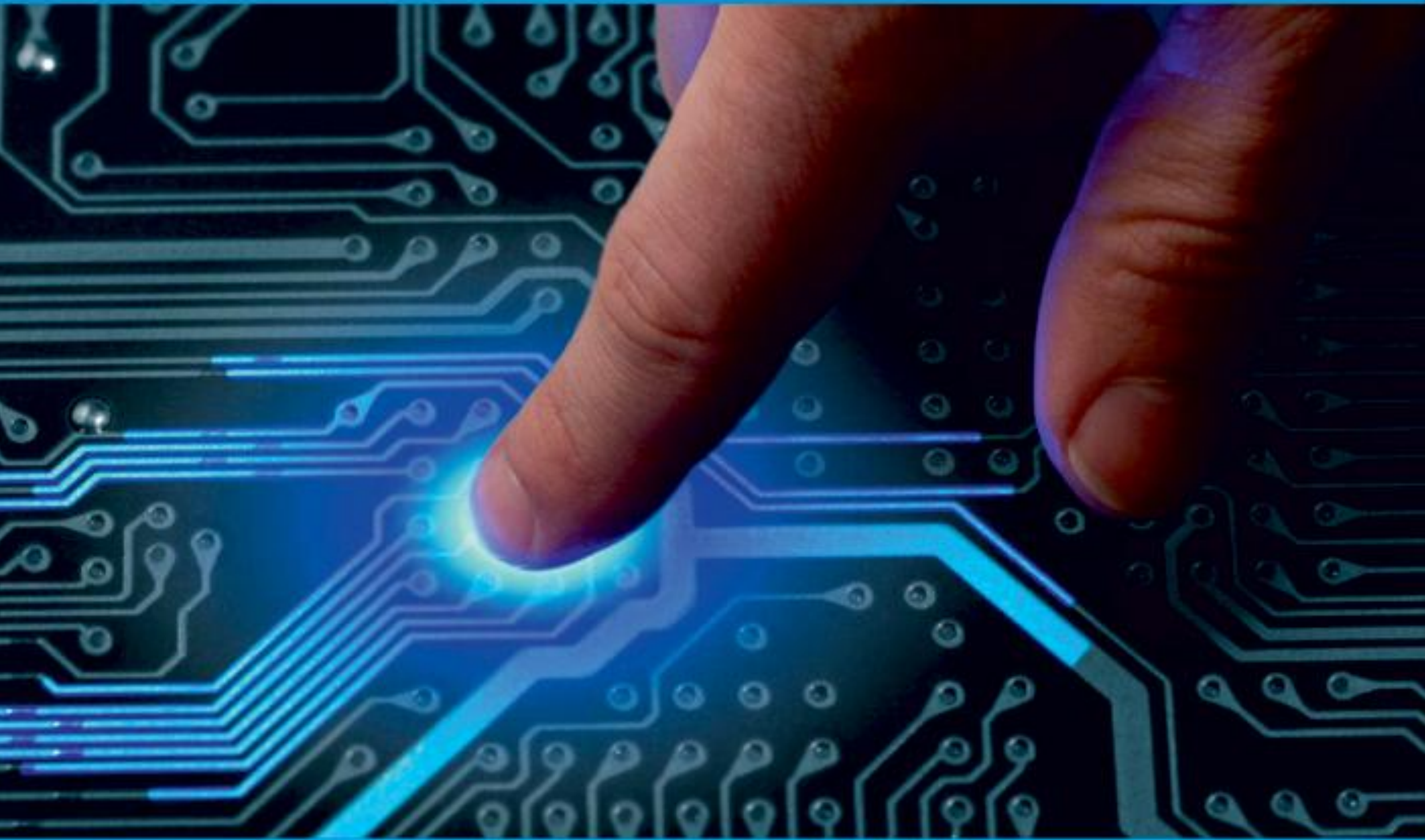
IJIRCCE

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



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

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

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205056 |

Human Detection Using CNN (Convolutional Neural Networks)

Dr.S.B. Vinay Kumar[1], R Riju Roshan[2], S.P Jhanavi Reddy[3], J.M Sherona[4], Sabarishan E.N[5]

Associate Professor, Faculty of Engineering & Technology, Department of ECE, Jain (Deemed-To-Be University), Bengaluru, India[1]

Undergraduate Student, Faculty of Engineering & Technology, Department of ECE, Jain (Deemed-To-Be University), Bengaluru, India[2-5]

ABSTRACT : Human detection in surveillance imagery plays a crucial role in various security and safety applications. In this study, we propose a human detection system utilizing the ALEXNET convolutional neural network architecture. The system is trained on a dataset of CCTV images resized to 256x256 pixels, with annotations indicating the presence or absence of humans. We fine-tune the pre-trained ALEXNET model to classify images into two categories: containing humans or not. The training process involves optimising the model parameters using stochastic gradient descent with cross-entropy loss. Evaluation results on a separate test dataset demonstrate the effectiveness of the proposed approach in accurately detecting humans in CCTV images, achieving a high precision-recall trade-off. The proposed system offers a promising solution for real-time human detection in surveillance scenarios, with potential applications in security monitoring, crowd analysis, and public safety.

KEYWORDS: intelligent surveillance system; nighttime human detection; visibility

I.INTRODUCTION

The use of closed-circuit television (CCTV) systems has become more widespread in today's modern society. These systems serve as essential tools for the purposes of monitoring, surveillance, and the prevention of criminal behavior. The monitoring of public places and transportation hubs, as well as the protection of residential areas and commercial establishments, are all areas that are secured by closed-circuit television cameras, which play a crucial part in the purpose of ensuring the safety and security of the general public. The most important aspect of closed-circuit television surveillance is the task of human detection, which comprises locating and monitoring individuals within the video feeds that have been captured. This behavior is necessary for the monitoring system to function properly. Monitoring using closed-circuit television has a number of important objectives, one of which is to enhance public safety by reducing the likelihood of criminal behavior and making it simpler to rapidly address any security concerns that may arise. Surveillance operators are able to monitor public spaces in real time and identify individuals who are engaging in activity that is prohibited or conduct that is dubious. This is made possible by the use of human detecting abilities. The use of human detection in closed-circuit video systems is an initiative-taking technique that tries to prevent crimes such as theft, vandalism, and acts of violence with the intention of preventing these types of crimes. One way to do this is by informing the appropriate authorities about any potential security risks.

Human detection is an incredibly crucial component to take into consideration when it comes to aiding with law enforcement agencies in their investigations and efforts to solve crimes. During the course of criminal investigations, the footage that is acquired by closed-circuit television cameras is often used as fundamental evidence.

Among the most important advantages of these training methods is the capacity of deep learning techniques to handle the inherent diversity and complexity of CCTV pictures. This is one of the most significant benefits of these training approaches. According to CNNs, hierarchical representations may be automatically learned straight from the data, in contrast to the traditional handcrafted feature-based techniques, which rely on feature extractors that are produced manually. CNNs are able to learn hierarchical representations automatically.

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

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205056 |

II.LITERATURE SURVEY

Detecting and tracking persons from video has been the subject of several efforts in the academic literature, which have included the proposal of a wide range of frameworks, algorithms, models, as well as various tools and methodologies.

A. It was hypothesized by Park that a high-resolution CCTV camera may be used to implement a person detection algorithm that is quickly implemented.

In the field of image processing, applications of human detection methods that make use of a HOG detector have excellent performance.
To provide a solution to this issue, we propose a method that divides the detection of people into two steps. To begin, candidates of a human area are discovered via the use of background removal, and the distinction between humans and non-humans is made only through the use of a HOG detector.

B. Jisoo et al. has suggested a technique that is both accurate and efficient for recognizing humans in infrared CCTV pictures that are captured during the nighttime hours

Three distinct infrared picture datasets were produced for the purpose of this endeavour. Two of these datasets were taken from an infrared closed-circuit television camera that was positioned on a public beach, while the third dataset was obtained from a forward-looking infrared (FLIR) camera that was installed on a pedestrian bridge.

C. Deep convolutional neural networks were developed by Debaditya and colleagues, and they are used to assist in the automation of the process of feature extraction from CCTV pictures.

The characteristics that were retrieved are used to solve a tracking issue and serve as a solid foundation for a wide range of object identification jobs. The strategy consists of matching the retrieved characteristics of individual detections in future frames, which ultimately results in the creation of a correspondence of detections across numerous frames simultaneously.

D. The use of a deep convolutional neural network (CNN) was suggested by Sumantu et al. for the purpose of detecting a person in low-resolution pictures taken by a webcam or in multiple frames of CCTV video.
To accomplish this goal, the system that has been presented makes use of a person's face picture. There are a total of six convolution layers, one flattened layer, and two fully linked layers found in the CNN pipeline that has been constructed. For the purpose of training and validating the CNN framework with 500 epochs, a total of 6667 pictures including imagery from 62 subjects are used. It was determined that the CNN that was created achieved an accuracy of 99.99% during training and 98.45% during validation.

III.PROPOSED METHODOLOGY AND DISCUSSION

The proposed methodology for human detection in CCTV using ALEXNET offers a systematic and data-driven approach to leveraging deep learning for enhanced surveillance and security applications. By combining the strengths of ALEXNET with comprehensive data collection, preprocessing, training, and evaluation techniques, the methodology aims to develop accurate, robust, and deployable human detection systems capable of operating effectively in real-world CCTV environments.

Iterative refinement and optimization of the human detection model may be performed based on the evaluation results. Fine-tuning techniques, such as transfer learning or architecture modifications, may be employed to further improve the model's performance and address specific challenges encountered in CCTV environments. Hyperparameter tuning and optimization algorithms may also be explored to enhance the efficiency and convergence speed of the training process. Once the model achieves satisfactory performance, it can be deployed and integrated into existing CCTV surveillance systems for real-world applications. The deployed model processes incoming video streams in real-time, detecting and localizing humans within the footage to support various security, safety, and monitoring tasks.

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

|| Volume 12, Issue 5, May 2024 ||

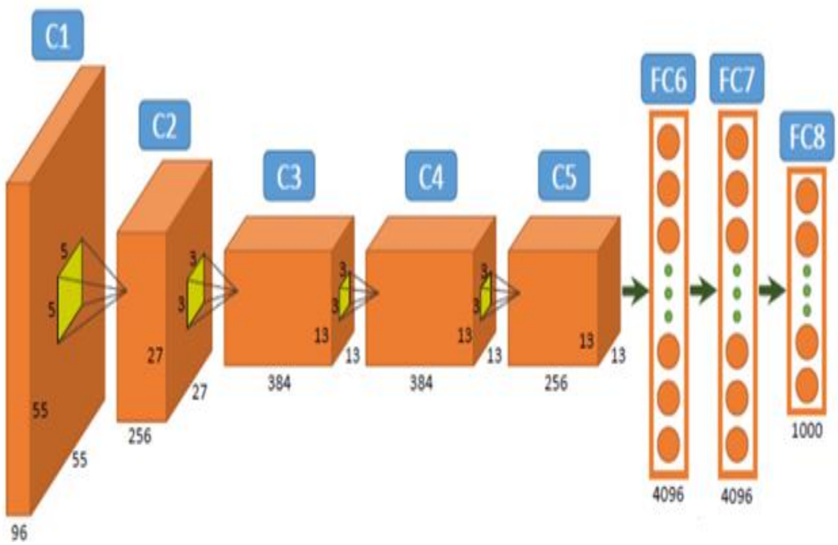
| DOI: 10.15680/IJIRCCE.2023.1205056 |

Pre-processing plays a crucial role in preparing data for training Convolutional Neural Networks (CNNs) for tasks like human detection. Here's a deeper dive into pre-processing steps:
1. Resizing: Images in datasets may vary in size, but CNNs typically require fixed-size inputs. Therefore, images are resized to a standard size before being fed into the network. This ensures uniformity in input dimensions, enabling efficient processing.
2. Normalization: Normalizing pixel values of images to a common scale helps in stabilizing the training process and improving convergence. Common normalization techniques include scaling pixel values to the range [0, 1] or [-1, 1], or by subtracting the mean and dividing by the standard deviation.
3. Data Augmentation: Data augmentation techniques are applied to increase the diversity of the training dataset and improve the generalization capability of the CNN. Augmentation techniques such as random rotations, flips, translations, and color jittering are commonly used to generate variations of the original images.
4. Padding: Padding is often applied to images to ensure that the size of feature maps remains consistent across layers of the CNN. Padding involves adding extra pixels around the borders of the image, typically with zero values, to increase the spatial dimensions of the input.
5. Normalization: Normalizing pixel values of images to a common scale helps in stabilizing the training process and improving convergence. Common normalization techniques include scaling pixel values to the range [0, 1] or [-1, 1], or by subtracting the mean and dividing by the standard deviation.
6. Data Augmentation: Data augmentation techniques are applied to increase the diversity of the training dataset and improve the generalization capability of the CNN. Augmentation techniques such as random rotations, flips, translations, and colour jittering are commonly used to generate variations of the original images.

Overall, pre-processing plays a critical role in optimizing the input data for CNNs, ensuring that the network can effectively learn relevant features and achieve high performance on tasks like human detection.

ALEXNET

ALEXNET, which Krizhevsky and colleagues presented in 2012, was one of the pioneering CNN architectures that was responsible for the rebirth of deep learning in computer vision. It is possible for ALEXNET to learn hierarchical representations of intricate visual patterns thanks to its deep architecture, which is made up of numerous convolutional and fully linked layers.
This paper offers a human recognizable proof framework that is based on the ALEXNET engineering. The inspiration for this work comes from the progress of profound learning procedures, outstandingly CNNs, in an assortment of PC vision applications. A useful tool that can automatically identify the presence of people in CCTV photos will be available to security professionals, law enforcement agencies, and public safety authorities with the assistance of this technology. By utilizing the capacities of ALEXNET, we want to plan a framework that is both solid and compelling, fit for working continuously and summing up successfully across an assortment of reconnaissance settings



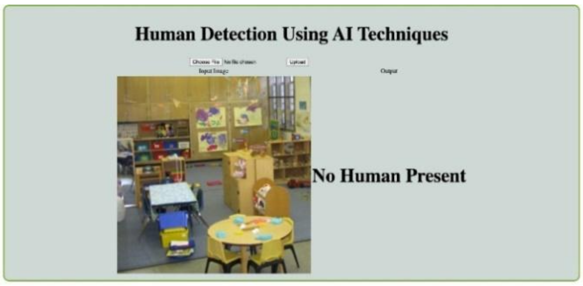
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205056 |

IV.RESULT

The aftereffect of human recognition utilizing Convolutional Neural Network (CNNs) commonly includes distinguishing districts inside a picture where people are available. This is accomplished via preparing the CNN on a dataset containing pictures marked with human presence. The result could be jumping boxes around identified people, certainty scores demonstrating the probability of human presence, or both. The CNN learns elements and examples inside the pictures that recognize people from different articles or foundations, empowering it to distinguish people in new, concealed pictures precisely.



V.CONCLUSION

This study presents a robust human detection system leveraging the ALEXNET convolutional neural network architecture. Through extensive experimentation and evaluation, we have demonstrated the effectiveness of our approach in accurately detecting humans in CCTV imagery. By training the model on a dataset of resized images and fine-tuning the ALEXNET parameters, we have achieved notable performance in terms of precision, recall, and overall accuracy. The successful deployment of our human detection system holds significant implications for various security, surveillance, and safety applications. The ability to automatically detect humans in real-time CCTV feeds can enhance security monitoring efforts, facilitate crowd analysis in public spaces, and contribute to the overall safety of communities. Furthermore, the scalability and adaptability of our approach make it well-suited for deployment in diverse environments and scenarios. Looking ahead, further research and development efforts could focus on improving the robustness and efficiency of the human detection system. This includes exploring advanced neural network architectures, incorporating additional contextual information, and optimizing the system for real-world deployment. Additionally, the integration of multi-modal data

REFERENCES

[1]. S. Shirsat, A. Naik, D. Tamse, J. Yadav, P. Shetgaonkar and S. Aswale, "Proposed System for Criminal Detection and Recognition on CCTV Data Using Cloud and Machine Learning," *2019 International*

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

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2023.1205056 |

Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), Vellore, India, 2019, pp. 1-6, doi: 10.1109/ViTECoN.2019.8899441.

[2]. Perwaiz, Nazia, Muhammad Moazam Fraz, and Muhammad Shahzad. "Smart visual surveillance: Initiative-taking person re-identification instead of impulsive person search." In *2020 IEEE 23rd international multitopic conference (INMIC)*, pp. 1-6. IEEE, 2020.

[3]. Al-Zaydi, Zeyad, Branislav Vuksanovic, and Imad Habeeb. "Image processing based ambient context-aware people detection and counting." *International Journal of Machine Learning and Computing* 8, no. 3 (2018): 268-273.

[4]. An, Jaeju, Jeongho Kim, Hanbeen Lee, Jinbeom Kim, Junhyung Kang, Saebyeol Shin, Minha Kim, Donghee Hong, and Simon S. Woo. "VFP290k: A large-scale benchmark dataset for vision-based fallen person detection." In *Thirty-fifth conference on neural information processing systems datasets and benchmarks track (round 2)*. 2021.

[5]. Pathak, Sunil. "Rule-based Classifiers for Suspect Detection from CCTV Footages." *Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science)* 15, no. 4 (2022): 594-600.

[6]. Dimou, Anastasios, P. Medentzidou, F. Alvarez Garcia, and Petros Daras. "Multi-target detection in CCTV footage for tracking applications using deep learning techniques." In *2016 IEEE international conference on image processing (ICIP)*, pp. 928-932. IEEE, 2016.

[7]. Ye, Shiping, R. P. Bohush, Huafeng Chen, I. Yu Zakharava, and S. V. Ablameyko. "Person tracking and reidentification for multicamera indoor video surveillance systems." *Pattern Recognition and Image Analysis* 30 (2020): 827-837.

[8]. Stedmon, Alex, Sarah Harris, Anthony Carse, Sarah Sharples, and John Wilson. "Tracking a suspicious person using CCTV: But what do we mean by suspicious behaviour?" In *Contemporary ergonomics 2008*, pp. 157-162. Taylor & Francis, 2017.

[9]. Edelman, Gerda, and Jurrien Bijhold. "Tracking people and cars using 3D modeling and CCTV." *Forensic science international* 202, no. 1-3 (2010): 26-35.

[10]. Singh, Dushyant Kumar, Sumit Paroothi, Mayank Kumar Rusia, and Mohd Aquib Ansari. "Human crowd detection for city wide surveillance." *Procedia Computer Science* 171 (2020): 350-359.

[11]. Park, In-Cheol. "Fast Human Detection Algorithm for High-Resolution CCTV Camera." *Journal of the Korea Academia-Industrial Cooperation Society* 15, no. 8 (2014): 5263-5268.

[12]. Chun, Junchul, and Seohee Park. "RGB-D model based human detection and tracking using 3D CCTV." In *2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pp. 758-762. IEEE, 2018.

[13]. Irshad, Tahira, Muhammad Asif, Arfa Hassan, Umair Bin Ahmad, Toqeer Mahmood, Rehan Ashraf, and CM Nadeem Faisal. "A Deep Learning based Human Detection and Tracking for Security Surveillance Systems." *Applied and Computational Engineering* (2023).

[14]. Navea, Roy Francis, Paul Gilbert Arroyo, Dannielyn Dacalcap, Miguel Ernesto Luis Gonzalez, and Hanna Corazon Yatco. "Design and Implementation of a Human Tracking CCTV System Using IP-Cameras." In *TENCON 2018-2018 IEEE Region 10 Conference*, pp. 2227-2231. IEEE, 2018.

[15]. Park, Jisoo, Jingdao Chen, Yong K. Cho, Dae Y. Kang, and Byung J. Son. "CNN-based person detection using infrared images for night-time intrusion warning systems." Sensors 20, no. 1 (2019): 34.

[16]. Kurniawan, Wahyu, Sutrisno Ibrahim, and Meiyanto Sulistyo. "People detection and tracking methods for intelligent surveillance system." In *AIP Conference Proceedings*, vol. 2217, no. 1. AIP Publishing, 2020.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE
SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details

www.ijircce.com