/10.48550/arXiv.1710.02726.
11) Xie Y, Wang Q, Chang Y, Zhang X. Fast Target Recognition Based on Improved ORB Feature. *Applied Sciences*. 2022;12(2):1–14. Available from: https://doi.org/10.3390/app12020786.
12) Pei H, Liu K, Zhao X, Yahya AA. Enhancing aphid detection framework based on ORB and convolutional neural networks. *Scientific Reports*. 2020;10(1):1–15. Available from: https://doi.org/10.1038/s41598-020-75721-2.
13) Rani R, Kumar R, Singh AP. Implementation of ORB and Object Classification using KNN and SVM Classifiers. *International Journal of Computer Sciences and Engineering*. 2019;7(3):280–285. Available from: https://doi.org/10.26438/ijcse/v7i3.280285.
14) Kuppala K, Banda S, Barige TR. An overview of deep learning methods for image registration with focus on feature-based approaches. *International Journal of Image and Data Fusion*. 2020;11(2):113–135. Available from: https://doi.org/10.1080/19479832.2019.1707720.
15) Gupta A, Peng Y, Watson S, Yin H. Multitemporal Aerial Image Registration Using Semantic Features. In: International Conference on Intelligent Data Engineering and Automated Learning, IDEAL 2019: Intelligent Data Engineering and Automated Learning – IDEAL 2019. Lecture Notes in Computer Science;Springer International Publishing. 2019;p. 78–86. Available from: https://link.springer.com/chapter/10.1007/978-3-030-33617-2_9.

16) Zou B, Li H, Zhang L. Self-Supervised SAR Image Registration With SAR-Superpoint and Transformation Aggregation. *IEEE Transactions on Geoscience and Remote Sensing*. 2022;61:1–15. Available from: https://ieeexplore.ieee.org/document/10002310.

17) Luo X, Lai G, Wang X, Jin Y, He X, Xu W, et al. UAV Remote Sensing Image Automatic Registration Based on Deep Residual Features. *Remote Sensing*. 2021;13(18):1–15. Available from: https://doi.org/10.3390/rs13183605.

18) Liu X, Han F, Ghazali KH, Mohamed II, Zhao Y. A review of Convolutional Neural Networks in Remote Sensing Image. In: Proceedings of the 2019 8th International Conference on Software and Computer Applications. ACM. 2019;p. 263–267. Available from: https://doi.org/10.1145/3316615.3316712.

19) Ye F, Su Y, Xiao H, Zhao X, W M. Remote Sensing Image Registration Using Convolutional Neural Network Features. *IEEE Geoscience and Remote Sensing Letters*. 2018;15(2):232–236. Available from: https://doi.org/10.1109/LGRS.2017.2781741.

20) Patel LK, Patel MI. Feature based image registration using CNN features for satellite images having varying illumination level. *International Journal of Advanced Technology and Engineering Exploration*. 2023;10(101):440–457. Available from: https://doi.org/10.19101/IJATEE.2022.10100139.

21) Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014. Available from: https://doi.org/10.48550/arXiv.1409.1556.

22) Zuliani M. Computational Methods for Automatic Image Registration. Santa Barbara, California, USA. 2007. Available from: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=JHUtlXoAAAAJ&citation_for_view=JHUtlXoAAAAJ:IjCSPb-OGe4C.