Indonesian Journal of Electrical Engineering and Computer Science
Vol. 36, No. 3, December 2024, pp. 1827~1835
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v36.i3.pp1827-1835

High-resolution aerial monitoring using DL for identifying abnormal activity based on visual patterns in drone videos

Mukesh Kumar Tripathi[1], Chellapilla V. K. N. S. N. Moorthy[2], Sandeep Kadam[3], Chaitali Shewale[4], Priya Shelke[4], Pravin R. Futane[4]
[1]Department of Computer Science and Engineering, Vardhaman College of Engineering, Hyderabad, India
[2]Department of Mechanical Engineering, Vasavi College of Engineering, Hyderabad, India
[3]Department of Computer Engineering, Keystone School of Engineering, Pune, India
[4]Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, India

Article Info

Article history:

Received Feb 16, 2024
Revised Aug 8, 2024
Accepted Sep 2, 2024

Keywords:

Average-based K-means
Clustering algorithm
Dirac depth silhouette function
OC-SVM
PEOO
UAVs

ABSTRACT

Unmanned aerial vehicles (UAVs) and sophisticated deep learning (DL) models have made the application of artificial intelligence (AI) more popular. This has resulted in an increase in the number of attempts to improve high-resolution aerial monitoring using DL for identifying abnormal activity based on visual patterns in drone videos. The study introduces a one-class support vector machine (OC-SVM) oddity locator for low-altitude, limited-scope UAVs used for ethereal video surveillance. The primary goal is to improve UAV-based observation capabilities by identifying areas or things of interest without prior knowledge, hence improving tasks like queue control, vehicle following, and hazardous product identification. The framework makes use of OC-SVM because of its quick and lightweight setup, making it suitable for continuous operation on low-computational UAVs. It empowers the identification of several peculiarities necessary for low-elevation reconnaissance by using textural characteristics to recognise both large-scale and tiny structures. Examine the UAV mosaicking and change location (UMCD) dataset to demonstrate the effectiveness of the framework, which achieves excellent accuracy and outperforms traditional methods by about one fifth in a variety of metrics. The suggested model compares with current methods, demonstrating superior accuracy and performance in recognition of peculiarities. Evaluation metrics include F1-score, review, exactness, and accuracy. The model demonstrates that it always encounters an oddity with a review compromise of up to seven on ten, achieving complete accuracy.

Corresponding Author:

Mukesh Kumar Tripathi
Department of Computer Science and Engineering, Vardhaman College of Engineering
Hyderabad, India
Email: mukeshtripathi016@gmail.com

1. INTRODUCTION

The rising usage of unmanned aerial vehicles (UAVs) or drones accentuates their portability and adaptability in different applications. Late headways have prompted the improvement of models that empower UAVs to handle constant insight, either freely or with the backing of a GPU-empowered ground station or wired PC. This capacity considers unassuming thinking in lightweight frameworks, upgrading the UAVs' capacity to adjust to dynamic conditions and tasks [1]-[3]. In addition, the joining of continuous knowledge empowers UAVs to follow occasions and give portrayals progressively. This usefulness is valuable in circumstances where human administrators might confront possible dangers. By recommending assessed ideal orders for activity, UAVs can act as allies to human administrators, improving situational mindfulness and supporting dynamic cycles. Generally speaking, the development of UAVs outfitted with constant knowledge capacities addresses a critical progression in advanced mechanics innovation, with expected applications across different businesses, including reconnaissance, search and salvage, ecological checking, and foundation examination [4]-[6].

The section highlights the rising pervasiveness of visual reconnaissance, principally using shut circuit TV (CCTV), intending to security concerns worldwide. As crime percentages rise, the reception of CCTV innovation has become typical in current life, particularly in areas requiring uplifted security measures [7]. In any case, successfully distinguishing irregularities or objects of interest from the broad volume of observation recordings represents a critical test. Ongoing headways in research have prompted the improvement of a different exhibit of present-day approaches for irregularity location (promotion) in observation films [8]. These methodologies are essential for improving safety efforts and distinguishing and anticipating mishaps, blockage, and different anomalies. Furthermore, promotion strategies can give significant factual bits of knowledge into street traffic status and other observed conditions. Thus, different PC vision-driven investigations and difficulties have been directed to address these requirements, including traffic checking, movement acknowledgment, crisis the board, human conduct examination, occasion discovery, and that's only the tip of the iceberg. These undertakings expect to use PC vision innovations to work on the productivity and exactness of inconsistency recognition frameworks, reinforcing general security and situational mindfulness across different spaces [9], [10].

The passage gives a clever outline of promotion inside the more extensive area of conduct getting it. It characterizes oddities as surprising, strange, sporadic, or eccentric occasions or things that stray from standard examples or datasets. These irregularities are set subordinate and contrast from winning examples inside the peculiarity of interest [11], [12]. In different spaces, for example, public security, wellbeing, sports examination, bunch movement checking, and visual observation, robotized promotion assumes a basic part. Computerized observation frameworks are valuable in clogged conditions as they can foresee phenomenal and complex circumstances, helping with pursuing reasonable well-being and crisis control choices. The section features the significance of observation procedures in testing and swarmed conditions like political conventions, occupied roads, air terminals, shopping centers, public festivals, and train stations. By recognizing and overseeing swarms, these reconnaissance methods add to public security and well-being and accomplish measurable targets. Coordinating computerized promotion inside reconnaissance frameworks improves situational mindfulness and supports proactive dynamics in different genuine world scenarios [13], [14].

The diagrams of the strategy and difficulties related to abnormality identification (promotion) models are grounded in conduct portrayal. Promotion models remove significant elements from observation information, like optical stream, spatiotemporal volume, shape, and direction [15]. These highlights act as contributions to the recognition cycle. Different methodologies, including principal component analysis (PCA), support vector machine (SVM), k-nearest neighbor (k-NN) calculation, and relationship examination, are utilized to use the extricated conduct highlights for abnormality identification. The essential goal of these structures is to identify peculiarities utilizing static and time-variation frameworks. Extra strategies like sliding windows (SWs) are incorporated with the methodologies above to address dynamic and time-variation oddities. Even with these endeavors, existing frameworks need assistance to catch time-variation model elements adequately, causing trouble describing strange contexts [16], [17]. The use of visual surveillance alongside cameras for security applications is becoming more expected with the ascent of wrongdoing all over the planet. Likewise, it has turned into a perceived piece of present-day life. CCTV is used for video surveillance in places that expect to be gotten. Consequently, successfully deciding potential oddities or objects of interest from many reconnaissance recordings has become arduous [18].

Dasariraju *et al.* [19] present a continuous group promotion framework using a convolutional long short-term memory (Conv-LSTM) organization. The methodology uses deep learning (DL) procedures to foresee rough exercises and continuously prepare partners for such events. The Conv-LSTM network is explicitly utilized to catch video outlines and recognize horrific acts inside the group. The framework proposed in the review shows further developed precision and speed contrasted with past techniques. Even with these headways, the framework's exactness should be improved by the intricacy of characterizing individual or gathering exercises precisely. This recommends that while Conv-LSTM networks offer promising capacities for continuous promotion in packed scenes, further refinement and enhancement might be essential to address the difficulties related to distinguishing and ordering different exercises inside the group. The review adds to the exploration of using profound learning procedures for constant group peculiarity recognition, featuring both the headway made and the continuous difficulties in accomplishing high precision in complex genuine world scenarios [20], [21].

While progress has been made in promotion models using conduct portrayal, difficulties, such as catching time-variation elements and tending to impediment issues, continue to happen. Further examination and progressions in artificial intelligence (AI) methods are essential to defeat these impediments and improve abnormality identification abilities in observation frameworks. The excess piece of this work is organized as follows: the recommended strategy is made sense of in segment 2; the proposed system's results and conversation focused on execution measurements are shown in area 3. At long last, segment 4 signifies the finish of the paper.

2. PROPOSED METHOD

The proposed methodology for a human anomaly detection (AD) system using UAV systems aims to enhance surveillance by identifying unusual human activities in video footage captured by UAVs [22]. The system is designed to operate in real time, leveraging advanced machine learning algorithms to detect anomalies without prior knowledge of the environment. The primary goals include developing a lightweight, efficient detection model suitable for the limited computational resources of small-scale UAVs and ensuring high accuracy and reliability in diverse and dynamic environments. The approach integrates state-of-the-art techniques in image processing and pattern recognition to effectively identify and classify anomalies, ultimately improving the efficacy of UAV-based surveillance applications. The nonlinear feature mapping (NFMF) component transforms input features into a more discriminative space, improving the model's ability to distinguish between normal and anomalous patterns. The double Bi-directional long short-term memory (DBiLSTM) architecture processes the sequential data in both forward and backward directions, ensuring a comprehensive understanding of temporal dependencies. This combined approach enables the detection system to robustly identify anomalies in real-time, even with the limited computational resources typical of UAV systems, thereby significantly improving surveillance capabilities.

Figure 1 visually represents the proposed human AD scheme for video surveillance using UAV systems. Block diagrams like this are essential for understanding complex systems' information flow and processing steps, offering a clear overview of component interactions and contributions to the method's functionality. This figure likely outlines the sequence of operations, starting from video input. Each block symbolizes a distinct processing step or component, illustrating how these elements integrate to achieve effective human AD in UAV-based video surveillance. The proposed NFMF-DBiLSTM model's working process for detecting human anomalies in video surveillance involves several key steps. Initially, video data is captured and preprocessed to enhance quality and remove noise. The NFMF module then transforms the preprocessed data into a more discriminative feature space. These transformed features are fed into the DBiLSTM network, which processes the data forward and backward, capturing temporal dependencies and patterns effectively. The DBiLSTM's bidirectional nature ensures a comprehensive sequence analysis, improving AD accuracy. Finally, the processed data is classified to identify anomalies, with the model leveraging its enhanced feature space and temporal understanding to provide accurate and efficient detection results suitable for real-time applications on UAV systems with limited computational resources.

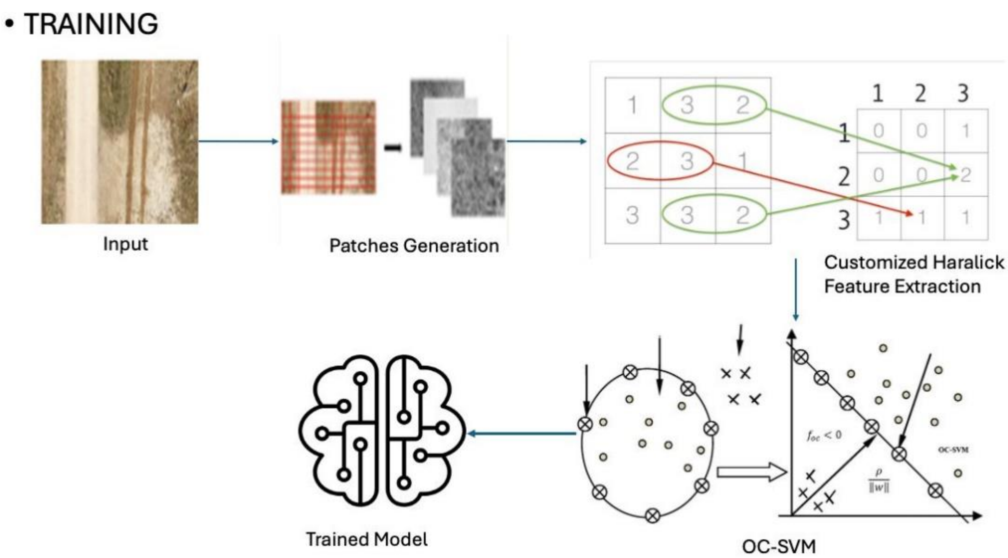

Figure 1. Block diagram of the proposed methodology

2.1. Frame conversion

In this process, since the video format is not suitable for classification purposes, the surveillance videos are transformed into a series of frames based on a specified time interval. This conversion ensures that each video is broken down into a manageable number of frames, which can then be used for further analysis. Let FFF denote the converted frames.

2.2. Preprocessing

These preprocessing steps are crucial for preparing the data for accurate and efficient AD. To detect betrayal behaviors, preprocessing is essential due to time constraints and the shadowing effect. This step involves enhancing the image quality and recognizing a set of targeted objects. The two main processes involved are:

– Duplicate frame removal: this step eliminates redundant frames to ensure that only unique and relevant frames are analyzed, thereby optimizing processing time and resources.
– Contrast enhancement: this step improves the visibility of the frames by increasing the contrast, which helps in better identification and differentiation of objects within the frames.

2.3. Background subtraction

For background and foreground subtraction using a Gaussian kernel, the density of an object within a group is estimated through a kernel function. The kernel function is K, the targeted object is x, and the bandwidth parameter h controls the smoothing. The bandwidth h is derived using probabilistic matrix factorization and the background-subtracted image is then used in subsequent phases for silhouette formation, where the contours and shapes of the detected objects are extracted for further analysis. After the background has been subtracted and the density of objects has been estimated using the Gaussian kernel, the next step is to form silhouettes of the objects [23]. This involves further processing of the background-subtracted image to extract the contours and shapes of the detected objects.

2.4. Clustering

The sorting and average-based K-means (SA-KM) clustering algorithm is employed following the silhouette function to cluster non-human, human, and suspected human categories from background-subtracted images [24]. Typically, K-means groups pixels into these categories by computing distances between incoming pixels and existing clusters. However, traditional models face the outlier problem. To address this, a sorting and median method is proposed. This method sorts the pixels and uses median values to improve the clustering process, effectively mitigating the impact of outliers and enhancing the accuracy of the human and non-human classification in the images.

2.5. Feature extraction

Feature extraction (FE) transforms raw data into numerical features that can be effectively processed while preserving the information contained in the original dataset. The data generated through FE typically yields better outcomes in classification tasks using machine learning than raw data [25]. The primary features extracted include the angle between joints, floor clip plane, height, width, velocity, acceleration, and differences between frames. Additional statistical features such as mean, standard deviation, skewness, and variance are considered. Moreover, dynamic features like the maximal Lyapunov exponent, correlation dimension, motion, size, and texture are also extracted to represent the data comprehensively.

2.6. Feature selection

Utilizing the prediction error-only optimization (PEOO) algorithm, vital features are chosen after FE. The PEOO algorithm mimics the food-searching behavior of Eurasian Oystercatchers probing for mussels. In this algorithm, each bird in the population serves as a search agent. The optimization ensures a balance in the mussels' size, calories, and energy. The conventional algorithm considers account length, power, and caloric value, all represented by random numbers ranging from 3 to 5. However, this research model selects these values using the Poisson function to enhance the selection process and prevent imbalances, ensuring optimal feature selection for subsequent classification tasks. Here, the mean number of iterations is notated as λ. Grounded on trial and testing, these values are chosen. The features are chosen resembling this way of hunting mussels. $\mathfrak{I}_n$ denotes the selected features. The proposed PERO pseudocode is given to Algorithm 1.

Algorithm 1. PERO pseudocode

Input: Extracted features $\aleph^{m}_{sus(ex)}$

Output: Selected features $\mathfrak{I}_n$

```
Begin
        Initialize population, E_fin
        Calculate the fitness of each search agent
        Identify the best solution ℵ^best_sus(ex)
        While i = 1
                For i = 1 to M
                        Compute time, energy, and caloric value
                        Update the position of the solution
                End for
        Calculate the fitness of each search agent
        End while
Return ℑ_n
End begins
```

3. RESULTS AND DISCUSSION

The proposed technique's performance is evaluated and compared with prevailing techniques implemented using the Python platform. The evaluation process involves several steps to comprehensively analyze the proposed human AD system. The proposed UAV systems involve several detailed steps, algorithms, and techniques to ensure replicability and validation of findings. High-resolution video data is initially collected from UAVs flying at low altitudes to capture detailed imagery. The NFMF algorithm transforms these frames into a more discriminative feature space, enhancing the model's ability to identify anomalies. Textural features are then extracted using local binary patterns (LBP) and gray-level co-occurrence matrix (GLCM), which capture the images' essential micro and macro structures [26]. The model is trained using an 80-20 train-test split, with the Adam optimizer and a learning rate 0.001 to minimize the loss function. The final model classifies frames as normal or abnormal, with performance evaluated using accuracy, precision, recall, and F1-score metrics. The system demonstrates superior performance compared to existing methods, with NFMF enhancing feature discrimination and DBiLSTM effectively capturing temporal patterns, thus validating the robustness and suitability of the process for real-time UAV-based surveillance applications.

3.1. Performance analysis

The method introduces a new spatial relationship, which refers to the arrangement of patches in the image, represented by a discretized circumference that considers displacements along all directions, thereby improving patch representation and ensuring rotation invariance. Generalized equations were developed to compute Haralick's textural features accurately within the proposed spatial relationship. Different patch sizes and circumferences with varying radii were evaluated to optimize performance. Overall, this approach demonstrates effectiveness in AD via textures, particularly in low-altitude aerial images, offering real-time capabilities and a lightweight implementation. This method leverages textural features and machine learning techniques with specific enhancements to traditional methods, significantly improving performance and providing a promising approach to AD in UAV-based surveillance applications. Here, the proposed model's clustering performance, centered on clustering accuracy and time, is analyzed and compared with existing methods such as K-means, fuzzy C-means (FCM), and farthest first clustering (FFC). The comparisons highlight the improvements in accurately categorizing non-human, human, and suspected human pixels and the efficiency in processing time against the traditional clustering techniques.

The images appear in Figure 2 demonstrate different stages or methods of processing aerial imagery for detecting and classifying regions, with a particular focus on identifying and highlighting the presence of a person. Figure 2(a) shows an aerial view of a road with a person walking along it. The image is in natural colors and appears to be taken during the day. A processed version of the aerial view overlaid with a grid and color-coded sections. The person is highlighted in red, indicating a detected region. The background is divided into green and blue sections, displaying different types of terrain or areas depicted in Figure 2(b). The Figure 2(c) grid and color-coding are present, with the person still highlighted in red. The green and blue sections vary slightly, potentially showing another type of terrain or region classification. Another variation of the processed aerial view with a grid and color-coded sections. The person is again highlighted in red, but there is also a yellow section near the person. The rest of the area is divided into green and blue sections, similar to the previous images depicted in Figure 2(d).

Table 1 compares the proposed model's performance against existing clustering methods, explicitly focusing on clustering accuracy. Clustering accuracy is a pivotal metric that evaluates how precisely objects are grouped into clusters while accounting for outliers. The proposed model achieves a remarkable clustering accuracy of 98.92%. This performance is notably superior, with the proposed model outperforming K-means by 2.92%, FCM by 3.80%, and FFC by 5.92 %. These improvements highlight the model's advanced capability in effectively handling outliers and ensuring accurate object clustering. The results underscore that the proposed model is significantly more efficient than the traditional clustering methods. K-means, while effective, need to catch up to the proposed model by a measurable margin, indicating its potential shortcomings in outlier management. Similarly, FCM, despite its nuanced approach to clustering through fuzzy logic, does not match the proposed model's precision. FFC, which relies on initializing clusters with the farthest points, exhibits the lowest accuracy among the compared methods, emphasizing the proposed model's superiority in complex clustering scenarios. Overall, the proposed model demonstrates exceptional performance, making it a highly reliable and efficient solution for clustering tasks that require high accuracy and robust outlier handling.

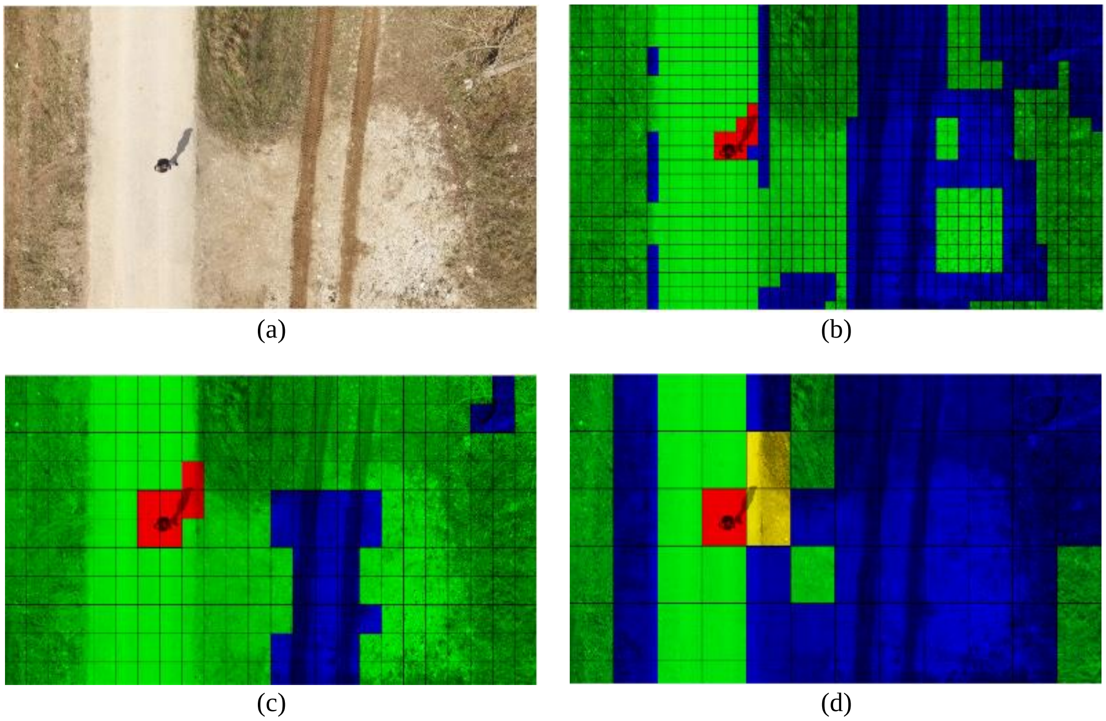

Figure 2. Different stages or methods of processing aerial imagery for detecting and classifying regions; (a) output generated using different patch sizes, (b) the red and green patches indicate, respectively, (c) anomalous and normal patches correctly classified, and (d) blue patches correspond to normal patches classified

Table 1. Clustering accuracy analysis

Method	Accuracy
Proposed SA-KM	98.92%
K-means	96%
FCM	95.12%
FFC	93%

The performance results of the proposed and existing methods are unveiled in Figure 3, with a specific focus on clustering time, which is the duration required to complete the grouping process. The proposed model demonstrates its superior efficiency through remarkably effective time management. The clustering time for the proposed model is 5,325 milliseconds (ms), which is substantially lower compared to the prevailing models. In comparison, the K-means algorithm requires 9,452 ms to complete the clustering process, highlighting the proposed model's significant time-saving advantage. Similarly, FCM has a clustering time of 10,769 ms, and FFC takes 13,544 ms, both of which are considerably higher than the proposed model. This stark difference in clustering times underscores the proposed model's efficiency in handling large datasets and complex clustering tasks more swiftly. The analysis proves that the proposed SA-KM model not only maintains high accuracy but also achieves faster clustering times, making it a highly efficient and practical solution for clustering applications that demand both precision and speed. The reduced clustering time of the proposed model enhances its usability in real-time applications and large-scale data processing, where time efficiency is crucial. Overall, the results validate the proposed model's superior performance in achieving faster clustering times without compromising accuracy, thereby establishing it as a more efficient alternative to traditional clustering methods.

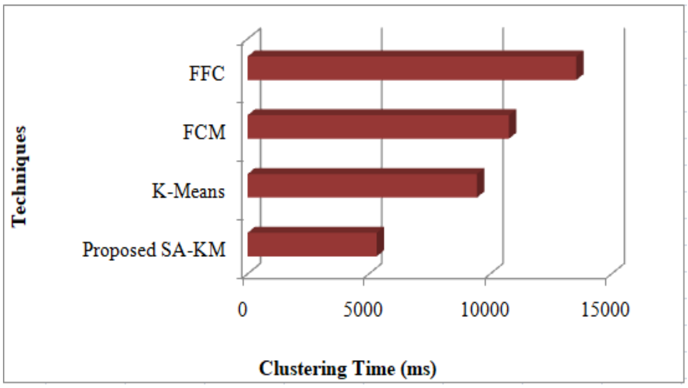

Figure 3. Clustering time analysis

4. CONCLUSION

This work introduces a novel, lightweight approach for real-time anomaly detection in low-altitude aerial images. The method begins by converting input frames to grayscale and dividing the photos into patches for detailed analysis. GLCMs are extracted from these patches, and textural statistics are derived to represent the patches. An OC-SVM is then used to detect anomalies based on the extracted textural statistics. This approach enhances traditional Haralick textural features, typically based on single offset displacements. Experiments conducted on the UMCD dataset demonstrate the effectiveness of this approach under various conditions, highlighting its novelty due to the lack of existing methods addressing anomaly detection through textures on this dataset. Comparisons with baseline Haralick textural features using single displacements confirm the proposed method's effectiveness. The importance of patch size and circumference radius was also identified, proving critical for achieving satisfactory performance, especially in security applications where anomaly detection is essential. Based on the results of all metrics, it is concluded that the proposed model is highly efficient compared to the prevailing models. However, this work does not focus on objects when they are in a group. Therefore, future work will be extended with advanced techniques to effectively consider and analyze people when they are in a group.

REFERENCES

[1] F. Zhu, X. Zhu, Z. Huang, M. Ding, Q. Li, and X. Zhang, "Deep learning based data-adaptive descriptor for non-rigid multi-modal medical image registration," *Signal Processing*, vol. 183, p. 108023, 2021, doi: 10.1016/j.sigpro.2021.108023.

[2] M. L. Terpstra, M. Maspero, A. Sbrizzi, and C. A. T. van den Berg, "⊥-loss: a symmetric loss function for magnetic resonance imaging reconstruction and image registration with deep learning," *Medical Image Analysis*, vol. 80, p. 102509, 2022, doi: 10.1016/j.media.2022.102509.

[3] Z. Zheng, W. Cao, Z. He, and Y. Luo, "Progressive anatomically constrained deep neural network for 3D deformable medical image registration," *Neurocomputing*, vol. 465, pp. 417–427, Nov. 2021, doi: 10.1016/j.neucom.2021.08.097.

[4] P. Singh, M. K. Tripathi, M. B. Patil, Shivendra, and M. Neelakantappa, "Multimodal emotion recognition model via hybrid model with improved feature level fusion on facial and EEG feature set," *Multimedia Tools and Applications*, pp. 1–36, 2024, doi: 10.1007/s11042-024-19171-2.

[5] A. Lara-Hernandez *et al.*, "Deep learning-based image registration in dynamic myocardial perfusion CT imaging," *IEEE Transactions on Medical Imaging*, vol. 42, no. 3, pp. 684–696, 2023, doi: 10.1109/TMI.2022.3214380.

[6] Q. Zheng, Q. Wang, X. Ba, S. Liu, J. Nan, and S. Zhang, "A medical image registration method based on progressive images," *Computational and Mathematical Methods in Medicine*, vol. 2021, pp. 1–9, 2021, doi: 10.1155/2021/4504306.

[7] G. A. Benvenuto, M. Colnago, M. A. Dias, R. G. Negri, E. A. Silva, and W. Casaca, "A fully unsupervised deep learning framework for non-rigid fundus image registration," *Bioengineering*, vol. 9, no. 8, p. 369, 2022, doi: 10.3390/bioengineering9080369.

[8] H. Siebert, L. Hansen, and M. P. Heinrich, "Learning a metric for multimodal medical image registration without supervision based on cycle constraints," *Sensors*, vol. 22, no. 3, p. 1107, 2022, doi: 10.3390/s22031107.

[9] Y. Fu, Y. Lei, T. Wang, W. J. Curran, T. Liu, and X. Yang, "Deep learning in medical image registration: a review," *Physics in Medicine and Biology*, vol. 65, no. 20, p. 20TR01, 2020, doi: 10.1088/1361-6560/ab843e.

[10] M. K. Tripathi, M. Neelakantappa, P. N. Mahalle, S. V. Channapattana, G. Deshmukh, and G. Prashant, "Detection and classification of mango fruit-based on feature extraction applying optimized hybrid LA-FF algorithms," in *Data-Centric Artificial Intelligence for Multidisciplinary Applications*, Boca Raton: Chapman and Hall/CRC, 2024, pp. 177–185.

[11] M. K. Tripathi, P. K. Reddy, M. Neelakantappa, C. V. Andhare, and Shivendra, "Identification of mango variety using near infrared spectroscopy," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 31, no. 3, pp. 1776–1783, 2023, doi: 10.11591/ijeecs.v31.i3.pp1776-1783.

[12] C. Papadakis, Y. Yin, M. Danikas, and C. Charalambous, "Surface discharges and flashover voltages in nanocomposite XLPE samples," *Engineering, Technology & Applied Science Research*, vol. 8, no. 6, pp. 3502–3504, 2018, doi: 10.48084/etasr.2186.

[13] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017, doi: 10.1016/j.media.2017.07.005.

[14] H. R. Boveiri, R. Khayami, R. Javidan, and A. Mehdizadeh, "Medical image registration using deep neural networks: a comprehensive review," *Computers and Electrical Engineering*, vol. 87, p. 106767, 2020, doi: 10.1016/j.compeleceng.2020.106767.

[15] X. Yang, R. Kwitt, M. Styner, and M. Niethammer, "Quicksilver: fast predictive image registration – a deep learning approach," *NeuroImage*, vol. 158, pp. 378–396, 2017, doi: 10.1016/j.neuroimage.2017.07.008.

[16] F. Zhu, M. Ding, and X. Zhang, "Self-similarity inspired local descriptor for non-rigid multi-modal image registration," *Information Sciences*, vol. 372, pp. 16–31, 2016, doi: 10.1016/j.ins.2016.08.031.

[17] M. K. Tripathi and Shivendra, "Neutrosophic approach based intelligent system for automatic mango detection," *Multimedia Tools and Applications*, vol. 83, no. 14, pp. 41761–41783, 2024, doi: 10.1007/s11042-023-17037-7.

[18] B. D. de Vos, F. F. Berendsen, M. A. Viergever, H. Sokooti, M. Staring, and I. Išgum, "A deep learning framework for unsupervised affine and deformable image registration," *Medical Image Analysis*, vol. 52, pp. 128–143, 2019, doi: 10.1016/j.media.2018.11.010.

[19] S. Dasariraju, M. Huo, and S. McCalla, "Detection and classification of immature leukocytes for diagnosis of acute myeloid leukemia using random forest algorithm," *Bioengineering*, vol. 7, no. 4, pp. 1–12, 2020, doi: 10.3390/bioengineering7040120.

[20] V. Lavanya P, M. K. Tripathi, E. P. Hemand, S. K, and J. V. N. Ramesh, "Hyperspectral crop image classification via ensemble of classification model with optimal training," *Web Intelligence*, pp. 1–31, 2024, doi: 10.3233/web-230209.

[21] A. T. Salawudeen, Mu'azu, M.B., Y. A. Sha'aban, and E. A. Adedokun, "On the development of a novel smell agent optimization (SAO) for optimization problems," 2018.

[22] G. Dhiman, M. Garg, A. Nagar, V. Kumar, and M. Dehghani, "A novel algorithm for global optimization: rat swarm optimizer," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 8, pp. 8457–8482, 2021, doi: 10.1007/s12652-020-02580-0.

[23] M. K. Tripathi and Shivendra, "Improved deep belief network for estimating mango quality indices and grading: a computer vision-based neutrosophic approach," *Network: Computation in Neural Systems*, vol. 35, no. 3, pp. 249–277, 2024, doi: 10.1080/0954898X.2023.2299851.

[24] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354–377, 2018, doi: 10.1016/j.patcog.2017.10.013.

[25] S. Albawi, O. Bayat, S. Al-Azawi, and O. N. Ucan, "Social touch gesture recognition using convolutional neural network," *Computational Intelligence and Neuroscience*, vol. 2018, pp. 1–10, 2018, doi: 10.1155/2018/6973103.

[26] F. Zöllner, "Multimodal ground truth datasets for abdominal medical image registration is taken from," *heiDATA*, 2022. https://heidata.uni-heidelberg.de/dataset.xhtml?persistentId=doi:10.11588/data/ICSFUS.

BIOGRAPHIES OF AUTHORS

Dr. Mukesh Kumar Tripathi received a Ph.D. in Computer Science and Engineering from VTU, Belagavi. He has 15 years of teaching and administrative experience. He has supervised and co-supervised more than five masters and 20 B.E. students. He is working as an assistant professor with the Department of Computer Science and Engineering, Vardhaman College of Engineering, Hyderabad, India. He is authored or co-authored more than twenty publications and more than 300 citations. His research interests include soft computing, machine learning, intelligent systems, image processing, and hyperspectral. He can be contacted at email: mukeshtripathi016@gmail.com.

Dr. Chellapilla V. K. N. S. N. Moorthy is working as Director RD, Vasavi College of Engineering, Hyderabad, Telangana, India. He received Master of Technology both in the fields of Computer Science Engineering and Heat Power Refrigeration and Air Conditioning. He received doctoral degree for research from GITAM University, and pursuing his doctoral degree in the field of machine learning too. He has research experience with a total research grant of 424.46 K USD from DST, government of India, more than 50 research publications, international research collaborations, awards and patents to his credit. His thrust areas of research include cognitive science, AI and ML, and computational fluid dynamics. He can be contacted at email: krishna.turbo@gmail.com.

Dr. Sandeep Kadam working as principal of Keystone School of Engineering, Pune, India. He is having 20+ years in academics and industry. His research area is cyber security and software engineering, include soft computing, machine learning, intelligent systems, image processing, and hyperspectral. He is Ph.D. research Co-Supervisor in Computer Engineering. He is authored or co-authored more than twenty-five publications and more than 100 citations. His research interests. He can be contacted at email: sandipkadam@gmail.com.

Prof. Chaitali Shewale is working as Assistant Professor in Information Technology Department at Vishwakarma Institute of Information Technology, Pune, India. She has completed her Masters in Computer Engineering and pursuing Ph.D. in Computer science and engineering. Her research area in machine learning, data science, cyber security and blockchain. She has published more 30 research articles in various national and international journals of repute. She is having more than 12 years of teaching experience in the field of information technology. She can be contacted at email: chaitali.shewale@viit.ac.in.

Dr. Priya Shelke holds a Doctor of Philosophy degree in Computer engineering from Savitribai Phule Pune University, Pune, India. She has received her master's (M.Tech. Computer Science and Engineering) degree from Visvesvaraya Technological University, Belgavi in 2009. She is working as an Associate Professor in Information Technology Department of Vishwakarma Institute of Information Technology, Kondhwa(bk), Pune. She is having 20 years of working experience. Her area of interest is image processing and blockchain technology. She has published over 40 papers in national and international conferences and journals. She is a life time member of ISTE. She can be contacted at email: priya.shelke@viit.ac.in.

Dr. Pravin R. Futane is currently associated with Vishwakarma Institute of Information Technology (VIIT), Pune as a Professor in Computer Engineering and Head, IT Department from last three years. Earlier have worked as a Professor in Computer Engineering at PCCoE for 3.2 years and Head of IT department for almost 2 years and was associated as Professor and Head of Department of Computer Engineering, Sinhgad College of Engineering, Pune. He is having total 24+ years of experience with 23 years of teaching experience including 15 years of administrative experience and 10 years of research experience parallelly. He can be contacted at email: pravin.futane@viit.ac.in.