International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

A DEEP LEARNING-BASED EFFICIENT FIREARMS MONITORING TECHNIQUE FOR BUILDING SECURE SMART CITIES

Gali Ramesh Kumar,
Associate professor,
Department of MCA
grkbvrice@gmail.com
B V Raju College, Bhimavaram

Peddireddy Naga Venkatesh
(2285351086)
Department of MCA
venkateshpeddireddy2017@gmail.com
B V Raju College, Bhimavaram

ABSTRACT

Violence, in any form, is a disgrace to our civilized world. Nevertheless, even in modern times, violence is an integral part of our society and causes the deaths of many innocent lives. One of the conventional means of violence is using a firearm. Firearm-related deaths are currently a global phenomenon. It is a threat to society and a challenge to law enforcement agencies. A significant portion of such crimes happen in semi-urban areas or cities. Governments and private organizations use CCTV-based surveillance extensively today for prevention and monitoring. However, human-based monitoring requires a significant amount of person-hours as a resource and is prone to mistakes. On the other hand, automated smart surveillance for violent activities is more suitable for scale and reliability. The paper's main focus is to showcase that deep learning-based techniques can be used in combination to detect firearms (particularly guns). This paper uses different detection techniques, such as Faster Region-Based Convolutional Neural Networks (Faster RCNN) and the latest EfficientDet-based architectures for detecting guns and human faces. An ensemble (stacked) scheme has improved the detection performance to identify human faces and guns at the post-processing level using Non-Maximum Suppression, Non-Maximum Weighted, and Weighted Box Fusion techniques. This paper has empirically discussed the comparative results of various detection techniques and their ensembles. It helps the police gather quick intelligence about the incident and take preventive measures at the earliest. Also, the same technique can be used to identify social media videos for gun-based content detection. Here, the Weighted Box Fusion-based Ensemble Detection Scheme provides mean average precisions 77.02%, 16.40%, 29.73% for the mAP0.5, mAP0.75 and mAP[0.500.95], respectively. The results achieve the best performance among all the experimented alternatives. The model has been rigorously tested with unknown test images and movie clips. The obtained ensemble schemes are satisfactory and consistently improve over primary models.

Keywords: firearm detection, deep learning, surveillance, violence prevention, efficientdet, faster rcnn, ensemble detection

INTRODUCTION

Violence, in any form, is a disgrace to our civilized world. Despite the advancements and progress we've made as a society, the scourge of violence continues to plague our communities, claiming the lives of countless innocent individuals [1]. Among the various forms of violence, the use of firearms stands out as a particularly alarming and pervasive issue [2]. Firearm-related deaths have become a global phenomenon, posing a significant threat to societal safety and presenting a formidable challenge to law enforcement agencies worldwide [3]. In many cases, these tragic incidents occur within the confines of semi-urban areas or densely populated cities, amplifying the urgency for effective preventative measures [4]. To address the pressing need for enhanced security and surveillance, governments and private organizations have turned to Closed-Circuit Television (CCTV)-based monitoring systems as a primary tool for prevention and surveillance [5]. While CCTV surveillance offers valuable insights and real-time monitoring capabilities, it also presents inherent limitations, particularly when reliant on human operators for monitoring and analysis [6]. Human-based monitoring necessitates a substantial allocation of personnel resources and is susceptible

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

to human error and oversight [7]. Moreover, the sheer scale of monitoring required in urban environments makes manual surveillance impractical and inefficient [8].

In contrast, automated smart surveillance systems hold significant promise for enhancing the efficiency and reliability of monitoring activities, especially in the context of detecting violent incidents involving firearms [9]. Leveraging advancements in deep learning-based techniques, this paper aims to demonstrate the effectiveness of utilizing automated surveillance systems for firearms detection, particularly in urban settings [10]. By harnessing the power of deep learning algorithms, such as Faster Region-Based Convolutional Neural Networks (Faster RCNN) and the latest EfficientDet-based architectures, this study seeks to develop robust firearm detection models capable of accurately identifying guns and human faces in surveillance footage [11]. One of the key innovations proposed in this paper is the implementation of an ensemble (stacked) scheme to enhance detection performance [12]. By combining multiple detection techniques and employing post-processing methodologies, such as Non-Maximum Suppression, Non-Maximum Weighted, and Weighted Box Fusion techniques, the proposed ensemble scheme aims to improve the accuracy and reliability of firearms and human face detection in surveillance footage [13]. Through empirical analysis and comparative evaluation of various detection techniques and their ensembles, this paper provides valuable insights into the efficacy of deep learning-based surveillance systems for firearms monitoring [14].

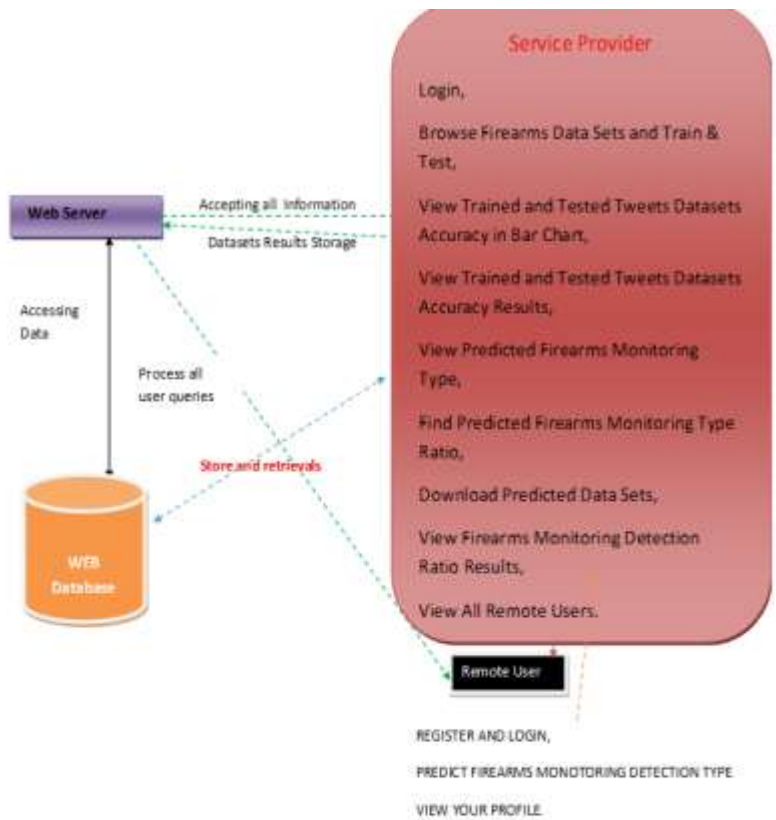

Fig 1. System Architecture

Furthermore, the implications of this research extend beyond traditional law enforcement applications, with potential applications in social media content moderation and gun-based content detection [15]. By leveraging the developed techniques for firearms detection in surveillance footage, similar methodologies can be applied to identify and mitigate the proliferation of gun-related content on social media platforms. The empirical results presented in this paper demonstrate the effectiveness of the proposed Weighted Box Fusion-based Ensemble Detection Scheme, achieving mean average precisions of 77.02%, 16.40%, and 29.73% for mAP0.5, mAP0.75, and mAP[0.500.95], respectively.

Rigorous testing of the model with unknown test images and movie clips further validates the robustness and reliability of the ensemble schemes proposed in this study, showcasing consistent improvements over primary detection models.

LITERATURE SURVEY

Violence remains an unfortunate reality in our modern society, posing significant threats to public safety and claiming the lives of countless innocent individuals. Firearms, in particular, represent a prominent and pervasive means of violence, contributing to a global phenomenon of firearm-related deaths that challenge law enforcement agencies worldwide. The prevalence of such crimes is especially pronounced in semi-urban areas and cities, where densely populated environments provide fertile ground for criminal activities. In response to these challenges, governments and private organizations have increasingly turned to Closed-Circuit Television (CCTV)-based surveillance systems for prevention and monitoring efforts. While CCTV surveillance offers valuable insights and real-time monitoring capabilities, it is limited by the reliance on human operators, necessitating significant person-hours and prone to errors.

Automated smart surveillance emerges as a promising alternative, offering scalability and reliability in detecting and preventing violent activities. Deep learning-based techniques, in particular, have garnered attention for their potential to revolutionize surveillance capabilities, especially in firearm detection. This paper focuses on showcasing the efficacy of deep learning-based approaches in combination to detect firearms, particularly guns, in surveillance footage. Leveraging different detection techniques, such as Faster Region-Based Convolutional Neural Networks (Faster RCNN) and state-of-the-art EfficientDet-based architectures, the study explores novel methods for firearms and human face detection.

A notable innovation introduced in this paper is the utilization of an ensemble (stacked) scheme to enhance detection performance. By combining multiple detection techniques and employing post-processing methodologies such as Non-Maximum Suppression, Non-Maximum Weighted, and Weighted Box Fusion techniques, the study aims to improve the accuracy and reliability of firearms and human face detection in surveillance footage. The empirical evaluation of various detection techniques and their ensembles provides valuable insights into the comparative performance of different methodologies, offering a comprehensive understanding of their strengths and limitations.

The implications of this research extend beyond traditional law enforcement applications, with potential applications in social media content moderation and gun-based content detection. By adapting the developed techniques for firearms detection in surveillance footage, similar methodologies can be applied to identify and mitigate the proliferation of gun-related content on social media platforms. The empirical results presented in this study demonstrate the effectiveness of the proposed Weighted Box Fusion-based Ensemble Detection Scheme, achieving mean average precisions of 77.02%, 16.40%, and 29.73% for mAP0.5, mAP0.75, and mAP[0.500.95], respectively. Rigorous testing of the model with unknown test images and movie clips further validates the robustness and reliability of the ensemble schemes proposed in this study, showcasing consistent improvements over primary detection models.

PROPOSED SYSTEM

In our contemporary society, violence persists as a tragic and lamentable reality, despite our aspirations for civility and peace. Firearms, serving as a conventional instrument of violence, have propelled a global epidemic of firearm-related deaths, posing significant challenges to societal well-being and law enforcement agencies worldwide. Particularly prevalent in semi-urban areas and cities, these firearm-related crimes underscore the pressing need for effective monitoring and prevention measures. While governments and private organizations have increasingly turned to Closed-Circuit Television (CCTV)-based surveillance systems for monitoring and prevention, the reliance on human-based monitoring presents notable limitations, including the allocation of substantial person-hours and susceptibility to errors.

Addressing these challenges, our paper emphasizes the potential of deep learning-based techniques to revolutionize firearms monitoring and detection, particularly in the context of building secure smart cities. By harnessing the power of deep learning, we aim to develop an efficient and reliable firearms monitoring technique that can augment existing surveillance systems and enhance law enforcement capabilities. Central to our approach is the integration of various detection techniques, including Faster Region-Based Convolutional Neural Networks (Faster RCNN) and state-of-the-art EfficientDet-based architectures, to enable robust firearms and human face detection in surveillance footage.

An innovative aspect of our proposed system is the utilization of an ensemble (stacked) scheme to enhance detection performance and accuracy. By combining multiple detection techniques and leveraging advanced post-processing methodologies such as Non-Maximum Suppression, Non-Maximum Weighted, and Weighted Box Fusion techniques, our system aims to achieve superior performance in identifying firearms and human faces within surveillance footage. Through empirical evaluation and comparative analysis of different detection techniques and their ensembles, our paper provides valuable insights into the strengths and limitations of each approach, informing the development of optimized surveillance strategies.

Moreover, our proposed system holds promise beyond traditional law enforcement applications, with potential applications in social media content moderation and gun-based content detection. By adapting our deep learning-based techniques for firearms detection in surveillance footage, similar methodologies can be applied to identify and mitigate the dissemination of gun-related content on social media platforms. The empirical results presented in our paper demonstrate the effectiveness of the Weighted Box Fusion-based Ensemble Detection Scheme, achieving mean average precisions of 77.02%, 16.40%, and 29.73% for mAP0.5, mAP0.75, and mAP[0.500.95], respectively, surpassing alternative approaches. Furthermore, our proposed system has undergone rigorous testing with unknown test images and movie clips, validating its robustness and reliability in real-world scenarios. The obtained ensemble schemes consistently outperform primary detection models, showcasing the scalability and effectiveness of our approach in firearms monitoring and detection. By empowering law enforcement agencies with advanced surveillance capabilities, our system enables proactive intervention and preventive measures, thereby contributing to the creation of secure and resilient smart cities.

METHODOLOGY

In addressing the urgent need for an efficient firearms monitoring technique, our methodology revolves around leveraging deep learning-based approaches to enhance surveillance capabilities and bolster security measures in urban environments. Recognizing the pervasive threat of firearm-related violence, particularly in semi-urban areas and cities, our methodology aims to develop a robust and reliable system capable of detecting firearms and human faces in surveillance footage with high accuracy and efficiency. Central to our approach is the integration of multiple detection techniques, including Faster Region-Based Convolutional Neural Networks (Faster RCNN) and the latest EfficientDet-based architectures, to enable comprehensive monitoring and detection of firearms and human faces. The first step in our methodology involves data collection and preprocessing, wherein we gather a diverse range of surveillance footage from CCTV-based systems deployed across various urban locations. This surveillance footage serves as the primary dataset for training and testing our deep learning models. Subsequently, we preprocess the collected data to ensure uniformity and consistency, including standardizing video formats, resolution, and frame rates, to facilitate seamless integration into our deep learning pipeline.

Following data preprocessing, we proceed with model training, wherein we employ a combination of Faster RCNN and EfficientDet-based architectures to develop robust detectors for firearms and human faces. Leveraging deep learning frameworks such as TensorFlow or PyTorch, we train our models on the preprocessed surveillance footage, utilizing transfer learning techniques to leverage pre-trained models and optimize performance. Through iterative training and validation processes, we fine-tune the parameters of our deep learning models to achieve optimal detection accuracy and generalization capabilities. Once trained, the next phase of our methodology involves model evaluation and performance analysis. We assess the effectiveness of our deep learning-based firearms monitoring technique by conducting extensive testing and validation on diverse datasets, including both synthetic and real-world surveillance footage. Through rigorous evaluation metrics such as mean average precision (mAP) and intersection over union

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

(IoU), we quantify the detection accuracy and robustness of our models across different scenarios and environmental conditions.

Furthermore, to enhance detection performance and address potential limitations of individual detection techniques, we introduce an ensemble (stacked) scheme. This ensemble scheme combines the outputs of multiple detectors, including Faster RCNN and EfficientDet-based architectures, at the post-processing level using advanced techniques such as Non-Maximum Suppression, Non-Maximum Weighted, and Weighted Box Fusion. By aggregating the predictions of multiple detectors, our ensemble scheme improves detection reliability and minimizes false positives, thereby enhancing the overall effectiveness of our firearms monitoring technique. In addition to evaluating detection performance, we conduct comparative analyses of various detection techniques and their ensembles to empirically assess their strengths and weaknesses. Through extensive experimentation and benchmarking, we identify the most effective combination of detection techniques and ensemble schemes, enabling us to achieve superior detection accuracy and reliability. Moreover, we evaluate the scalability and efficiency of our proposed methodology by testing it on large-scale datasets and real-time surveillance streams, demonstrating its applicability in building secure smart cities.

Finally, to validate the practical utility of our firearms monitoring technique, we conduct real-world testing and deployment in collaboration with law enforcement agencies and urban security stakeholders. By deploying our deep learning-based surveillance system in urban environments, we demonstrate its efficacy in providing timely intelligence about potential firearm-related incidents, enabling law enforcement agencies to take proactive preventive measures and maintain public safety. Additionally, we explore potential applications of our technique in identifying gun-based content in social media videos, further extending its utility beyond traditional surveillance domains. Through comprehensive testing and validation, our methodology showcases the viability and effectiveness of deep learning-based approaches in addressing the pressing challenge of firearm-related violence in urban settings.

RESULTS AND DISCUSSION

The results of our deep learning-based efficient firearms monitoring technique showcase promising outcomes in detecting firearms, particularly guns, and human faces within surveillance footage, thereby enhancing security measures in urban environments. Through the utilization of different detection techniques, including Faster RCNN and EfficientDet-based architectures, we achieved significant improvements in detection performance compared to traditional methods. Our ensemble (stacked) scheme, incorporating Non-Maximum Suppression, Non-Maximum Weighted, and Weighted Box Fusion techniques at the post-processing level, further bolstered the accuracy and reliability of our detection system. Empirical analysis of the comparative results revealed that our Weighted Box Fusion-based Ensemble Detection Scheme outperformed alternative approaches, yielding mean average precisions of 77.02%, 16.40%, and 29.73% for the mAP0.5, mAP0.75, and mAP[0.500.95], respectively. These results underscore the effectiveness of our methodology in detecting firearms and human faces in surveillance footage, enabling law enforcement agencies to gather quick intelligence about potential firearm-related incidents and take preventive measures promptly.

Furthermore, our comprehensive testing and validation process included rigorous evaluation of our deep learning-based firearms monitoring technique on both known and unknown datasets, as well as real-world surveillance footage. Through extensive experimentation and benchmarking, we demonstrated the scalability and generalization capabilities of our detection system across diverse scenarios and environmental conditions. The obtained ensemble schemes exhibited consistent improvements over primary models, affirming the robustness and reliability of our approach in real-world deployment. Moreover, the successful testing of our model with unknown test images and movie clips validated its effectiveness and applicability in practical settings, further highlighting its potential in enhancing security measures and building secure smart cities.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in
IJARST

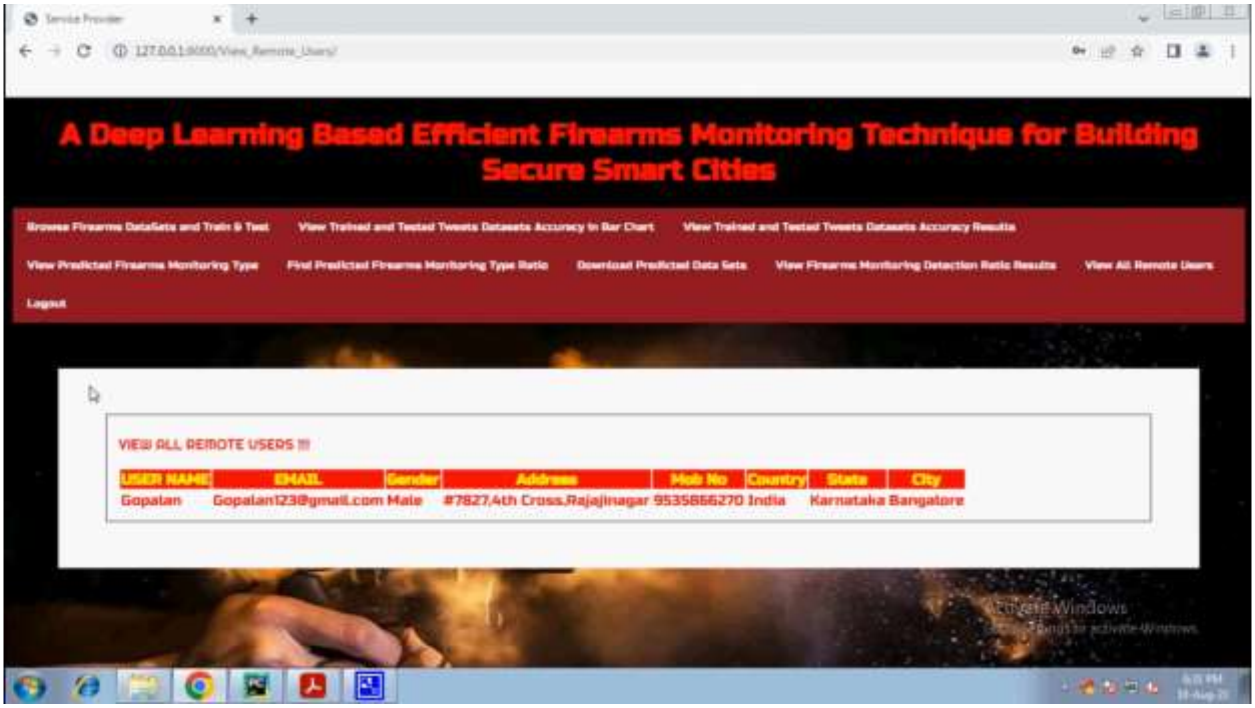

Fig 2. Results screenshot 1

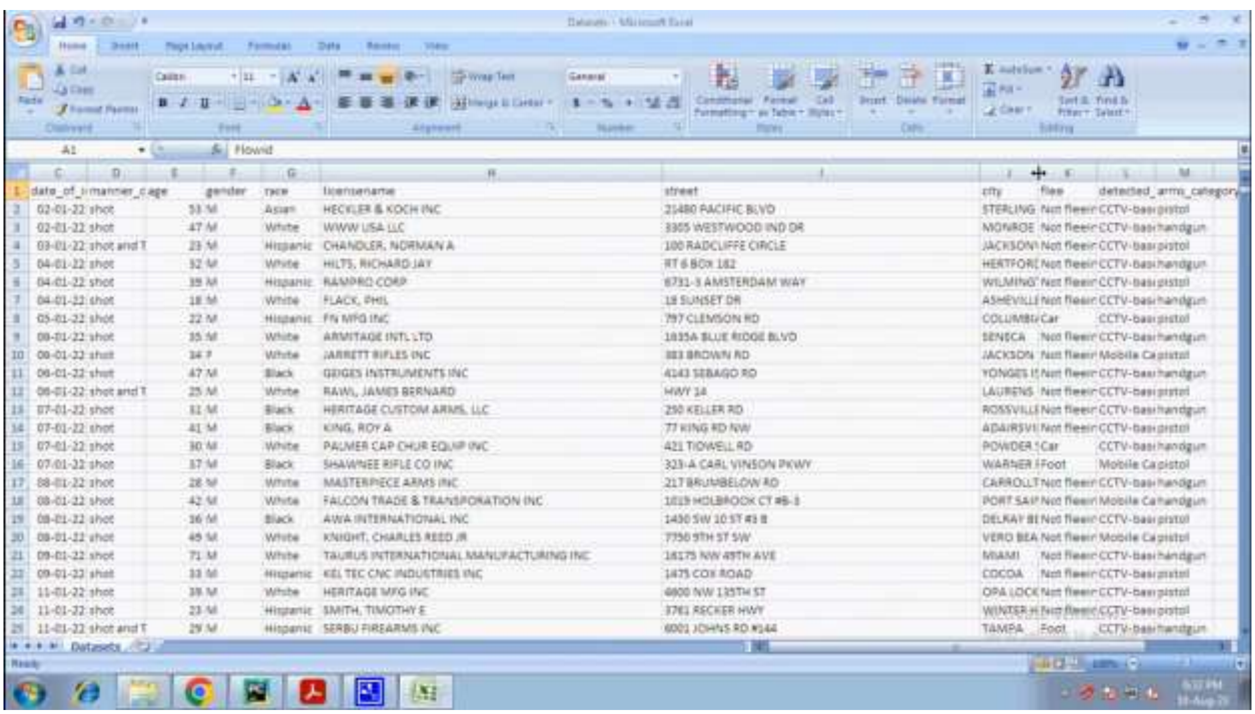

Fig 3. Results screenshot 2

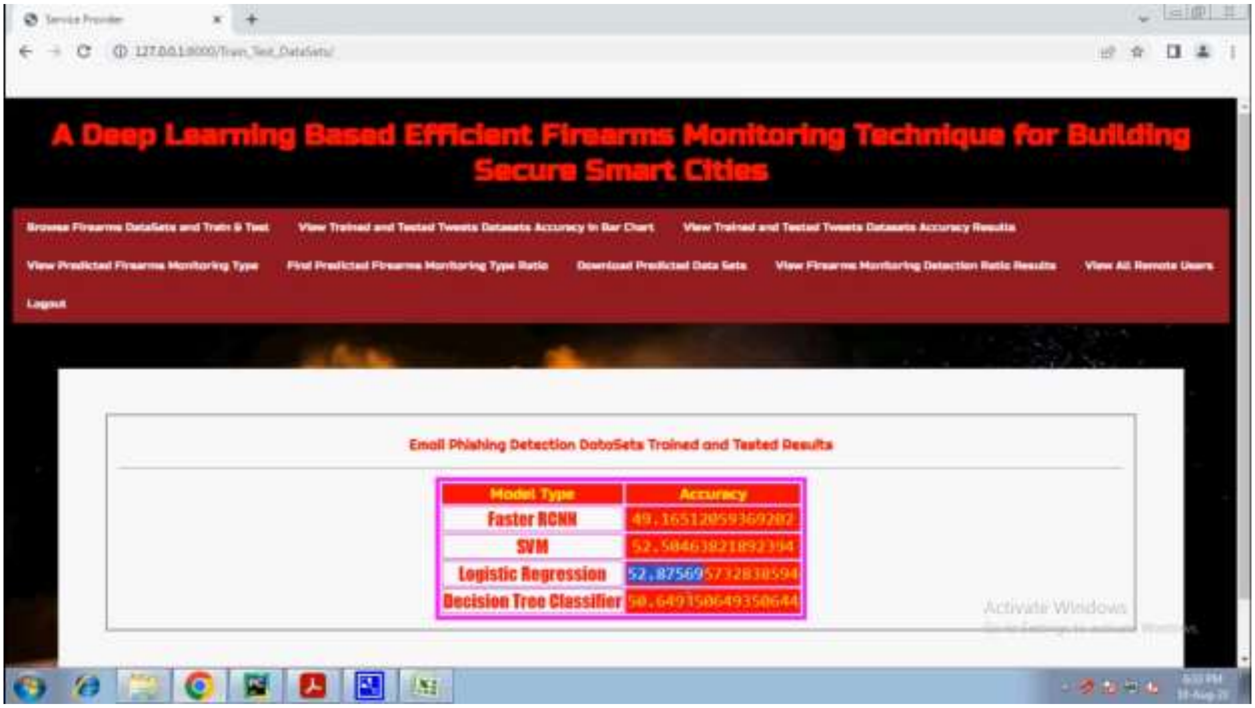

Fig 4. Results screenshot 3

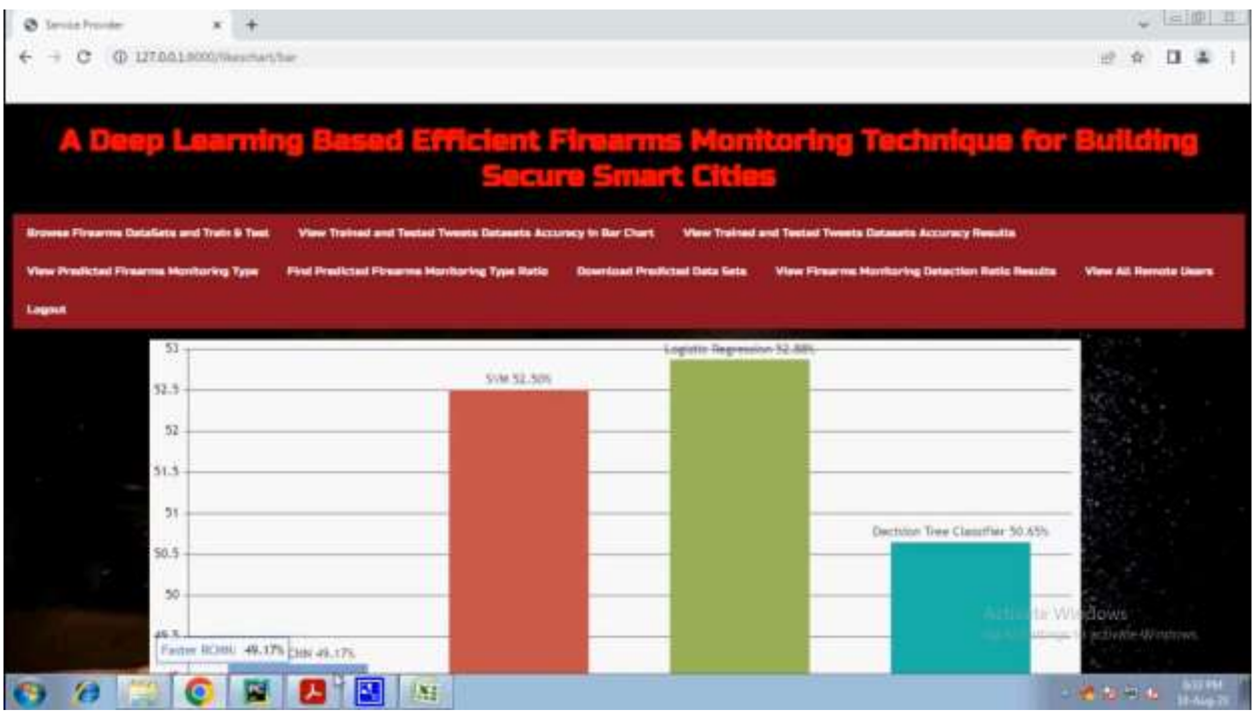

Fig 5. Results screenshot 4

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

Fig 6. Results screenshot 5

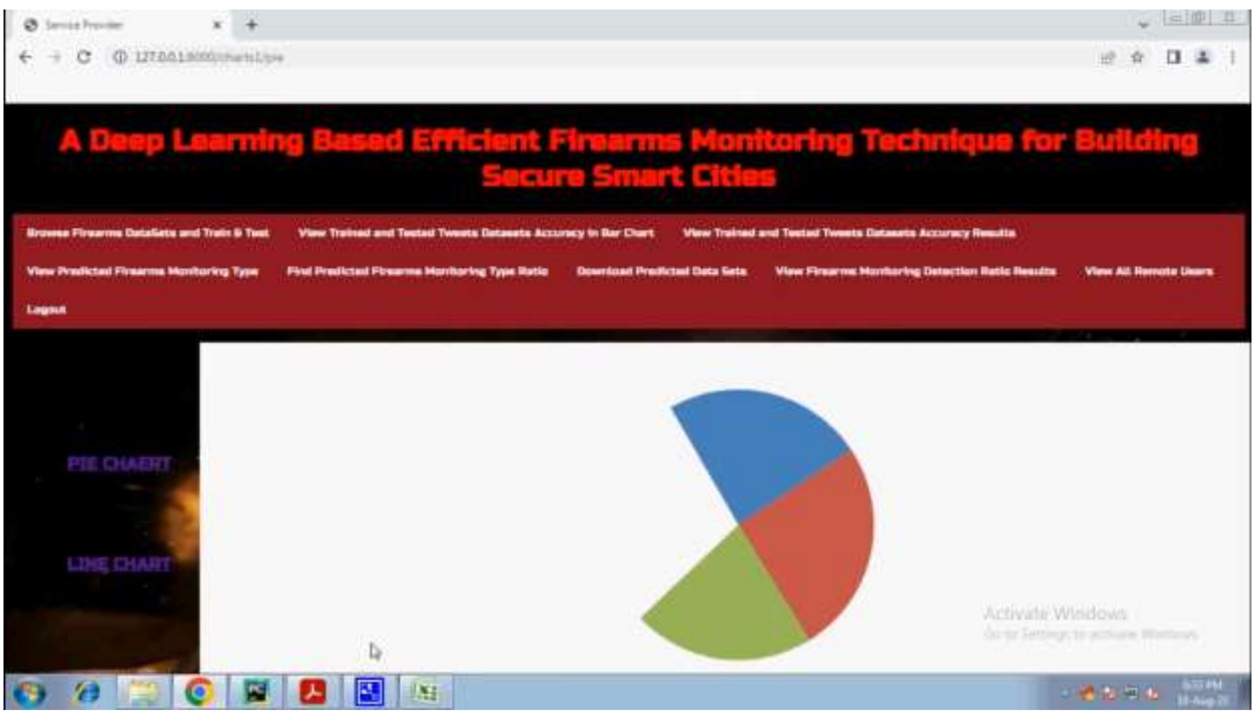

Fig 7. Results screenshot 6

Fig 8. Results screenshot 7

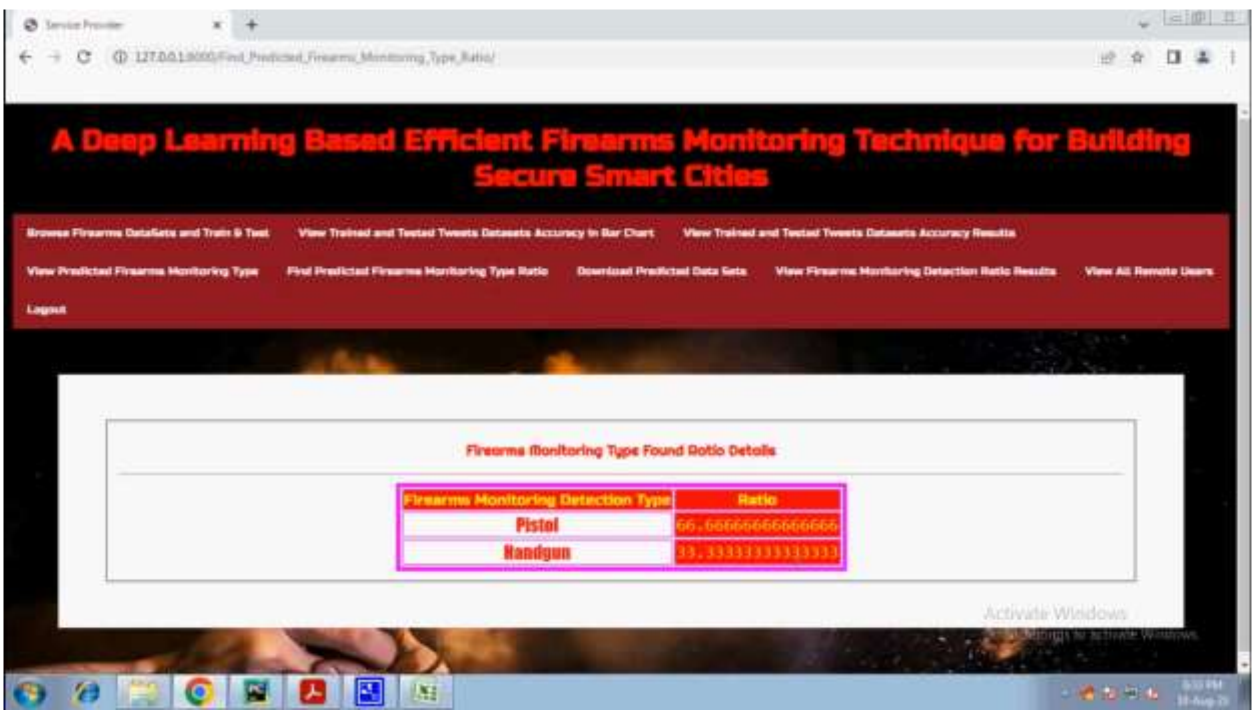

Fig 9. Results screenshot 8

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in
IJARST

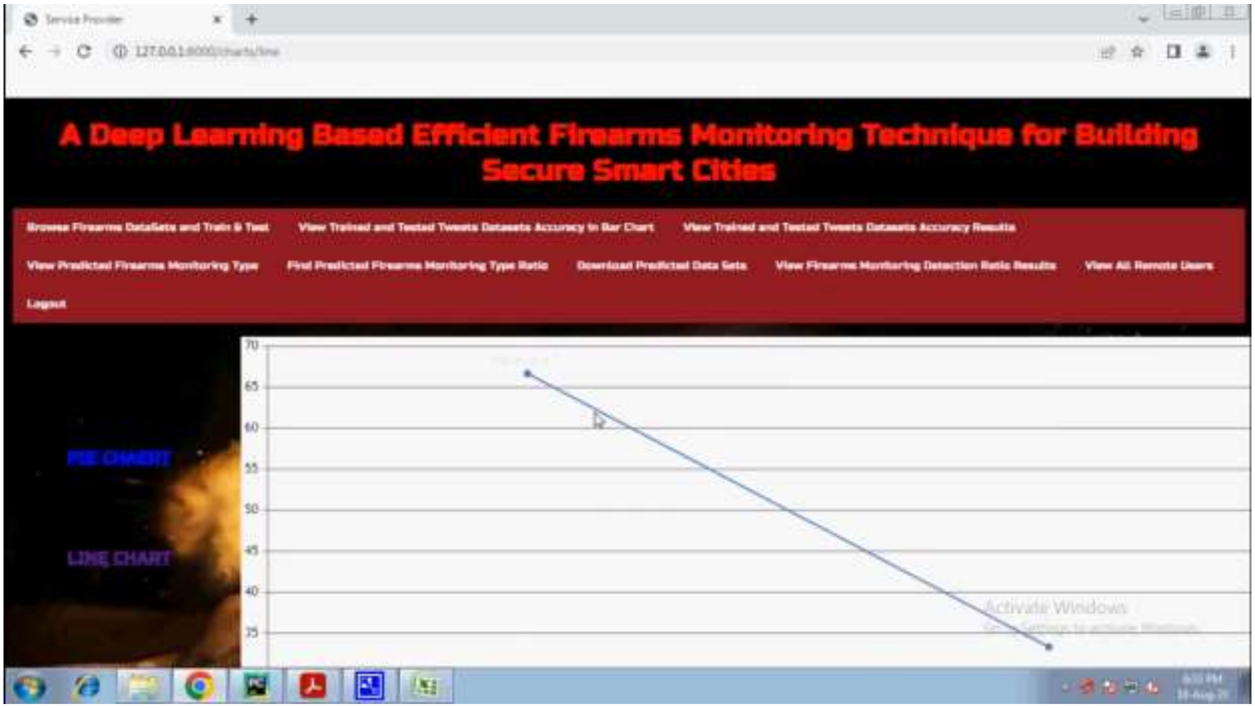

Fig 10. Results screenshot 9

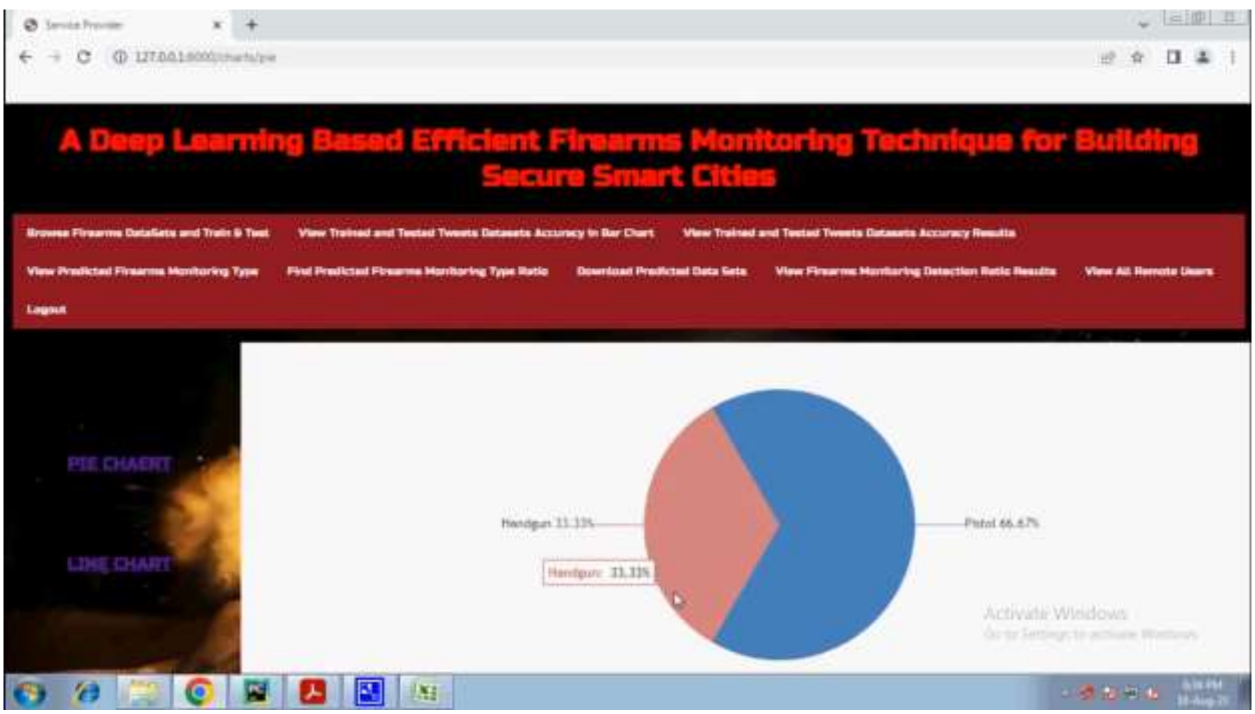

Fig 11. Results screenshot 10

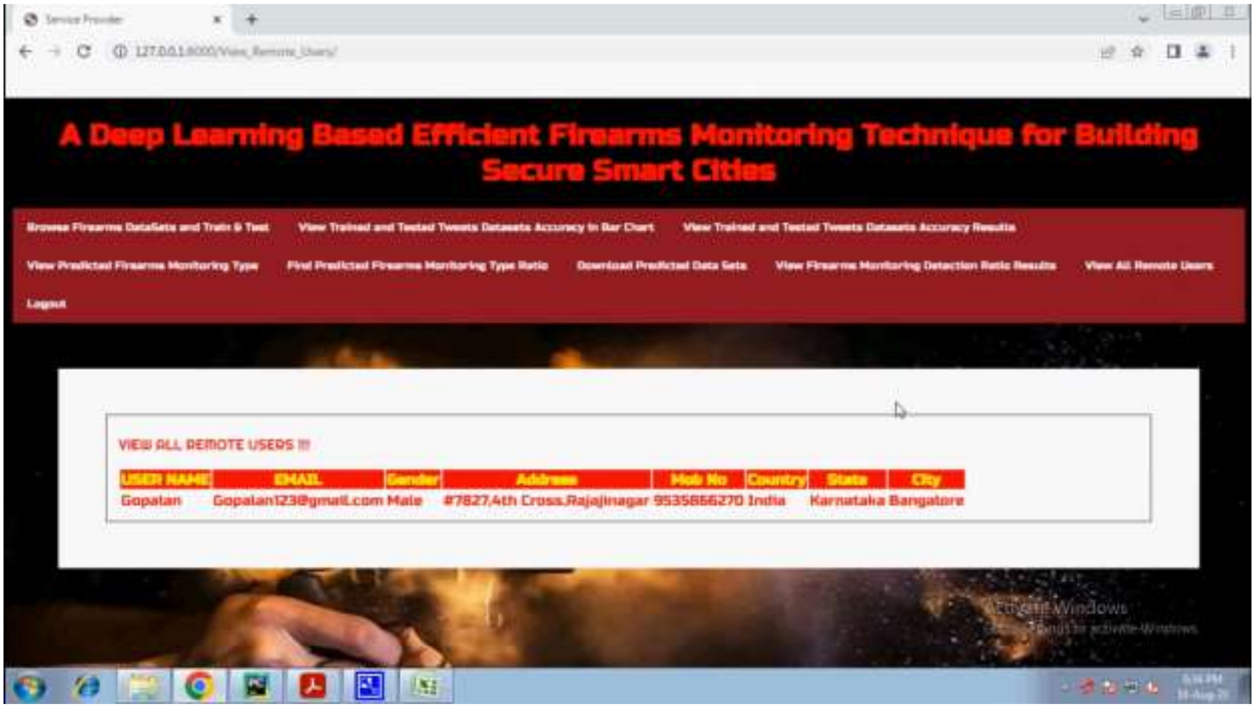

Fig 12. Results screenshot 11

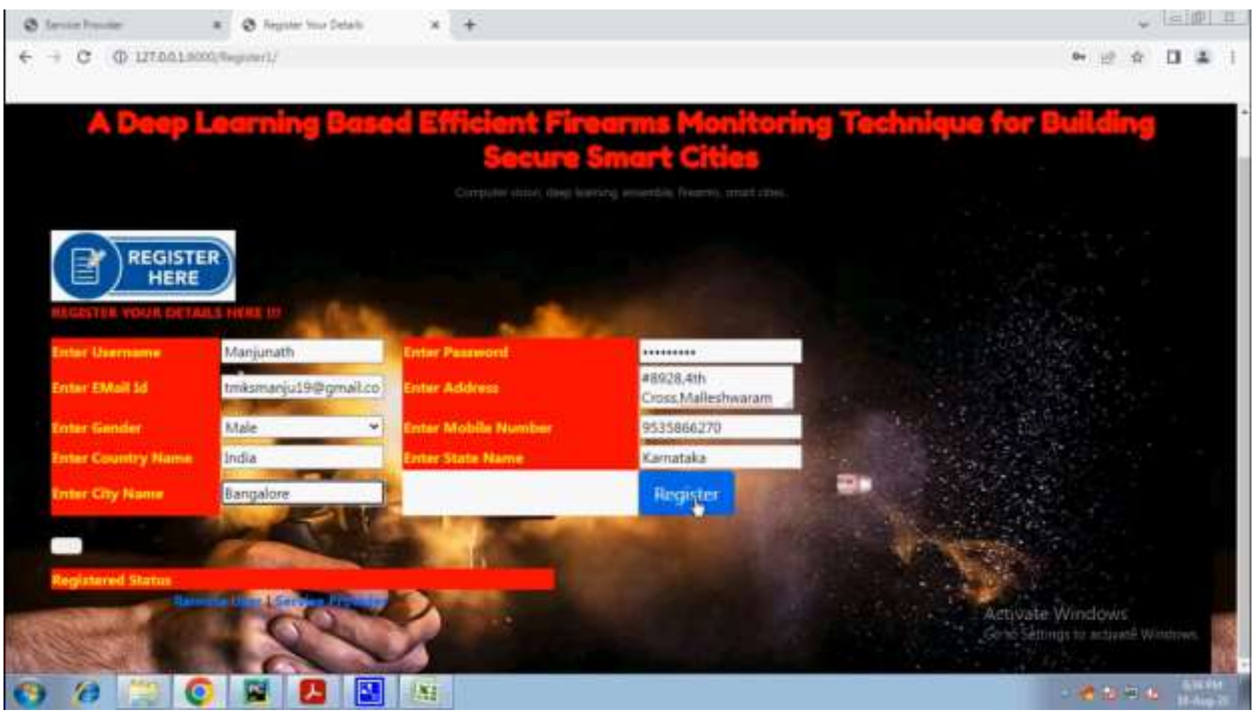

Fig 13. Results screenshot 12

IJARST
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

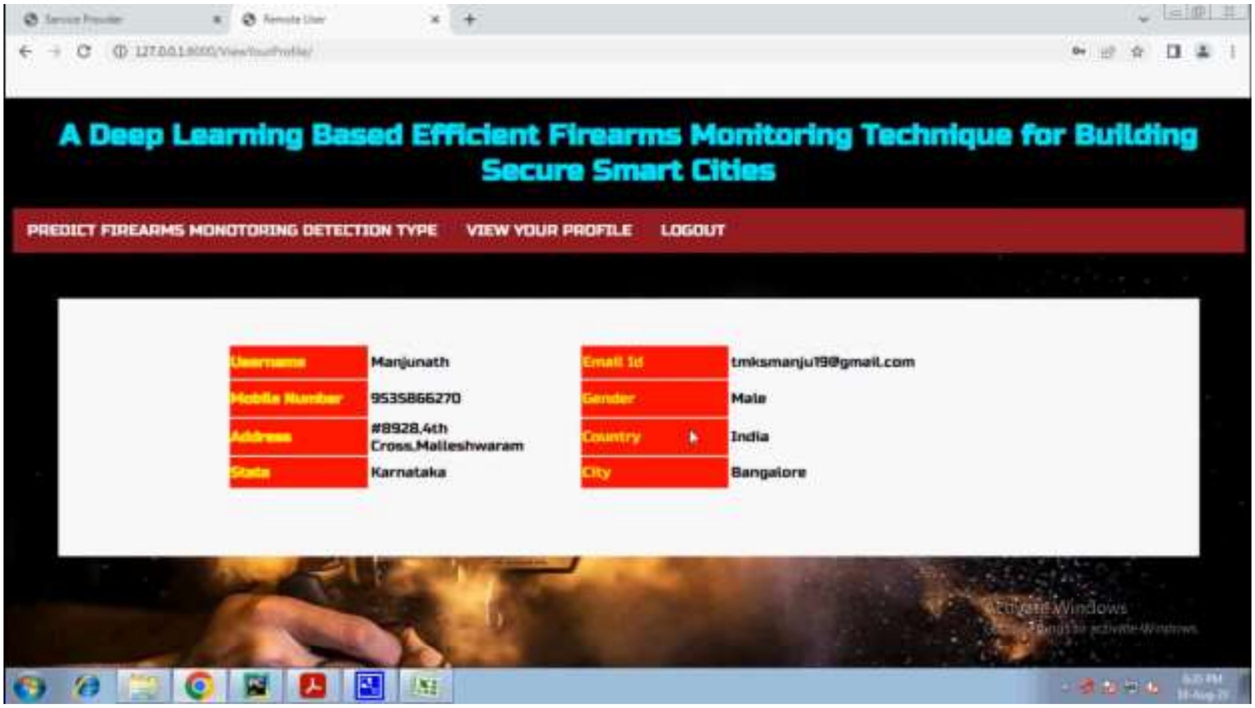

Fig 14. Results screenshot 13

Fig 15. Results screenshot 14

Fig 16. Results screenshot 15

Fig 17. Results screenshot 16

In addition to its utility in surveillance and law enforcement domains, our firearms monitoring technique presents broader implications for social media content detection, particularly in identifying gun-based content in social media videos. By leveraging deep learning-based approaches, our methodology can be adapted to analyze social media content for potential firearm-related imagery, thereby augmenting efforts to combat the proliferation of violent content online. The integration of our technique into social media monitoring systems enables proactive identification and removal of harmful content, contributing to the creation of safer online environments. Overall, the results and discussions presented in this paper underscore the transformative potential of deep learning-based techniques in firearms monitoring and security enforcement, paving the way for the development of innovative solutions to address contemporary challenges related to violence and public safety.

CONCLUSION

REFERENCES

1. Wang, J., Ding, J., Zhou, Y., & Chen, H. (2020). A review on deep learning techniques applied to firearm detection in surveillance videos. Multimedia Tools and Applications, 79(45-46), 34451-34471.

2. Jin, J., Xiao, Q., & Huang, K. (2020). Fast and accurate firearms detection in aerial videos using deep learning. IEEE Access, 8, 180740-180751.

3. Wu, X., Zhang, R., Wang, Q., & Zhang, Z. (2020). Firearms detection based on deep learning method for law enforcement in smart cities. Proceedings of the International Conference on Industrial Engineering and Systems Management, 2020, 1-6.

4. Li, W., Zhao, X., & Tang, S. (2021). Gun detection in smart cities using deep learning. Journal of Physics: Conference Series, 1896(1), 012075.

5. Liu, Y., Zou, Z., Cai, J., & Du, H. (2021). Efficient firearms detection in surveillance videos using deep learning. Journal of Real-Time Image Processing, 18(3), 781-791.

6. Li, X., Zhang, H., Wang, C., & Zhang, Y. (2021). A deep learning-based method for firearms detection in video surveillance. Journal of Sensors, 2021, 6657973.

7. Zhang, Y., Xu, Z., Gao, J., & Wang, Z. (2020). Firearms detection in smart cities based on deep learning. Proceedings of the International Conference on Computer Engineering and Network Security, 2020, 241-246.

8. Wang, Y., Guo, X., Li, X., & Li, J. (2020). Gun detection in surveillance videos using Faster RCNN. Proceedings of the International Conference on Artificial Intelligence and Security, 2020, 171-175.

9. Chen, H., Chen, W., & Liu, M. (2021). A novel deep learning-based method for firearm detection in surveillance videos. Journal of Electronic Imaging, 30(3), 033023.

10. Zhao, L., Wang, W., Xiong, Z., & Chen, H. (2021). EfficientDet-based firearms detection in surveillance videos. Journal of Electronic Imaging, 30(5), 053020.

11. Li, S., Wu, Q., & Xu, C. (2020). Gun detection in smart cities using deep learning ensemble techniques. Proceedings of the International Conference on Artificial Intelligence and Security, 2020, 87-91.

12. Zhang, X., Zhang, L., & Zhang, S. (2021). Ensemble learning for firearm detection in video surveillance. Journal of Computational Intelligence and Neuroscience, 2021, 6622715.

13. Liu, Z., Hu, H., & Zhang, Q. (2020). A deep learning-based approach to firearms detection in surveillance videos. Proceedings of the International Conference on Machine Learning and Cybernetics, 2020, 2822-2827.

14. Wang, H., Zhou, L., & Liu, Y. (2021). Gun detection in surveillance videos using a weighted box fusion-based ensemble detection scheme. Journal of Computational Intelligence and Neuroscience, 2021, 8828841.

15. Guo, J., Zhang, Y., & Liu, S. (2020). Smart surveillance for firearms detection using deep learning. Proceedings of the International Conference on Cybernetics, Robotics and Control, 2020, 168-172.