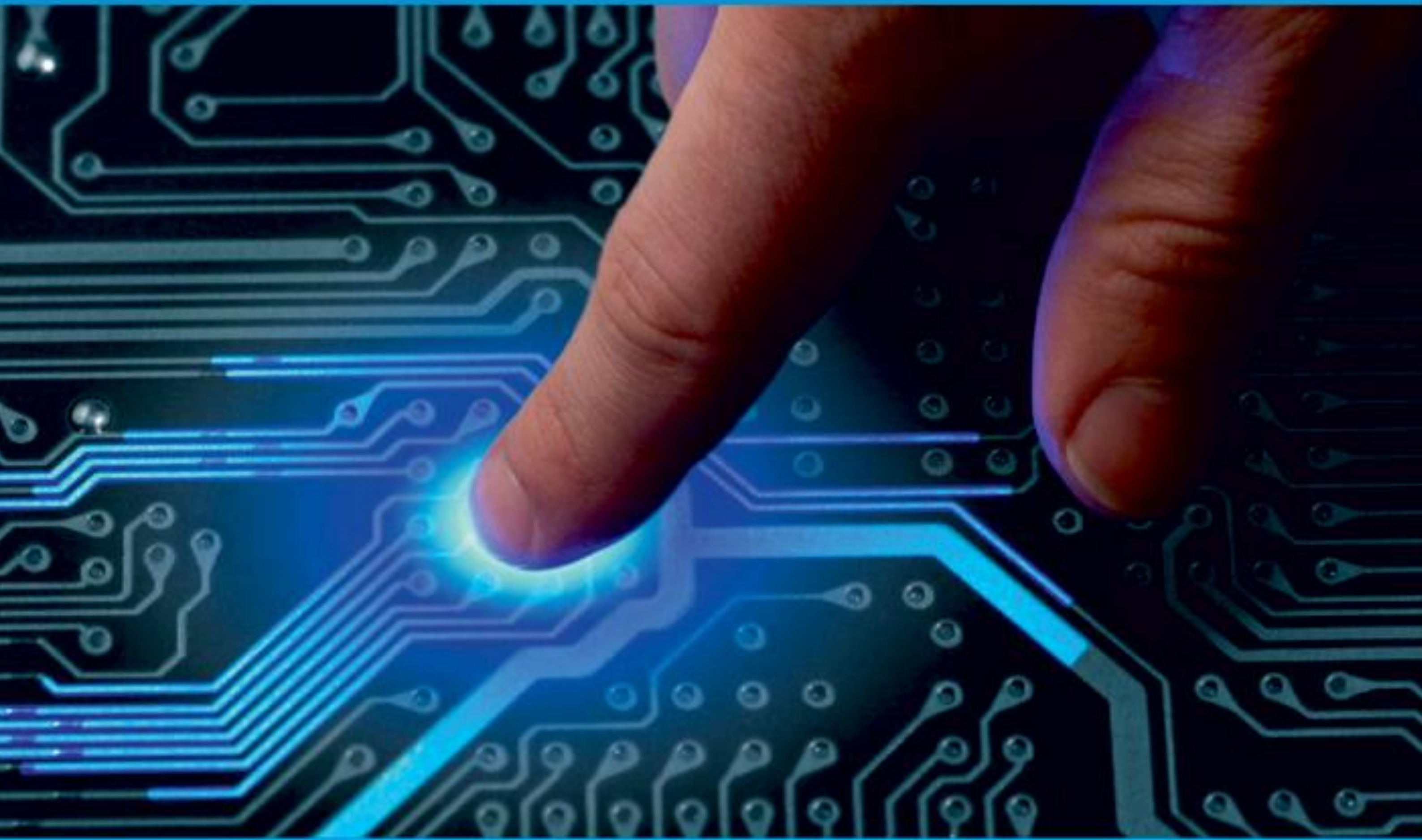
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207094|

Image Registration Using Deep Learning

Pragati Patel

Dr. Harisingh Gour University Sagar, Department of Computer Science and Application, Sagar, India

ABSTRACT: Image registration is a term used in computer vision and medical imaging to describe the process of analysing and aligning matches between multiple images. Owing of its robustness against subtle geometric distortion and grayscale variations, feature-based image registration is a popular method. The precision of transformation model approximations grounded in Random Sample Consensus (RANSAC) may be impacted by noise, occlusions, shadows, and modifications in picture content that contaminate the relevant feature point set, in practical scenarios.

This research aims to minimize the labour-intensive labelling procedure and enable deep neural network training by offering an automated approach for generating the training data. Through hypothesis sampling, learning guidance is used to define the probabilistic model estimation based on RANSAC. For generating the sampling probabilities, RANSAC is used with ProbNet neural networks for accurate estimates with a minimum set. Both qualitative and quantitative investigations are conducted to show the effective performance of the recommended model. Qualitative testing shows, the suggested approach performs better than competing techniques, and end-to-end learning is made possible by the integrating the model with the deep-learning framework, which further improves image registration accuracy.

KEYWORDS: Deep Learning, Random sample consensus, Image Registration, ProbNet, Neural Guided RANSAC.

I. INTRODUCTION

Image registration is a technique of overlapping multiple photographs of a scene, aiming to align the reference image and detected image geometrically. It involves the geometric adjustment and alteration of two images, the reference image and the sensed image, in order to account for variations in the conditions of their acquisition. Over a thousand publications on image registration have been published in the last ten years, according to the Institute of Scientific Information (ISI) database, demonstrating the substantial and ongoing interest in this field of study [8]. In many domains, programmers advance image registration, object identification, and data processing. The various image alignment and registration algorithms which exist are listed below: -

I. In order to align images, feature-based algorithms frequently use key point detection, local invariant descriptors, and keypoint matching techniques. These include Difference of Gaussians (DoG), Speeded up Robust Feature (SURF), Harris, Task Graph Based Fast Fourier Transform (GFFT) [9], Scale Invariant Features Transform (SIFT) [10].

II. Similarity measurements, such as mutual information, cross-correlation, and sum of squared intensity differences, are frequently used in medical applications for image registration.

III. Deep learning has gained popularity because it employs neural networks to automatically learn homography.

Using feature-based techniques for image alignment and registration, keypoints are used to identify significant elements in an input image. Local invariant descriptors quantify the area around each keypoint, and techniques like RANSAC and SIFT features compare and find correspondences. Here, we are predicting weights for observations using the combination of RANSAC and neural networks. The resulting sampling of minimal sets marks the algorithm as RANSAC with Neural Guidance [2].

International Journal of Innovative Research in Computer and Communication Engineering
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |
|| Volume 12, Issue 7, July 2024 ||
| DOI: 10.15680/IJIRCCE.2024.1207094|

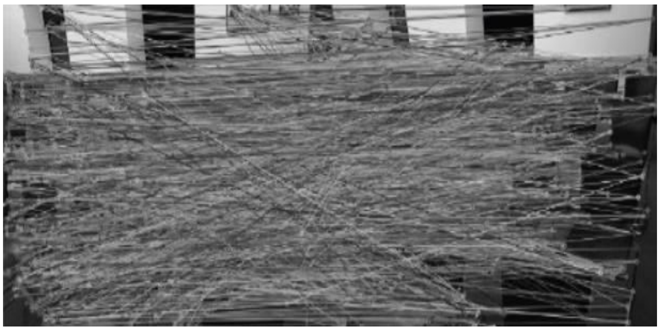


Fig 1. SIFT Correspondences

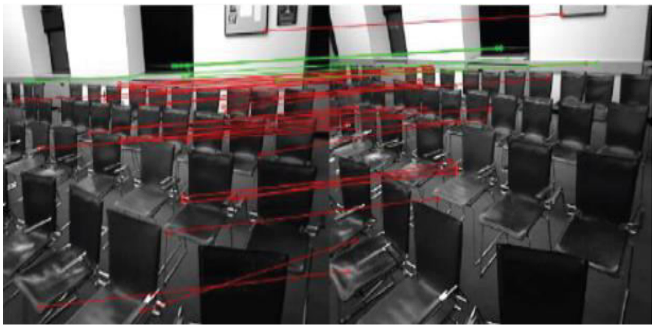


Fig 2. RANSAC

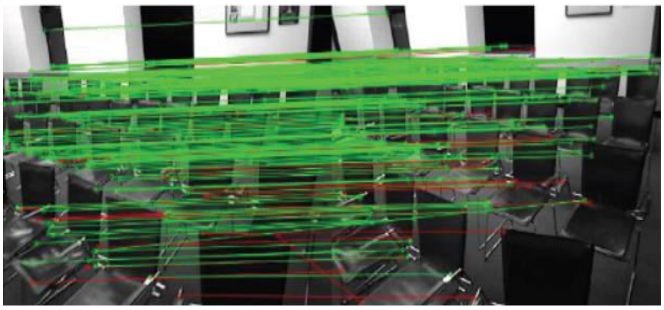


Fig 3. RANSAC with Neural Guidance

II. LITERATURE REVIEW

[1] A paper written by Zitová, Barbara & Flusser, Jan. reviews conventional and contemporary image registration techniques, focusing on feature-based approaches. It assesses their effectiveness using key processes like feature identification, matching, resampling, mapping function design, and transforming images. The study addresses common problems and provides valuable insights for future research.
[2] "Digital Image Registration Using Projections, by S. Alliney and C. Morandi in IEEE," This paper discusses a new picture registration algorithm using one dimensional Fourier transformations for image processing in aeronautical and clinical applications. It addresses challenges in observing unpredictable camera and scene
relationships, requiring preliminary rotation calculations for stable displays. Initial tests are described.
[3] Lisa Gottesfeld Brown, from Columbia Univ., New York, mentioned in a paper that aligning multiple photos under different circumstances is crucial in the image processing and evaluation for large-scale systems. Picture registration is essential for target recognition, global land use monitoring, stereo image matching, and medical image alignment, supporting accurate analysis and decision-making.
[4] A. K. Jain, Jianchang Mao and R. P. W. Duin, Research work states that "Pattern recognition involves supervised or unsupervised classification, with statistical approaches being the most studied. Neural network techniques and learning theory are increasingly used in recognition system design. Despite 50 years of research, complex patterns remain unsolved". This review paper compares well-known methods and highlights research topics and applications in this challenging field.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207094|

[5] Marc L. Kessler, Kristy K. Brock, Todd R. McNutt, Sasa Mutic and Hua Li reported in a paper that Radiotherapy software systems improve treatment planning and delivery by combining diverse picture data. Algorithms align planning and in-room images, aiding patient positioning and adaptive radiation. Real-time estimations and dose estimations facilitate effective plan adaptation. Rigid commissioning and quality assurance procedures ensure secure integration.

[6] L. Bruzzone, R. Guan, C. Yang, M. Wang and H. Zhao authored in a paper that Semantic segmentation in VHR remote sensing applications requires large parameter estimations. A MSFFL model, LiANet, improves accuracy with reduced parameters and enhanced attention modules. The LiANet, an efficient lightweight attention network, achieves promising performances.

[7] A paper work by R. Liu, L. Mi and Z. Chen proposes an adaptive fusion network (AFNet) for segmentation in remote sensing image, using a multiple level of architecture and scale-feature attention module. Extensive tests on two publicly available data sets demonstrate its effectiveness.

II. PROPOSED METHODLOGY

This section provides a detailed presentation of the ProbNet-RANSAC feature-based image registration pipeline, with an emphasis on the geometric transformation model evaluation procedure.

3.1 Convolutional neural networks (CNN): - Convolution neural networks are mostly used to extract features from images; convolutional, pooling, and fully connected layers are further classifications for these layers. The feature matrix is divided into discrete blocks by the pooling layer, which then uses the maximum or average value to reduce dimensionality. The fully connected layer nonlinearly mixes channel feature matrices and local characteristics to generate the output. To estimate the feature points using RANSAC from a probabilistic viewpoint, ProbNet was created.

3.2 Method: - Using the ProbNet probability, we created an accurate RANSAC transformation model. RANSAC is a reliable estimating technique that was developed by Fischler and Bolles [1]. This method offers a means of differentiating between the initial matchings' mismatches and the subset of matches. The model transformation is analysed to remove mismatches and estimate the parameters. This approach is reliable, noise-resistant, and removes mismatches. RANSAC places a lot of emphasis on choosing the proper threshold value. When a modest value is used, many matches are eliminated as mismatches, which lower the percentage of right matches and the overall number of matches. Additionally, choosing a high number for the threshold raises the alignment error and mismatch rate, both of which have a negative effect on the registration process. However, because it relies on the type of image, image distortion, and image attributes, optimum threshold value determination manually and experimentally is quite challenging.

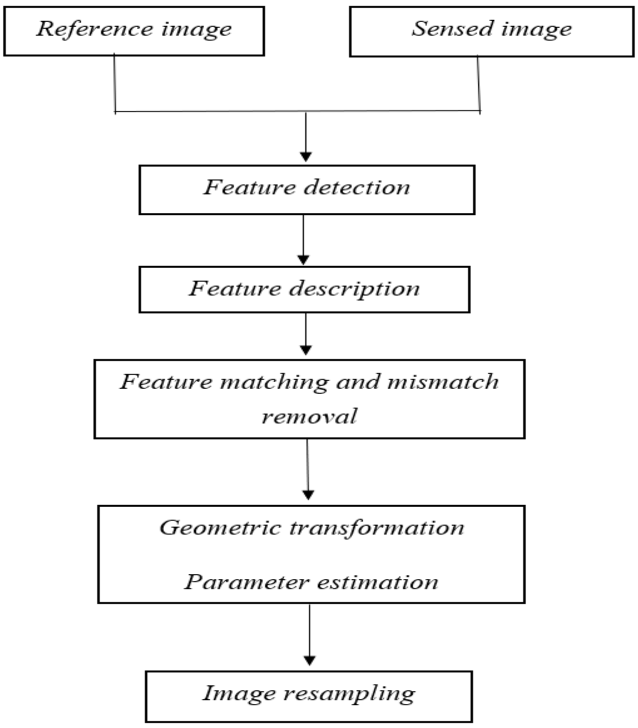


Fig 4. Flowchart for Image Registration Method.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |
|| Volume 12, Issue 7, July 2024 ||
| DOI: 10.15680/IJIRCCE.2024.1207094|

3.3 Neural Guidance: - The hypothesis pool S is created using RANSAC, that uses uniform random selection of observations. Instead, we want to sample data based on a learning distribution, which is defined by neural network parameters 'x'. Hence, the observations are selected as per the probability distribution, y ~ q (y; x). On the discrete set of observations C, q (y; x) is a categorical distribution. The goal is to maximize the likelihood of choosing outlier-free minimal sets, that will give precise estimation of 's', by learning parameters 'x' in this manner. By sampling observations and minimum sets separately, we sample a hypothesis pool S in accordance with q (S; x) [2], i.e.

$$q(S; x) = \prod_{i=1}^{M} q(s_i; x), \text{ with } q(s; x) = \prod_{j=1}^{N} q(y_j; x) \quad (1)$$

We suppose that during training, a task loss function ℓ(s), can be used to assess the estimate's quality. Our goal is to learn the distribution q (S; x) so that we have a high chances of getting a minimal task loss. Our goal is to minimize task loss, which we characterize as [3]:

$$L(x) = H_{S\sim q(S;x)}[\ell(s)] \quad (2)$$

We calculate the gradients between the network parameters and the estimated task loss as

$$\frac{\partial}{\partial x} L(x) = H_S[\ell(s) \frac{\partial}{\partial x} \log q(S; x)] \quad (3)$$

It is not practical to integrate over all potential hypothesis pools in order to compute the expectation. Consequently, we draw K samples of Sk~q (S; x) to approximate the gradients:

$$\frac{\partial}{\partial x} L(x) \approx \frac{1}{K} \sum_{k=1}^{K} [\ \ell(s) \frac{\partial}{\partial x} \log q(S_k; x)] \quad (4)$$

Equation 4's gradient variance may be high because of the sampling approximation. By deducting a baseline b, we employ a standard variance reduction method from reinforcement learning [4]:

$$\frac{\partial}{\partial x} L(x) \approx \frac{1}{K} \sum_{k=1}^{K} [\ \ell(s) - b] \frac{\partial}{\partial x} \log q(S_k; x)] \quad (5)$$

We discovered that, b = ℓ', provides a straightforward baseline [2].

3.3.1 Essential Matrix Estimation: -

Geometry between two images that are from the same scene is called epipolar geometry [5]. Specifically, two points, p and p′, in the left and right images, meet the condition $p'^T Mp = 0$, where M is the 3×3 matrix. We can use eight correspondences or seven correspondences with numerous solutions to estimate M uniquely [5]. The calibration parameters F and F′ of both cameras, S is a particular case of the fundamental matrix: $S = F'^T MF$. Based on five correspondences, one can approximate the essential matrix [6]. Initially, we evaluate RANSAC with Neural Guidance for the defined setting and compute necessary matrices using SIFT correspondences [7]. We closely follow their evaluation methodology and make comparisons with their findings.

IV. IMPLEMENTATION AND RESULT

4.1 Algorithm for Image Registration

Parameters: μ_0, n_{max}
Input: S, S', C_{ini}
Output: p_{fin}, C_{fin}

I. Step 1: $C_{best} \leftarrow \emptyset$
Step 2: n←0, initially
Step 3: reiterate
Step 4: Calculate the pose p that minimizes a specific cost function by selecting at random a sub-set A of the set C_{ini} that consists of three distinct correspondences.

$$M(p, A) = \sum_{(i,j)\in A} f(r_i, Q_{r,i}; r'_j; Q'_{r,j}; p) \quad (6)$$

where the function f defines how much each feature pair contributes to the total cost. Repeat if pose p obtained is not valid.

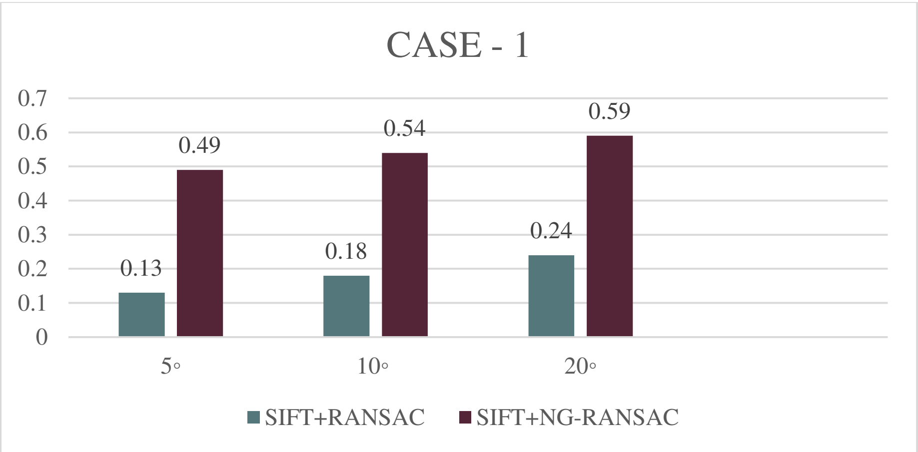
International Journal of Innovative Research in Computer and Communication Engineering
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |
|| Volume 12, Issue 7, July 2024 ||
| DOI: 10.15680/IJIRCCE.2024.1207094|

Step 5: Deduce the consensus set P(p).
Step 6: if $|P(p)| > |C_{best}|$ then
Step 7: $C_{best} \leftarrow P(p)$
Step 8: $p_{best} \leftarrow p$
Step 9: terminate if
Step 10: $n \leftarrow n+1$
Step 11: unless $n = n_{max}$
Step 12: Calculate the pose p_{fin} which reduces equation 6th for the set C_{best} .
Step 13: Evaluate the consensus set $P(p_{fin})$.
Step 14: $C_{fin} \leftarrow P(p_{fin})$
Step 15: return p_{fin},C_{fin}
Here, μ_0 is the predefined threshold,
and S, S' are the feature sets which are obtained from two different views.

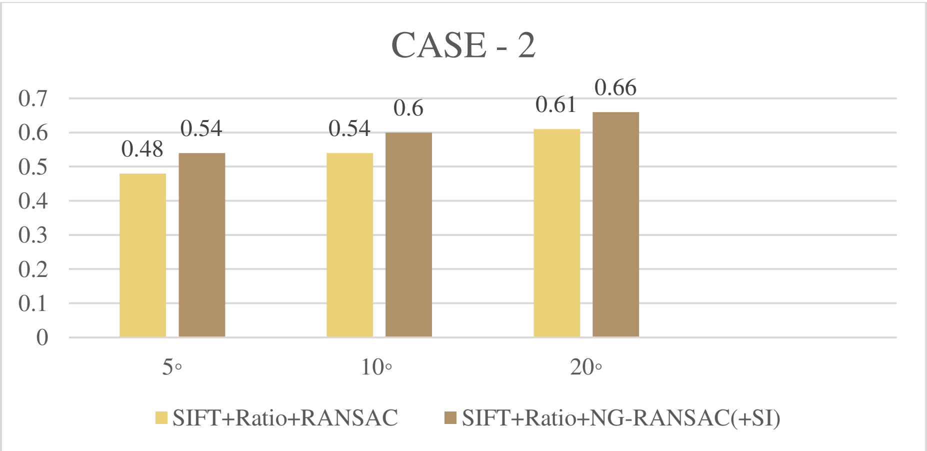
4.2 Evaluation Metric: We are able to retrieve the camera pose up using the essential matrix, which we then compare to the ground truth pose. We calculate the angular error in degrees between the pose translation vectors and the pose rotations. For the final angular error, we use the greater of the two numbers. For each test sequence, we compute the cumulative error curve and the area under the curve (AUC) up to a 5°, 10°, or 20° threshold [2]. Lastly, for every test sequence, we present the average AUC.

4.2.1 Graph: We calculate the pose of image from the essential matrix. We calculate the AUC of the total angular error up to a 5°, 10°, or 20° threshold.

a) without side information



b) with side information



| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |
|| Volume 12, Issue 7, July 2024 ||
| DOI: 10.15680/IJIRCCE.2024.1207094|

c) self-supervised

4.2.2 Results from graph: -

RANSAC ranks as the poorest approach since it produces subpar outcomes across all criteria. In the first case, we presume that no side information regarding correspondence quality is available. For the most part, NG-RANSAC performs better than Yi et al.'s InClass [7]. Both rely on SIFT correspondences, have the same network design, and employ RANSAC during testing. In the second case, we observe an enhancement in all methods. Classic techniques are better than learned method and hence side information is easily incorporated using NG-RANSAC [2]. In the third case, we guide NG-RANSAC by Self-supervised Learning. NG-RANSAC outperforms the opponent. Compared to supervised NG-RANSAC, unsupervised NG-RANSAC attains marginally lower accuracy.NG-RANSAC [2] optimizes the right target by including all of these elements throughout its training process.

4.3 Key Performance Indicators (KPIs)

Key Performance Indicators (KPIs) for image registration generally focus on the accuracy and efficiency of the registration process. Here are some important KPIs for image registration:

4.3.1 Sum of square distance: The sum of squared distances is frequently used as an objective function in the context of image registration to assess how dissimilar or mismatched two images are that need to be registered or aligned. Finding the best transformation to minimize the difference between the two photos is the objective. A popular metric for this is the sum of squared distances.

Calculate SSD: For each corresponding point (a_ref, b_ref) in the reference image and its transformed counterpart (a_target, b_target) in the target image, calculate the squared difference between their intensities (pixel values):

$$SSD = (T\_ref\ (a\_ref,\ b\_ref) - T\_target\ (a\_target,\ b\_target))^{2x}$$

where T_ref is the intensity value at point (a_ref, b_ref) in the reference image, and T_target is the intensity value at point (a_target, b_target) in the target image.

Sum the SSD: Sum up all the squared differences for all the corresponding points. This will give you the total sum of squared differences.

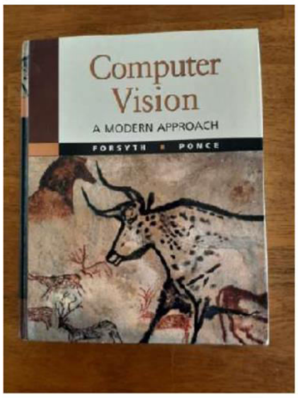
Result: - Sum of Squared Differences (SSD) for image registration: 11698868224.0

International Journal of Innovative Research in Computer and Communication Engineering

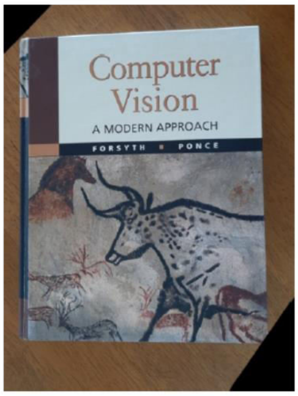
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207094|



img 1 Target image



img 2 Reference image

4.3.2 RMSE (Root Mean Squared Error): Registration Accuracy is a crucial KPI evaluating the alignment of features in input images, measured using metrics like RMSE, MAE, or Hausdorff distance. Root Mean Squared Error (RMSE) is a popular similarity metric used in image registration to measure the difference between comparable pixels in the two aligned pictures. The overall distance between matching pixels' intensity values is what is measured.

Use the steps below to compute the RMSE for image registration:

a. The two photos you want to register must be loaded.
Using the registration method of your choice, align the photos.
b. Identify the matching pixel positions where the difference between the aligned images is calculated

squared_diff = (image1 - image2) ** 2

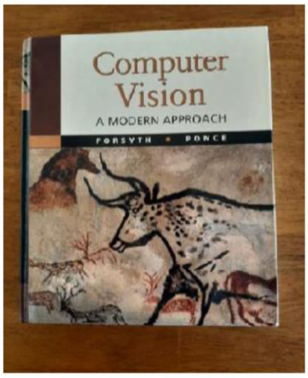
c. Calculate the differences' squared mean.

mean_squared_dif=np.mean(squared_diff)

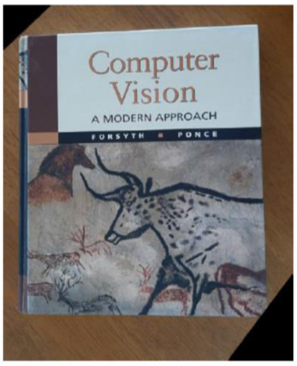
d. To find the RMSE, take the mean's square root.

RMSE = np. Sqrt (mean_squared_diff)

Result: - Root Mean Squared Error (RMSE) for image registration: 30.97567367553711



img 3 Target image



img 4 Reference image

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |
|| Volume 12, Issue 7, July 2024 ||
| DOI: 10.15680/IJIRCCE.2024.1207094|

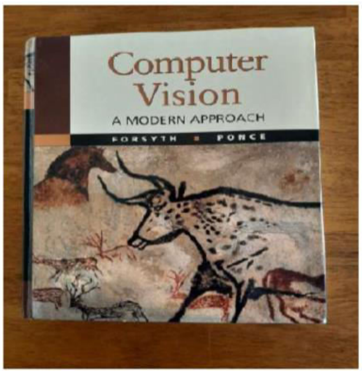
4.3.3 The Normalised Cross-Correlation (NCC) is a similarity statistic used in image registration to assess how well two images line up with one another. It measures how similar corresponding pixel values in the two images are, and is frequently used as an objective function to improve registration parameters. You can do the following procedures to calculate the normalised cross-correlation for picture registration:

The two photos you want to register must be loaded.
Using the registration method of your choice, align the photos.
Cross-correlation between the aligned pictures should be calculated.
To get the Normalised Cross-Correlation, normalise the cross-correlation.

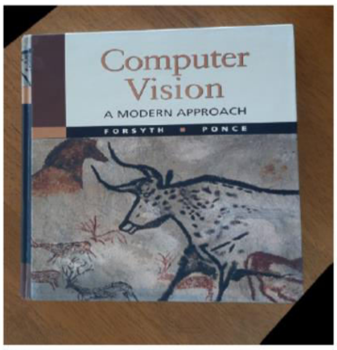
$$NCC = \Sigma\, ((\text{image1} - \mu 1) * (\text{image2} - \mu 2)) / (\sigma 1 * \sigma 2)$$

Where:

- $\mu 1$ and $\mu 2$ are the mean intensities of image1 and image2, respectively.
- $\sigma 1$ and $\sigma 2$ are the standard deviations of image1 and image2, respectively.



img 7 Target image



img 8 Reference image

Result: - Normalized Cross-Correlation for image registration: 0.9692180752754211

Table 1: Key Performance Indicators value for Image Registration

S.No.	Key Performance Indicator	Result
1.	Sum of Squared Differences	11698868224.0
2.	Root Mean Squared Error	30.97567367553711
3.	Normalized Cross-Correlation	0.9692180752754211

V. CONCLUSION

For this study we constructed a deep convolutional neural network based on ProbNet to predict the sampling probability of each associated feature point in the counterfeit related feature point collection. To increase the accuracy of image registration, a more accurate estimation of the transformation model was obtained by sampling the minimum set of RANSAC using the anticipated probability as a guide. Overall, the ProbNet-generated probabilities are crucial to the structure of RANSAC, it demonstrated how deep convolutional neural networks

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207094|

may be used for image registration. Neural Guided RANSAC is a robust estimator which uses guided hypothesis sampling as per the learned probabilities. We can apply Neural Guided RANSAC to many computer vision applications and can achieve a steady improvement over RANSAC alone.

REFERENCES

[1] Martin A. Fischler and Robert C. Bolles. Random Sample Consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM,1981.

[2] E. Brachmann, C. Rother, "Neural-Guided RANSAC: Learning Where to Sample Model Hypotheses", ICCV 2019.

[3] Eric Brachmann and Carsten Rother. Learning less is more- 6D camera localization via 3D surface regression. In CVPR, 2018.

[4] Richard S. Sutton and Andrew G. Barto. Introduction to Re- inforcement Learning. MIT Press, 1998.

[5] Richard I. Hartley and Andrew Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, 2004.

[6] David Nister. An efficient solution to the five-point relative pose problem. TPAMI, 2004.

[7] David G. Lowe. Distinctive image features from scale invariant keypoints. IJCV, 2004.

[8] Zitová, Barbara & Flusser, Jan. (2003). Image Registration Methods: A Survey. Image and Vision Computing. 21. 977-1000. 10.1016/S0262-8856(03)00137-9.

[9] Qinglin Lu, Xinyu Wang, Wenjing Ma, Yuwen Zhao, Daokun Chen, and Fangfang Liu. 2023. GFFT: a Task Graph Based Fast Fourier Transform Optimization Framework. In *52nd International Conference on Parallel Processing (ICPP 2023), August 07--10, 2023, Salt Lake City, UT, USA***.** ACM, New York, NY, USA 11

[10] X. Guo, J. Yang and H. Lin, "Image registration method based on improved SIFT algorithm and essential matrix estimation," 2017 IEEE International Conference on Information and Automation (ICIA), Macao, China, 2017, pp. 814-815, doi: 10.1109/ICInfA.2017.8079015.

[11] Dong, Y.; Liang, C.; Zhao, C. A Novel Remote Sensing Image Registration Algorithm Based on Feature Using ProbNet-RANSAC. *Sensors* **2022**, *22*, 4791.

[12] Cupec, Robert & Nyarko, Karlo & Kitanov, Andreja & Petrovic, Ivan. (2009). RANSAC-Based Stereo Image Registration with Geometrically Constrained Hypothesis Generation. Automatika: Journal for Control, Measurement, Electronics, Computing and Communications. 50. 195-204.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com