Novel Surveillance System for Suspicious Activities Analysis using Deep Learning

Aditya Kakde, Bhavana Kaushik, Deepika Koundal, Durgansh Sharma, and Neelu Jyothi Ahuja

Abstract Surveillance is about tracking suspicious actions rather than recognizing objects in the scene. Current surveillance systems employ classification and detection procedures, but fail to explain why they are used, given that the human eye is capable of seeing them as well. For surveillance to be effective, it must not only report an occurrence but also warn if there is a risk of an incident occurring. As a result, suspicious actions must be investigated by comparing current data to past data in the form of a picture. Human involvement cannot compare this past data with the most recent data from millions of data points since it would take too long. As a result, we suggest a system that uses an end-to-end system to analyse some questionable actions in real-time. To accurately detect any suspicious actions using sensor data, the proposed surveillance system employs Content-based Image Retrieval (CBIR) combined with a deep learning algorithm. Using real-time graphical analysis and feature extraction also allows for improved administration of results and data. The suggested system completed CBIR with deep learning and demonstrated its graphical analysis and feature extraction in real-time, demonstrating its uniqueness over previous systems. The suggested system includes a dashboard that can be used to analyse not only what happened at that specific time, but also what happened in the previous days and how they differed from what happened today. With this suspicious analysis method, one can not only determine whether or not

Aditya Kakde
University of Petroleum and Energy Studies, Dehradun, India e-mail: adityakakde100@gmail.com

Bhavana Kaushik
University of Petroleum and Energy Studies, Dehradun, India e-mail: bkaushik@ddn.upes.ac.in

Deepika Koundal
University of Petroleum and Energy Studies, Dehradun, India e-mail: Deepika Koundal@ddn.upes.ac.in

Durgansh Sharma
Christ (Deemed to be University), Ghaziabad, India e-mail: durgansh.sharma@christuniversity.in

Neelu Jyothi Ahuja
University of Petroleum and Energy Studies, Dehradun, India e-mail: neelu@ddn.upes.ac.in

an incident has occurred, but also receive a warning if there is a possibility of an incident.

Key words: Content Based Image Retrieval, Computer Vision, Convolution Neural Network (CNN), Real-time Feature Extraction, Graphical Analysis, Satellite Imagery, Surveillance System

1 Introduction

Remote sensing has a fascinating and rich history. It all began during World War I when aerial photography proved to be an effective instrument for exploration and observation. The pre-Hispanic Civilization of Nazca started distant observation long back. 'Earth observation' as a source of traditional indicators to create geoglyphs (known as Nazca Lines) that could only be seen from a certain height was used in ancient days. However, in today's world, this type of monitoring is still carried out in the form of satellite photographs that are taken from satellites to witness the earth's motion and suspicious events. It is always crucial to observe the ground movements for identifying suspicious activity for example construction of new buildings, storms, wildfires and rising sea levels. From farmers analysing their crops to urban architects properly charting roadways, satellite pictures have a wide range of uses. Increased sea levels, hurricanes, and wildfires all can be detected with the help of satellites that monitor the environment. Radar satellites are being used by geologists to anticipate volcanic eruptions and locate fault lines. Military satellites, which are generally used for surveillance and investigation by intelligence specialists, and satellites that are generally used for communication and entertainment are termed commercial satellites, GPS satellites are used for direction-finding applications, and scientific satellites are used for weather studies, planetary research, and assessing agricultural patterns. The current research proposes a technique for detecting suspicious behaviour using satellite pictures (that are openly available on Google). Surveillance is said to be perfect and efficient when it detects any anomalous or distrustful activities precisely. Current surveillance systems are generally operated by humans, and they involve incessant human attention to detect any unusual/suspicious activity. As humans are intervening, the efficiency of the system decreases with time due to the exhaustion and tiredness factor of humans. The above challenge can be resolved by the automation of the surveillance system. The purpose of the automated mechanism is to give a warning in the form of an alarm or any other method when a predefined abnormal action happens [1]. Satellite imagery is fetching more researchers and scholar community for various challenges and issues, and it has been used for provincial-level mapping, location planning, and defencelessness or destruction valuations in the latest events. In non-emergency situations, satellite imagery has been used to estimate population estimates [2]. Currently, some 5,300 satellites orbit Earth, which means that thousands of cameras are taking real-time photographs above you. Satellite photography provides a unique perspective for photographing the planet, which can aid scientists and others in identifying patterns and trends [3]. Furthermore, to make educated guesses about unusual and unusual behaviour. As part of the country's coastal security fortification, Indian Space Research Organisation (ISRO) satellite imageries will quickly monitor distrustful vessels and boats heading into the seas, according to the home ministry [4]. Using satellite photography to correctly detect changes to the Earth's surface may help with everything from climate change studies and farming to human migration patterns and nuclear non-proliferation. However, dynamically integrating photos from a variety of sources — for example, those that indicate surface changes (such as new building development) against those that show substance changes (such as water to sand) was unfeasible [5]. With a fresh suggested model capability, what's suspicious happening on the ground may now be easily spotted in terms of a new building and military movements for security purposes using satellite photos. The suggested system employs a deep Convolutional Neural Network to perform Content-Based Image Retrieval (CBIR). CBIR uses query pictures to search a huge database for images, and content denotes the shape, colour, texture, and other associated information of both the query and stored photos. To conduct Content-Based Image Retrieval, the proposed system uses a deep Convolutional Neural Network (CBIR). CBIR searches a large database for photographs using input images, where content refers to various characteristics of an image like shape, colour and texture connected with both the input and stored photos [6]. In [7], the characteristics of satellite photographs maintained in the record are mined using a VGG19 convolution neural network model. VGG19 was chosen above other extremely deep learning models due to the systems specification constraint. It's a 19-layer and 5-pooling layers CNN with both the convolutional and fully connected layers using ReLU activation functions, and the output layer using softmax activation functions to acquire critical characteristics to recognise alike images from kept images based on an input image. The query photographs' characteristics are compared to stored database picture features during the testing phase, and similarity is calculated. The similarity is used to find the most comparable photos.

2 Literature Review

It was proposed in [8] that a deep CNN be utilised with a 4-layer neural model using the ReLU function as activation was employed on the CIFAR-10 data, and it had a six times quicker training error rate. The unique architecture, which consisted of 5 convolutional layers with overlapping pooling and local response normalisation, was then evaluated on the ILSVRC-2010 data, and it obtained the lowest error percentage. And after that, it is put to the test on the ILSVRC-2012, where it attained the lowest error percentage of 67 and 40.9. In [7], the ILSVRC-2012 dataset was used to test a deep learning model. The grouping accuracy improved after increasing the layers of the deep learning model. When pictures were scaled to (256,512), the 19 layers of CNN succeeded with the lowest top-1 testing error of 23.7% and top-5 testing error of 6.8%. A review of CBIR was given in [6], which discussed the sorts of characteristics that are utilised to identify similarities. Among the characteristics are colour, texture, shape, spatial, low-level cues, region-based approaches, and extraction of features using a deep learning model. Using a combination of colour and form data, the deep learning approach was employed in [9] to locate the most comparable photographs to the query image. The images used in the experiment were sourced from the internet. CBIR used a transfer learning technique to recover brain tumours [10]. A unique paradigm was given to aid the radiologist in recognising tumour types in unclear instances. In [11], CBIR was employed as a transfer learning method for fetching trademarks as an image. In [12], the transfer learning approach was used with the CBIR mechanism on biological and remote sensing data, resulting in the conclusion that CNN is more dependable than typical CBIR systems. In [13], a comparative analysis of CNN and CNN-SVM was studied, and it was determined that CNN-SVM should be used. When CNN was combined with GPU, however, it yields better results as compared to CNN-SVM by 0.5 per cent and consumed a reduced amount of time. The article by [14] proposes a detection approach for a ship as an object based on spectral reflectance in challenging settings such as darkness, fog, and hazes using multi-spectral satellite pictures. A neural network dubbed LFNet (lightweight fusion network) was also developed to validate regions with ships by combining ship reflectance and picture colour information. In addition, the proposed research shows that they can detect adverse weather. Continuously moving satellite images and videos are utilised for a lot of monitoring and tracking. A work published in [15] suggests that micro vehicle recognition be done on satellite videos utilising a multi-morphological-cue-based discriminating algorithm to isolate the vehicle from the background noise. The effectiveness of maritime surveillance systems has improved in recent years, and one suggested system, which utilises a Haar-like classifier for boat identification, is provided in [16]. They also proposed an upgrade in [17], which was based on detection and integrated data from diverse sources to get improved results. The results, which include images shot in a variety of lighting conditions and with a variety of camera settings, highlight the technique's utility. In the [18], routine satellite remote sensing surveillance on oil spills using SAR photographs is carried out, which assists in determining the percentage of oil spillage by monitoring satellite data for the previous five years in the area of the Bohai Sea and north of the Yellow Sea. As observed in [19], one of the study projects in the field of surveillance focuses on marine traffic understanding utilising multi-sensor satellite data processing, which recommends the programmed identification of the vessel's movement-related characteristics as well as vessel velocity vector calculation. The gradient-based method improves the accuracy of the estimate of wake motion-related characteristics. In [20], issues such as concerns and difficulties, distrustful movement identification of humans, the methodology for suspicious activity identification, datasets, with assessment methods were discussed, as well as a comparison of numerous surveillance systems. In [21], to detect violent actions in real-time, a crowd detection system based on convolutional long short-term is applied. In [22], a new deep YOLOv3 detection technique with 106 convolutional layers and 2 dense connected layers has been developed. The surveillance system is based on a detecting system and is utilised with a small drone. In [23], using Quantized SSD Mobilenet V2 and Tiny YOLOv3 models, a surveillance system based on the detection and categorization of military vehicles was presented, with the Tiny YOLOv3 model providing the best performance. In [24], for surveillance using UAVs, an ideal deep learning algorithm called Optimal UAV-based Layer Distribution (OULD) and OULD with Mobility Prediction (OULD-MP) is utilised to decrease latency during data classification. In [25], a monitoring and surveillance system is proposed that uses a deep learning model on aerial photos, with the best performance coming from YOLOv4 twice. A detection mechanism underpins the suggested model. However, none of them discusses real-time feature extraction or real-time graphical analysis of confidence ratings. The classification portion was also discussed, but not the analysis of historical data and how it might be compared to current data. Furthermore, no logic has been provided as to why a bounding box is required over an object that can be seen with one's own eyes. Some discussed detection in multi-spectral images, however, it was proposed as video surveillance, for which this study gave reasons in Section 3, why it isn't very trustworthy. The tabular analysis of the literature review can be seen in table 1.

3 Proposed Work

To conduct Content-Based Image Retrieval, the proposed system employs an algorithm with pictures as a query. The fact is emphasised that video is not essential for the identification of suspicious behaviour because video monitoring may produce challenges such as object tracking of multiple objects, shadow identification, fuzzy objects, clutter, and gatherings [20]. The convolutional layer examines the whole image and fetches the critical characteristics and features of an image. The characteristics of pictures in the database are trained using the VGG19 deep learning model. Throughout the validation phase, the details of the input pictures are compared with the characteristics of the testing images, and likeness is calculated between both sets of images. The similarity is used to find the most comparable photos. In addition, the prototype can extract characteristics in real-time. It uses the scheduler principle, which involves creating a folder in which a scheduler may be started. First, a time is assigned to a scheduler. When someone uploads a photo to that folder, the scheduler checks to see if it was uploaded, and if it was, the training will begin immediately. The size of the filter, strides, and padding values are other convolutional layer characteristics. The padding value is set to 1, which preserves the convolutional layers' and strides' spatial resolution as shown in Figure 1. Using the receptive field, the convolutional layer analyses the pictures part by part, forming a 3 x 3 matrix. Because s = 1, the value of stride in the given picture is 1. This indicates that another matrix will have a change of one column and one row which is described in Figure 2. The process of convolution involves the dot product of part of an image with the selected filters to give a single value after summing up. As a consequence, we receive the extracted feature. This method is known as convolu-

Table 1 Tabular analysis of the literature review

Author	Title	Dataset	Methods
A. Latif et. al [6]	Content-based Image Retrieval and Feature Extraction: A Comprehensive Review	Not Applicable	CBIR
K. Simonyan and A. Zisserman [7]	Very Deep Convolutional Network for Large Scale Image Recognition	ILSVRC	VGG
A. Krizhevsky et. al [8]	ImageNet Classification with Deep Convolution Neural Network	ILSVRC	AlexNet
R. Rajkumar and M.V. Sudhamani [9]	Content based Image Retrieval System using Combination of Color and Shape Features and Siamese Neural Network	Data scrapped from world wide web (www)	CBIR
Z. Swati et. al [10]	Content-Based Brain Tumor Retrieval for MR Images Using Transfer Learning	CE-MRI	VGG19-based novel feature extraction framework
S.Hasan et. al [11]	Trademark Image Retrieval using Transfer Learning	FlickerLogos-32 and Logos-32	TIR system using AlexNet
P. Sadeghi-Tehran [12]	Scalable Database Indexing and Fast Image Retrieval Based on Deep Learning and Hierarchically Nested Structure Applied to Remote Sensing and Plant Biology	MalayaKew and UCM	CBIR with FE-CNN, BOVW and MFF
O. Mohamed et. al [13]	Content-Based Image Retrieval Using Convolutional	ImageNet and Caltech256	CNN-SVM
X.Xie et. al [14]	Ship Detection in Multispectral Satellite Images Under Complex Environment	Multi-spectral images of ships from 4 satellites	LFNet
W. Ao et. al [15]	Needles in a Haystack: Tracking City-Scale Moving Vehicles from Continuously Moving Satellite	Satellite videos	Novel algorithm based on local noise modelling
D. Bloisi et. al [16]	Automatic Maritime Surveillance with Visual Target Detection	AIS	Haar-Cascade
D. Bloisi et. al [17]	Enhancing Automatic Maritime Surveillance Systems with Visual Information	EO and IR data	Visual detection, visual tracking, and data fusion
L.Bing et. al [18]	Spatial Distribution Characteristics of Oil Spills in the Bohai Sea Based on Satellite Remote Sensing and GIS	SAR images	Framework based on remote sensing and geographical information system (GIS)
M. Reggiannini and L. Bedini [19]	Multi-Sensor Satellite Data Processing for Marine Traffic Understanding	Remote sensing data of ships	Tailored gradient estimator in the early processing stages
R. K. Tripathi et. al [20]	Suspicious human activity recognition: a review	Not Applicable	Surveillance
T.Saba [21]	Real time anomalies detection in crowd using convolutional long short-term memory network	Standard crowd anomaly dataset	Conv-LSTM
K. Madasamy et. al [22]	OSDDY: embedded system-based object surveillance detection system with small drone using deep YOLO	Data of open field and marine environment from drones	YOLOv3
P. Gupta et. al [23]	Edge device based Military Vehicle Detection and Classification from UAV	Novel data of military vehicles	Quantize SSD Mobilenet v2 and Tiny YOLOv3
M. Jouhari et. al [24]	Distributed CNN Inference on Resource-Constrained UAVs for Surveillance Systems: Design and Optimization	Data collected from UAVs	LeNet and VGG166
H. Gupta and O. Verma [25]	Monitoring and surveillance of urban road traffic using low altitude drone images: a deep learning approach	AU-AIR	YOLOv4

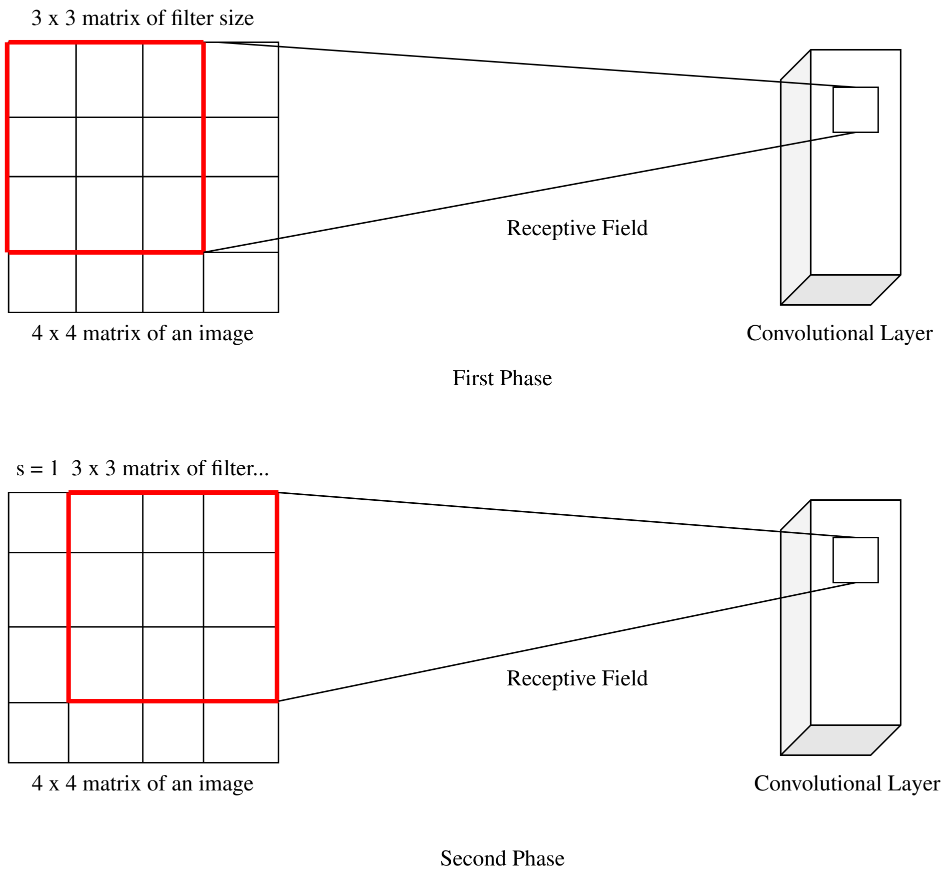

Fig. 1 Process of Feature Extraction

tion. The image's dimension is steadily reduced during the process, leaving just the most significant elements which can be seen through Eq. 1.

$$\frac{n_h + 2_p - f}{s} + 1 \times \frac{n_h + 2_p - f}{s} + 1 \tag{1}$$

where nh stands for the height of the image and nw is the width of an input image, padding is denoted by p, f stands for a dimension of the kernel, and s stands for stride. To reduce the input image's size pooling mechanism is used which is described in Figure 3.

In the pooling process, max pooling is applied where 4 cells get retrieved from the 2 x 2 matrix and assess the cell with the greatest value which can be seen in Eq. 2.

$$max(Cell1, Cell2, Cell3, Cell4) \tag{2}$$

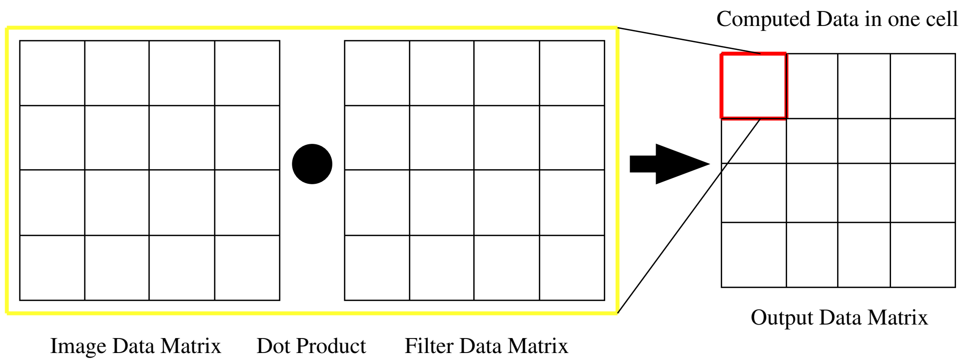

Fig. 2 Process of Convolution

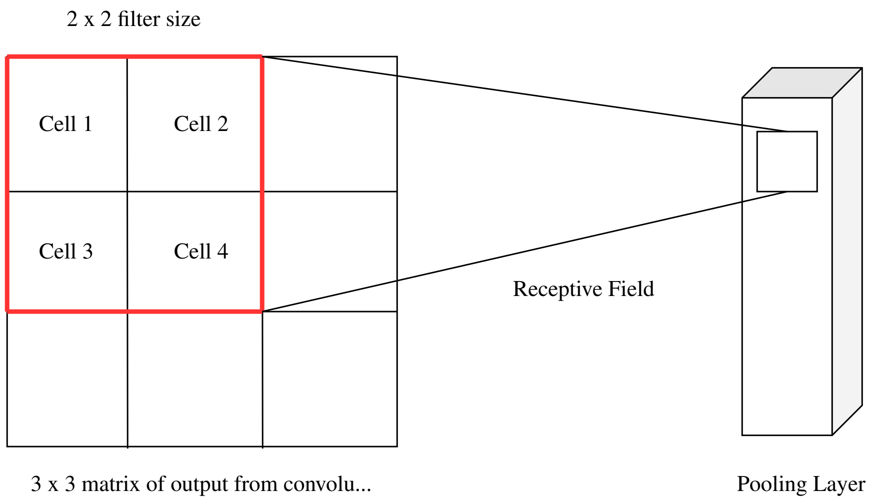

Fig. 3 Process of Pooling

In this phase, the image is likewise down-scaled to capture the most important parts. The following formula seen in Eq. 3 may be used to compute it:

$$\frac{n_h - f}{s} + 1 \times \frac{n_h - f}{s} + 1 \tag{3}$$

where nh stands for the height of the image and nw is the width of an input image, f is the dimension of the kernel and s = Strides (should be a completely divisible integer) and it is shown in Figure 4.

The dot product, also known as convolution, will be used to produce the original picture data and filter data first, as illustrated in the flowchart above. The data is subsequently transferred to the pooling layer, which takes into consideration the largest number value from the resulting matrix. The pictures are not trained in the

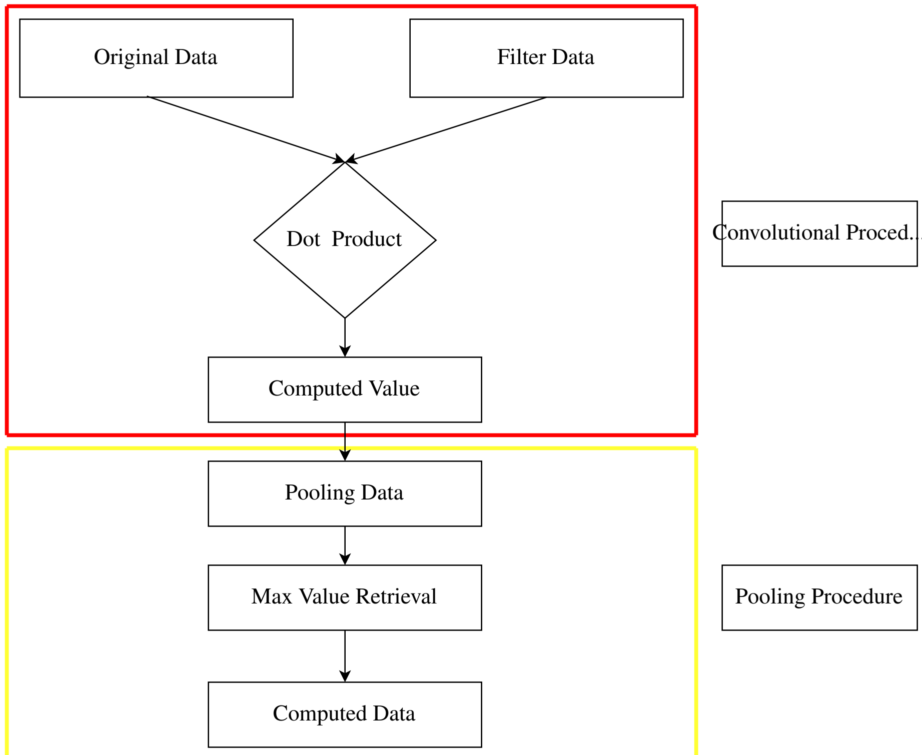

Fig. 4 Flowchart of the feature extraction process.

suggested method. When we train a neural network, we enable it to learn data, but in this situation, all we do is acquire the features of the photos so that we may compute their similarity. After using this approach for fetching the features on both database and query photographs, the dot product of their feature vectors will be determined which can be seen in Eq. 4.

$$\sum_{i=1}^{n} q_i . \sum_{i=1}^{n} d_i .T \tag{4}$$

where q is the feature set of the input picture, d is the feature set of the stored picture, and T is indicating the transposition of the feature set. Then, because we need to order the most comparable photos, a quick approach is utilized.

$$(N-1) + \frac{2}{N} \sum_{k=0}^{N-1} Q_k \tag{5}$$

where N = array size and k ∈ 0,...,n-1. To organise an array in ascending order as given in Eq. 5, we reverse this procedure to make a decreasing ordered list to obtain the most comparable picture.

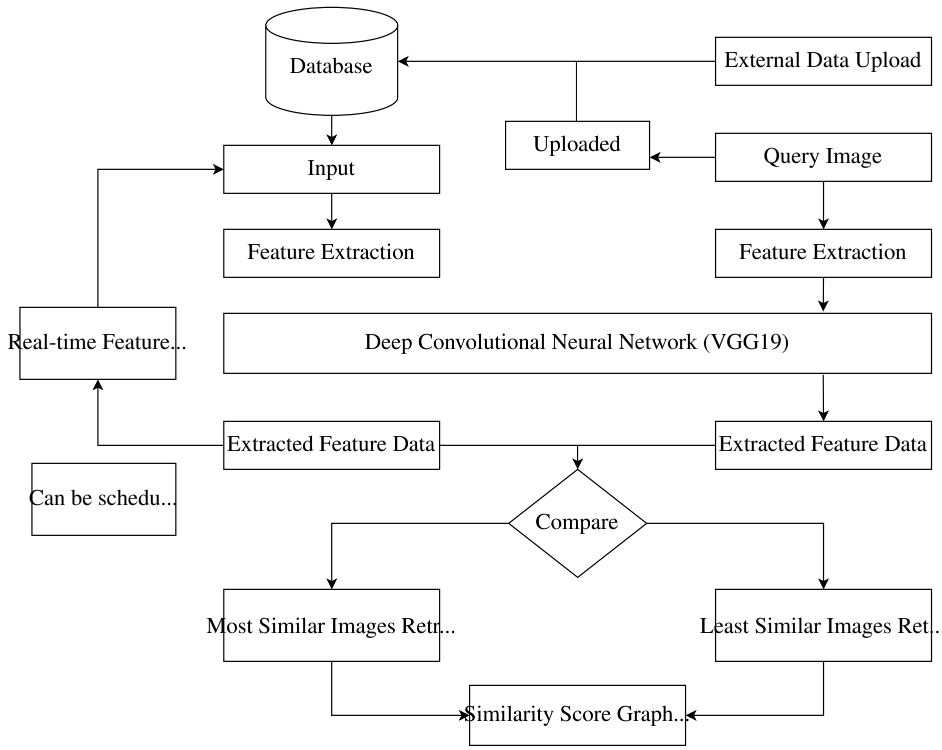

Fig. 5 Flowchart of proposed surveillance systems.

4 Experimental Setup

The uniqueness of the suggested monitoring mechanism, which is fully described in Fig. 5 and the experimental setup are discussed in this section The information is initially uploaded to the dashboard and then stored in the record. Images are sent into the deep learning CNN model to extract features of all stored images. These characteristics include colour, texture, patterns, and so on, unlike typical CBIR, where only a single kind of feature is extracted for matching purposes. The feature vector is fetched with the same CNN model when we submit a query or a test picture to the database. The stored images and input data's likeness are then calculated. If a match is discovered, the most similar image is returned, and if a large number of photographs are required, the match is returned in decreasing order. In terms of real-time feature extraction, the training phase is planned daily at noon (though this may be altered), and it constantly scans the given photographs and initiates training at an infinite range.

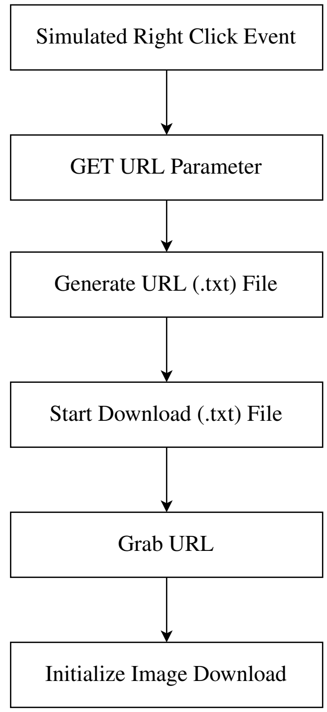

Fig. 6 Steps of fetching the images from Google.

5 Results and Discussions

This section discusses about the dataset used, work output generated, and the analysis of the outcome.

5.1 Dataset

This module discusses the dataset as well as the desired outcome. The photographs are acquired from the internet because the collection of satellite photos of the site is not available. The approach is depicted in the flowchart in Fig. 6. This process simulates a right-click activity for receiving the URL of each picture from the context menu without navigating to another page. The second phase includes the Get URL method to get a URL parameter from a query string since Google keeps the whole picture URL in a query parameter. After that creation of the text file is done and downloaded. The next step includes fetching all URLs once all are collected. We utilise the request parameter to download the.txt file once we have all of the URLs in it. For the proposed system, 530 satellite images were used to generate characteristics.

5.2 Work Output

For surveillance, CBIR with deep learning is used to analyse not only what happened at that exact instant, but also what happened in the prior days and how they differ from what happened today. Assume we've gotten a significant number (perhaps millions) of satellite photos depicting a group of militants nearing the border. However, rather than detecting anything, we want to use changes in a specific region's activity as a monitoring tool for analysing changes. We'll need the most comparable photos from the prior actions to compare with the most current ones. It will take much too long to locate the most comparable photos if someone manually compares a query image to millions of other photographs in the folder. By taking input, a dashboard has been given to prevent manual interruptions.

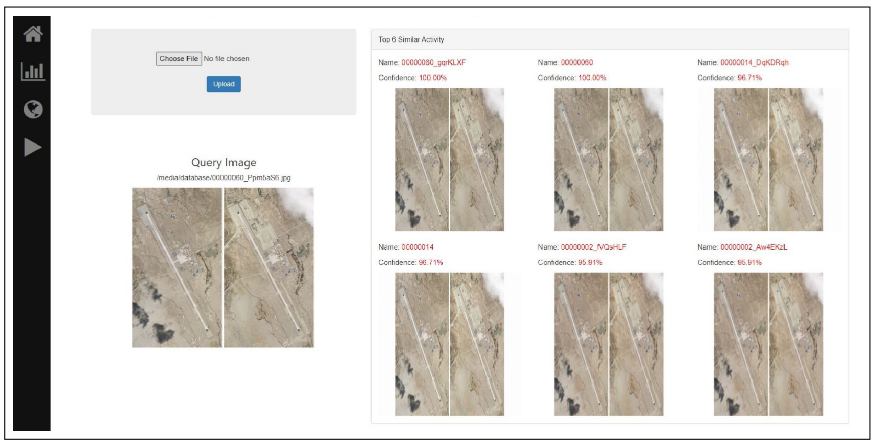

Fig. 7 The dashboard's main screen.

In Figure 8, a satellite picture is given as input, and accordingly received the utmost comparable images linked with that event. The six most similar images were gathered to show the prototype, with the first being highly matched, the second being bit lesser alike than the first, and so on. This dashboard makes use of tabs to allow for easy navigation from one section to the next and, as a consequence, removes the need to refresh the web page in Fig. 7.

A graphical breakdown of the outcome may be shown at the same time in Fig. 8. After that, you'll find the about section and real-time feature extraction. It is possible to see the number and names of images that have been lately posted in Fig. 9. If you submit the same image more than once, it will be re-titled and will be taken as a new picture. After all these the images are stored in a database and further utilized to automate the scheduler. When the scheduled interval arrives, it searches for recently stored images before starting the training process, which eliminates the need for human participation.

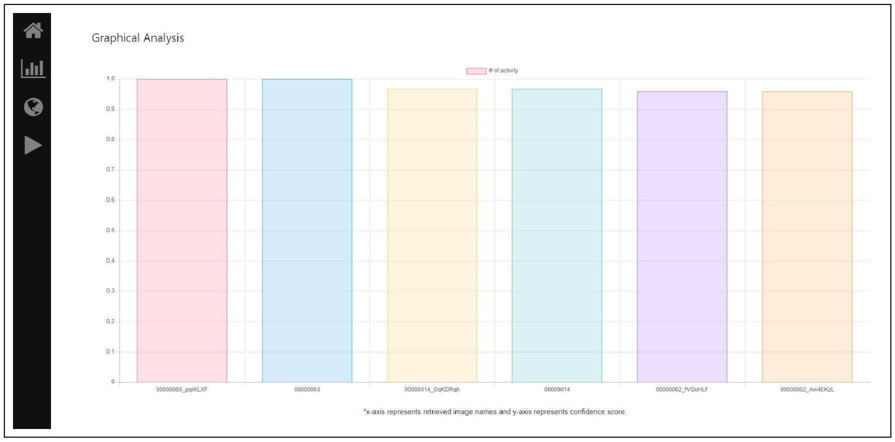

Fig. 8 Real-time graphical analysis of test results in real-time.

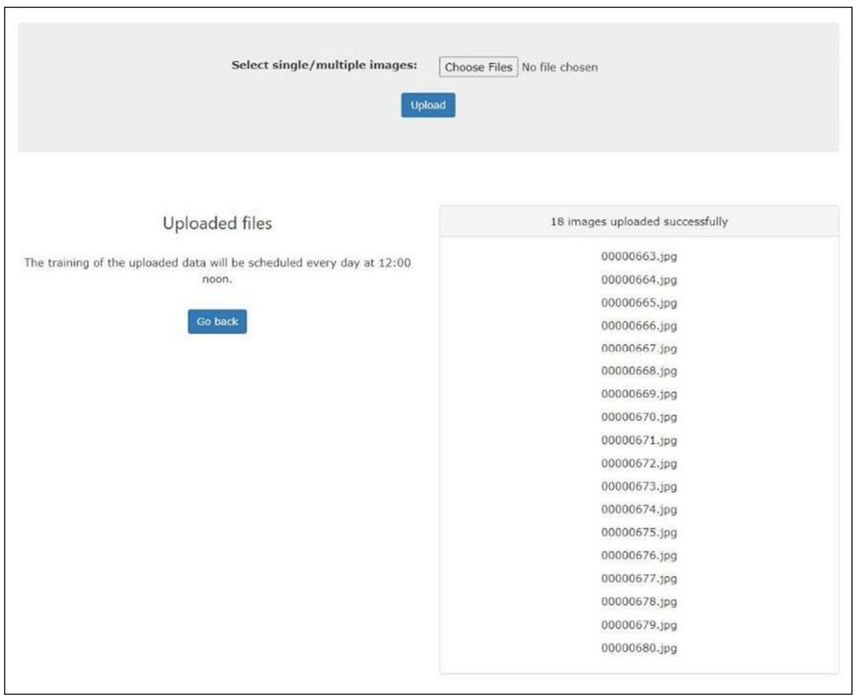

Fig. 9 Real-time feature extraction.

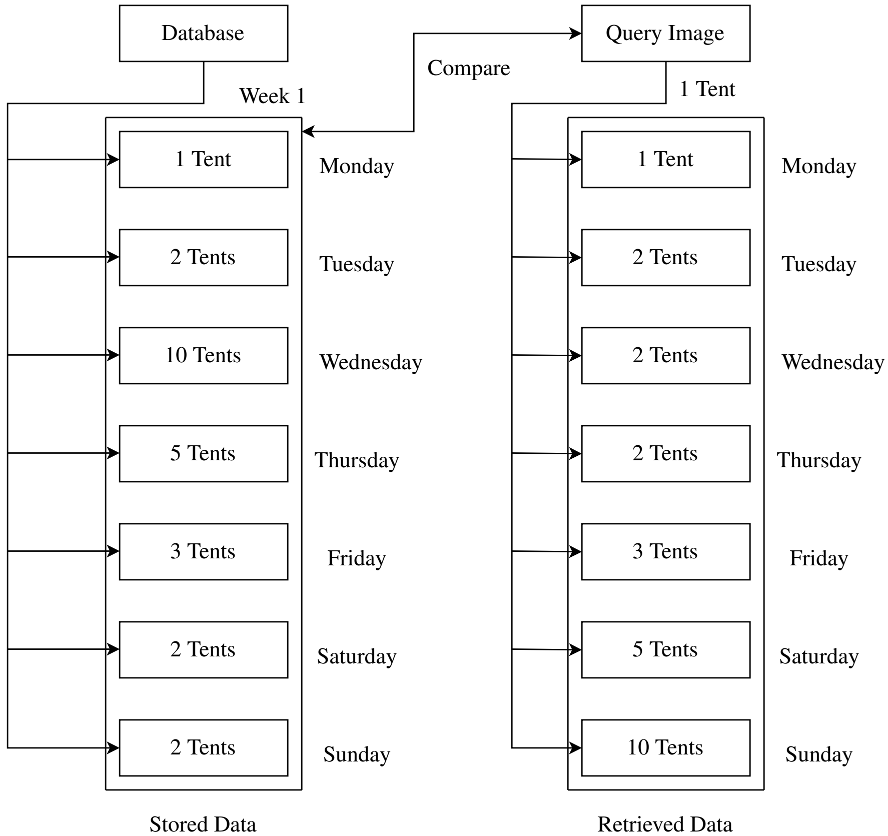

Fig. 10 Example analysis of suspicious activity.

5.3 Analysis

This part will show you how to get the analysed result from the returned images acquired with CBIR and deep learning. A circumstance in which n number of tents are placed in an area for a week has been presented as an example. This example assumes that the subject is well-understood. Fig. 10 shows how input or query pictures can be utilized to recognise doubtful activity on earth faster and more efficiently. The labels in Fig. 10 are only there to help you understand what is happening to check the suspicious activity. Week 1 data may be displayed, with activities such as 1 Tent, 2 Tents, and so on, with Monday through Sunday indicating the day those activities occurred. The results are matched to the given test image from the first day of week 2. Assume the query image has the activity 1 Tent. It is then compared to the database image, which retrieves the most similar images, which can now be analysed to identify the variation in movement, which reveals that tents increased in the early days of the past week and steadily lessened by the completion of the week. The database visualisations in Fig. 10 are only illustrative. It won't simply be a few photographs in the database; it might be hundreds of thousands. As a result, analysing and making conclusions only based on data is impossible. In Table 2 all the recent work published in area surveillance is utilized to identify the techniques they are using to compare with the proposed method.

Table 2 Comparative analysis of the proposed surveillance system with other state-of-the-art systems

Author	CBIR	DL	Real-time GA	Real-time FE
T.Saba [21]	✗	✓	✗	✗
K. Madasamy et. al [22]	✗	✓	✗	✗
P. Gupta et. al [23]	✗	✓	✗	✗
M. Jouhari et. al [24]	✗	✓	✗	✗
H. Gupta and O. Verma [25]	✗	✓	✗	✗
Proposed Work	✓	✓	✓	✓

The components considered are CBIR, DL, real-time GA, and real-time FE, where DL stands for deep learning, GA for graphical analysis, and FE for feature extraction. A tick indicates that a system has used these parameters, whereas a cross indicates that it has not.

6 Future Work

The suggested system introduces new methods for carrying out surveillance activities, although there is always an opportunity for creativity. As a result, further research may be done on:

1. Enhancement of the feature extraction process of CBIR.
2. Enhancement of the deep neural network architecture.
3. Using detection in CBIR with deep learning when dealing with multi-spectral images.

7 Conclusion

Content-Based Image Retrieval (CBIR) combining deep learning techniques with real-time feature extraction and real-time graphical analysis has never been used in a surveillance system. As a result, the proposed system employs CBIR and deep learning algorithms to identify suspicious behaviour in real-time while extracting data characteristics. It may be used to compare and contrast acts that occurred not only at that exact time, but also in prior days, and how they differ from what is being done now. The recommended answer is in the shape of a dashboard, which, thanks to its swift navigation capabilities, allows all of these actions to be accomplished quickly. This suspicious analysis technique not only determines whether or not an event has occurred, as categorization systems do but also provides a warning if an occurrence is likely to occur.

References

1. S. Mubareka, D. Ehrlich, B. Ferdinand and F. Kayitakire, "(Settlement location and population density estimation in rugged terrain using information derived from Landsat ETM and SRTM data," International Journal of Remote Sensing, vol. 29, no. 8, pp. 2339-2357, 2008.
2. f. S. B. T. Checchi and J. Palmer, "Validity and feasibility of a satellite imagery-based method for rapid estimation of displaced populations," International Journal of Health Geography, vol. 12, no. 4, pp. 1-12, 2013.
3. S. Brown, "Satellite surveillance may be less of a privacy concern than you think – for now," 29 October 2019. [Online].
Available: https://www.cnet.com/science/turns-out-satellite-surveillance-only-sounds-like-a-major-privacy-concern/ [Accessed 26 May 2022]
4. R. Venkatesan, "ISRO satellite imageries to monitor suspicious vessels in Indian waters," 8 January 2018. [Online].
Available: https://www.thehindubusinessline.com/news/isro-satellite-imageries-to-monitor-suspicious-vessels-in-indian-waters/article9972309.ece [Accessed 21 May 2022]
5. A. Ziemann, "Integrating diverse satellite images sharpens our picture of activity on Earth," 21 December 2021. [Online].
Available: https://www.lanl.gov/discover/science-columns/top-columns-and-blogs/2021/space-satellite-images.php [Accessed 25 May 2022].
6. A. Latif, A. Rasheed, U. Sajid, J. Ahmed, N. Ali, N. I. Ratyal, B. Zafar, S. H. Dar, M. Sajid and T. Khalil, "Content-Based Image Retrieval and Feature Extraction: A Comprehensive Review," Mathematical Problems in Engineering, pp. 1-21, 2019.
7. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arxiv 14091556, 2014.
8. A. Krizhevsky, l. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolution Neural Network," Communication of ACM, vol. 60, no. 6, pp. 84-90, 2017.
9. R. Rajkumar and M. V. Sudhamani, "Content based Image Retrieval System using Combination of Color and Shape Features and Siamese Neural Network," International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 9, no. 2, pp. 71-77, 2019.
10. Z. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed and J. Lu, "Content-Based Brain Tumor Retrieval for MR Images Using Transfer Learning," IEEE Access, vol. 7, pp. 17809 - 17822, 2019.
11. S. J. Hassen, A. Taha and M. M. Selim, "Trademark Image Retrieval using Transfer Learning," Journal of Engineering and Applied Sciences, vol. 14, no. 18, pp. 6897-6905, 2019.
12. P. Sadeghi-Tehran, P. Angelov, N. Virlet and M. Hawkesford, "Scalable Database Indexing and Fast Image Retrieval Based on Deep Learning and Hierarchically Nested Structure Applied to Remote Sensing and Plant Biology," Journal of Imaging, vol. 5, no. 3, pp. 33-54, 2019.
13. O. Mohamed, E. A. Khalid, O. Mohammed and A. Brahim, "Content-Based Image Retrieval Using Convolutional Neural Networks," Lecture Notes in Real-Time Intelligent Systems, 2019.
14. X. Xie, B. Li and X. Wei, "Ship Detection in Multispectral Satellite Images Under Complex Environment," Remote Sensing, vol. 12, no. 5, p. 792, 2020.

15. W. Ao, Y. Fu, X. Hou and F. Xu, "Needles in a Haystack: Tracking City-Scale Moving Vehicles from Continuously Moving Satellite," IEEE Transactions on Image Processing, 2019.
16. D. Bloisi, L. Iocchi, M. Fiorini and G. Graziano, "Automatic Maritime Surveillance with Visual Target Detection," Computer Science Journal, 2011.
17. D. D. Bloisi, F. Previtali, A. Pennisi, D. Nardi and M. Fiorini, "Enhancing Automatic Maritime Surveillance Systems with Visual Information," IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 4, pp. 1-10, 2017.
18. L. X. Q. L. X. and Z. N. (. S. D. C. o. O. S. i. t. B. S. B. o. S. R. S. a. G. J. o. C. R. Bing, "Spatial Distribution Characteristics of Oil Spills in the Bohai Sea Based on Satellite Remote Sensing and GIS," Journal of Coastal Research, vol. 90, pp. 164-170, 2019.
19. M. Reggiannini and L. Bedini, "Multi-Sensor Satellite Data Processing for Marine Traffic Understanding," Electronics, vol. 8, no. 2, pp. 152-170, 2019.
20. R. K. Tripathi, A. S. Jalal and S. C. Agrawal, "Suspicious human activity recognition: a review," Artificial Intelligence Review, vol. 50, no. 2, pp. 283-339, 2018.
21. T. Saba, "Real time anomalies detection in crowd using convolutional long short-term memory network," Journal of Information Science, 2021.
22. K. Madasamy, V. Shanmuganathan, V. Kandasamy, M. Y. Lee and M. Thangadurai, "OSDDY: embedded system-based object surveillance detection system with small drone using deep YOLO," EURASIP Journal on Image and Video Processing, vol. 19, no. 1, 2021.
23. P. Gupta, B. Pareek, G. Singal and D. V. Rao, "Edge device based Military Vehicle Detection and Classification from UAV," Multimedia Tools and Applications, vol. 81, no. 14, p. 19813–19834, 2022.
24. M. Jouhari, A. Al-Ali, E. Baccour, A. Mohamed, A. Erbad, M. Guizani and M. Hamdi, "Distributed CNN Inference on Resource-Constrained UAVs for Surveillance Systems: Design and Optimization," IEEE Internet of Things Journal, pp. 1-16, 2022.
25. H. Gupta and O. Verma, "Monitoring and surveillance of urban road traffic using low altitude drone images: a deep learning approach," Multimedia Tools and Application, vol. 81, pp. 19683-19703, 2022.