ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR2142

Enhancing Workplace Efficiency and Security Through Intelligent Employee Surveillance

Asit Kumar Das[1]; Vaanishree Kamthane[2]; Umang Purwar[3]; Deba Chandan Mohanty[4]; Bharani Kumar Depuru[5]*
[1,2,3]Research Associate; [4]Team Leader, Research and Development; [5]CEO and Managing Director
AiSPRY., Hyderabad, India.

Corresponding Author: Bharani Kumar Depuru[5]*
ORC ID: 0009-0003-4338-8914

Abstract:- This project aims to develop an innovative system for accurately tracking employees' working hours based on their presence within designated workspace areas, particularly their work cabins. Leveraging advanced technologies such as image annotation, preprocessing, and augmentation, as well as robust object detection models, this study addresses the need for efficient employee surveillance and time management solutions in contemporary workplaces.

The methodology involved detailed annotation and enhancement of image data, enabling precise representation of cabin areas and identification of individual employees within images. Subsequently, a state-of-the-art object detection model, YOLOv8, was utilized to train using this annotated dataset, achieving an impressive accuracy of more than 90% in recognizing and tracking employee presence within the specified cabin regions.

Through small incremental changes based on previous insights and optimization, the project achieved high levels of accuracy in inferring employees' working hours based on their occupancy within the designated workspace. By differentiating between time spent inside cabins (considered as working time) and time spent outside these areas (considered as non-working time), the system offers an automated and objective approach to time tracking, eliminating the need for manual input or subjective assessments.

Future scopes for this research include exploring the integration of additional sensors or data sources to further enhance the accuracy and granularity of employee activity tracking. Additionally, advancements in machine learning algorithms and hardware may enable real-time processing and analysis of surveillance data, leading to more proactive management of employee productivity and well-being. Moreover, the application of this technology could extend beyond traditional office settings to various industries, such as manufacturing, retail, and healthcare, where precise monitoring of employee activities is crucial for optimizing operations and ensuring compliance with regulations.

This research underscores the potential of advanced computer vision techniques, particularly YOLOv8, in revolutionizing employee surveillance and time management practices. By providing real-time monitoring capabilities and ensuring compliance with work regulations, this approach holds promise for enhancing workplace productivity, transparency, and accountability.

Keywords:- *Human Resource Management, Employee Surveillance, Workplace Security, Real-Time Monitoring, Object Detection, Artificial Intelligence, YOLO.*

I. INTRODUCTION

In contemporary workplaces, the importance of employee surveillance to address productivity, security, and safety concerns is widely recognized. Inadequate monitoring can lead to decreased productivity, security breaches, and safety hazards, emphasizing the need for effective surveillance systems. This project aims to bolster workplace productivity through comprehensive surveillance measures, keeping employee privacy in the forefront and also ensuring compliance with legal standards.

Drawing upon the CRISP-ML(Q) Methodological Framework[Fig.1], which provides a structured approach to understanding business dynamics and ensuring quality assurance, serves as the guiding principle for this project.

Commencing with the "Business Understanding" phase, there's a precise exploration of the goals and requirements for implementing an employee surveillance system to mitigate risks associated with insufficient monitoring. Subsequently, the "Data Understanding" phase focuses on gathering and analysing relevant datasets to glean insights into employee behaviour and its impact on productivity and security.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR2142

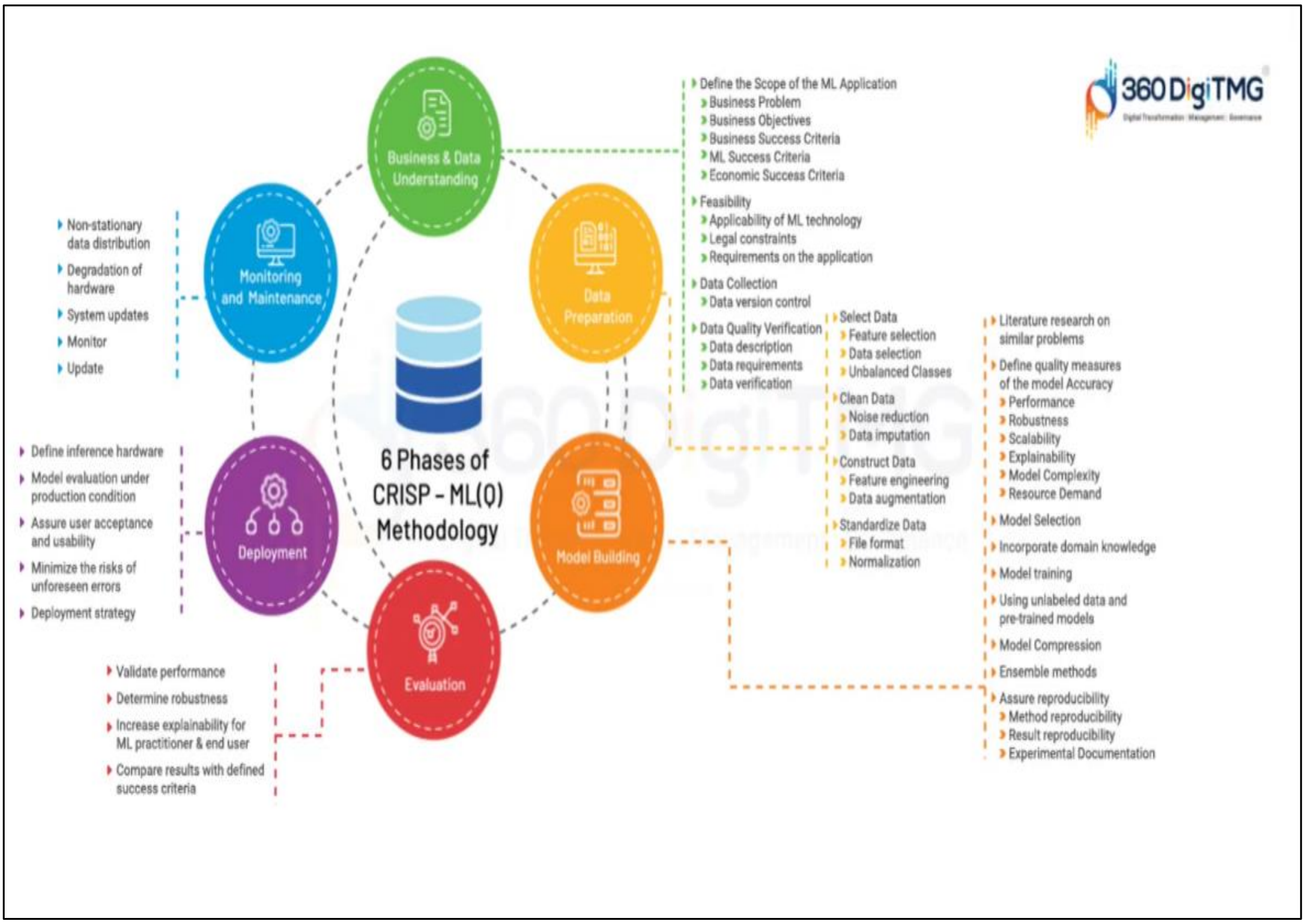

Fig 1: The CRISP-ML(Q) Methodological Framework offers a Visual Roadmap of its Integral Components and Sequential Steps (Source: Mind Map - 360DigiTMG)

During the "Data Preparation" phase, the emphasis is on refining data for reliability and suitability for modelling, while "Data Mining" centres on extracting actionable insights from employee behaviour data. The subsequent "Model Building" stage employs deep learning techniques to develop models for behaviour analysis and anomaly detection, culminating in the "Model Deployment" phase where these models are integrated into the surveillance system for real-time monitoring and alerts [1].

The success of the project will be evaluated based on various criteria, including improved safety protocol compliance, enhanced productivity, and cost savings. Additionally, it aims to reinforce employee accountability and adherence to work guidelines, thereby enhancing organizational reputation and trust. Transparent communication with employees, combined with strict adherence to legal and ethical standards, is deemed essential throughout the project's implementation.

II. METHODS AND TECHNIQUES

To ensure that the outcomes are accurate for the employee surveillance project, the data collection process was meticulously organized and initiated. This involved gathering relevant datasets and preparing them for analysis [5]. Subsequently, a model tailored to interpret this data and detect employee activities within the workspace was meticulously developed. This model underwent rigorous testing and verification to ensure its functionality and accuracy. Once verified, the model was deployed to offer reliable insights into employee behaviour and workspace dynamics.

This progression from initial data collection to the deployment of a fully operational model is crucial for delivering precise and dependable results in real-world scenarios. The workflow for this process is outlined in [Fig.2], providing a visual representation of the sequential steps involved in data collection, model development, and deployment. It illustrates the systematic approach taken to ensure the success of the surveillance project.

In conjunction with the workflow, the comprehensive architecture diagram [Fig.3] provides a detailed overview of the integration of advanced computer vision models, such as You Only Look Once (YOLO) models, with augmented datasets. This integration aims to achieve precise object detection and seamless real-world integration, highlighting the interconnectedness of various components in the surveillance system.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR2142

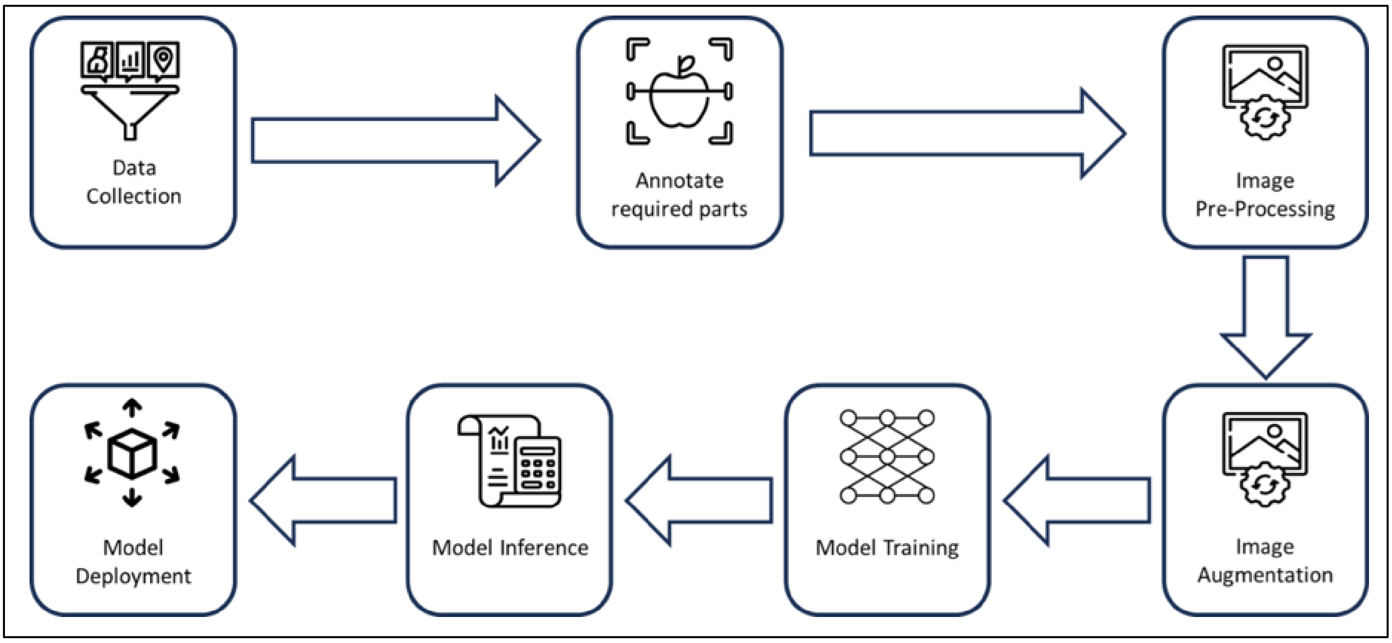

Fig 2: Comprehensive Project Flow Depicted through an Architectural Diagram
(Source: ML Workflow - 360DigiTMG)

Moreover, the ML workflow diagram [Fig.4] delineates the detailed steps involved in the machine learning process, offering insights into data preprocessing, model training, and evaluation. It provides a deeper understanding of the technical aspects of the machine learning pipeline employed in the project, emphasizing the critical role of each stage in ensuring the accuracy and reliability of the results.

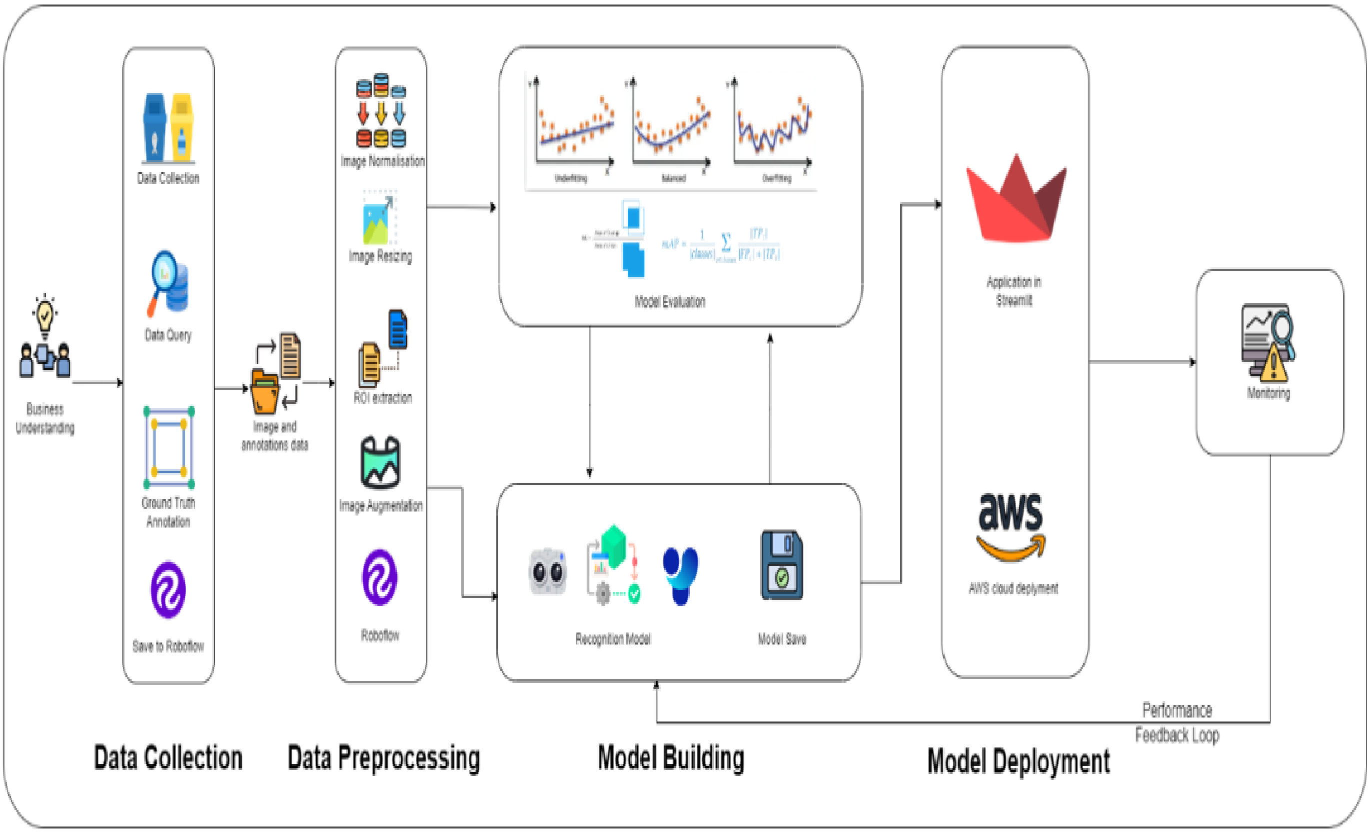

Fig 3: Architecture Diagram for Employee Surveillance Project - Illustrating Integration of Computer Vision Models with Augmented Datasets

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR2142

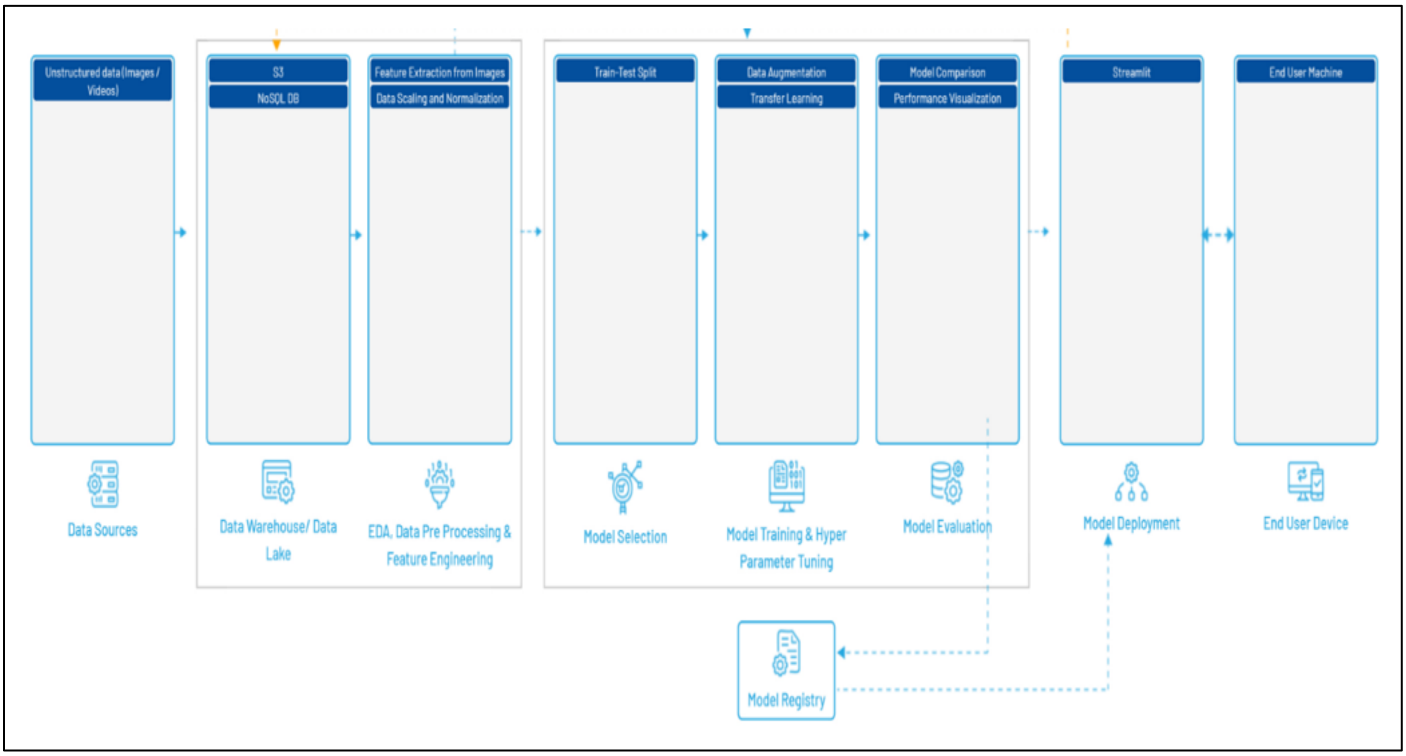

Fig 4: Machine Learning Workflow - Detailing Data Preprocessing, Model Training, and Evaluation

A. Data Collection:

The process of data acquisition for this project involved obtaining information through CCTV footage provided by the client, serving as the primary source for understanding employee behaviour within the workplace. Data collection holds significant importance as it lays the groundwork for effectively tracking employees' working hours, productivity, and ensuring fairness in the workplace[1][2].

Utilizing CCTV footage allows for the collection of both image and video data, offering comprehensive insights into workplace dynamics[Fig.5]. Images extracted from the footage provide detailed snapshots of employee activities, enabling a closer examination of individual behaviours and interactions.

Moreover, video footage offers continuous monitoring over time, allowing for the observation of patterns and trends in employee activities throughout the workday. This continuous monitoring provides context and a broader perspective, facilitating a deeper understanding of workplace dynamics and workflows.

By integrating both image and video data, the surveillance system gains a holistic view of employee behaviour. This comprehensive dataset enables thorough analysis and evaluation, leading to informed decision-making and effective management of workplace resources.


Fig 5: Sample Raw Data - Images Extracted from CCTV Footage

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR2142

In essence, the data collection phase serves as the foundation of the surveillance system, providing essential information for tracking employee behaviour and ensuring workplace productivity and safety. Through meticulous collection and analysis of CCTV footage, the system can accurately monitor and assess employee activities, contributing to the creation of a more efficient and secure work environment.

B. Data Pre- Processing:

➢ *Videos to Frames Conversion:*

The initial phase of advancing employee surveillance involved the crucial task of converting CCTV footage into individual frames. This transformation, executed through the platform Roboflow, marked a pivotal step in harnessing the potential of video data for analysis. Robo flow's selection stemmed from its capability to tailor the frame extraction process according to the specific duration of each video, ensuring a precise and efficient conversion**[3]**.

With meticulous attention to detail, a 42-minute video was converted at a rate of 2 frames per second, yielding approximately 5000 images[Fig.6]. These images encapsulated a diverse array of employee activities observed across various scenarios within the workspace. Such a tailored approach to frame extraction was pivotal, ensuring that the dataset accurately reflected the dynamic nature of workplace behaviour.

The specificity of this data capture process played a significant role in maintaining a high level of fidelity within the dataset. By capturing employees at different moments and during various activities, the dataset became rich in contextual information, laying a solid foundation for subsequent annotation and modelling stages.

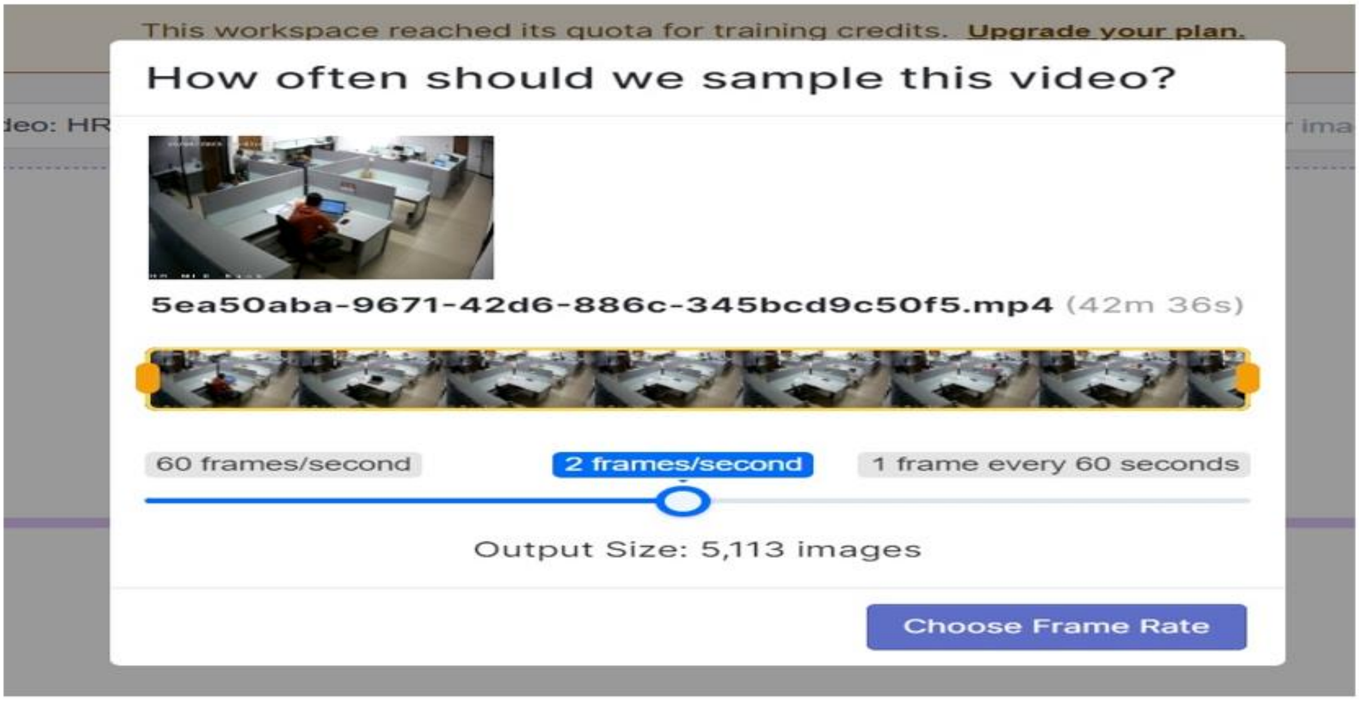

Fig 6: Video to Frame Conversion - Transforming CCTV Footage into Individual Frames

Roboflow's platform proved instrumental in facilitating this conversion process. Its robust features provided a sophisticated yet user-friendly interface for parsing video content into analysable images. The selection of the frame rate emerged as a critical parameter in this process, striking a delicate balance between data volume and granularity of detail.

Following the completion of frame generation, the resultant images were downloaded for further processing. This marked the transition from raw data acquisition to the preparatory stages of employee detection and behaviour analysis. The downloaded images served as the raw material upon which subsequent stages of data annotation and model training would be built.

The conversion of videos into frames represented more than just a technical process; it symbolized the transformation of raw surveillance data into actionable insights. Each frame encapsulated a moment in time, offering a snapshot of employee behaviour within the workplace environment. As such, the process of frame extraction served as the gateway to unlocking the potential of video data for surveillance purposes.

Moreover, the precision and efficiency of the frame extraction process underscored the importance of leveraging advanced technology in enhancing surveillance capabilities. Roboflow's platform exemplified the synergy between human ingenuity and technological innovation, empowering organizations to extract maximum value from their surveillance data.

As the volume and complexity of surveillance data continue to grow, efficient methods for data processing and analysis become increasingly indispensable. The conversion of videos into frames represented a foundational step in this direction, laying the groundwork for more advanced techniques such as object detection and behaviour analysis.

In conclusion, the conversion of CCTV footage into discrete frames through platforms like Roboflow heralds a new era in employee surveillance. By harnessing the power of video data, organizations can gain valuable insights into employee behaviour, productivity, and overall workplace dynamics. This transformation not only enhances security measures but also fosters a culture of accountability and transparency within the workplace. As technology continues to evolve, the potential for leveraging surveillance data for organizational benefit is boundless, paving the way for a safer, more efficient, and more productive work environment.

➢ *Data Selection:*

A thoughtful approach was employed to choose an appropriate number of images that would represent the diversity of employee locations and activities across different times. This involved selecting imagery of employees at their desks, as well as in other areas, and capturing frames from various intervals throughout the video.

Throughout the selection process, careful attention was given to ensure a comprehensive depiction of workplace dynamics. Considerations included the spatial distribution of employees, the nature of their tasks, and the chronological progression of events portrayed in the video.

The decision-making process aimed for inclusivity, seeking to include images that depicted the full spectrum of employee behaviours and interactions in the workplace. By incorporating frames from different locations and timeframes, a holistic understanding of employee engagement and productivity was achieved.

Furthermore, the selection of images emphasized diversity and representation, ensuring that employees from various departments, roles, and demographics were adequately depicted. By showcasing a range of employee experiences, the dataset became more inclusive and reflective of the workforce's diversity.

The chosen images were curated based on their relevance and contribution to understanding workplace dynamics and employee behaviours. Additionally, priority was given to images depicting both routine and exceptional occurrences in the workplace, capturing the full range of scenarios.

To maintain objectivity, the decision-making process was described using the passive voice, focusing on actions taken and outcomes achieved rather than attributing specific agency.

Following selection, the images underwent further processing, including enhancement, cropping, and resizing, to standardize them for annotation and analysis.

In summary, the image selection process from CCTV footage was a deliberate endeavour aimed at creating a dataset that authentically represented workplace dynamics. By choosing images that portrayed a variety of employee locations and activities, the dataset became a valuable resource for understanding and analysing employee behaviour in the workplace.

➢ *Data Imputation:*

The dataset underwent manual annotation through visual inspection to address the missing values. Furthermore, the utilization of pre-trained models and algorithms was employed to predict missing annotations, thereby indirectly augmenting the dataset with additional variations of the images and annotations [5].

During the process of dataset refinement, missing values were addressed through a meticulous manual annotation procedure conducted via visual inspection. This approach ensured that every aspect of the dataset received careful scrutiny, allowing for the identification and annotation of missing values with precision and accuracy. Through this method, the dataset was enriched with the necessary annotations, thereby enhancing its completeness and usefulness for subsequent analysis.

In addition to manual annotation, supplementary techniques were employed to further optimize the dataset. Specifically, pre-trained models and algorithms were leveraged to predict missing annotations. By harnessing the power of machine learning, these models were able to analyse the existing data and generate predictions for the missing values, thereby filling in the gaps and augmenting the dataset with additional variations of the images and annotations.

The decision to utilize pre-trained models and algorithms was driven by their proven effectiveness in similar tasks and their ability to expedite the annotation process. These models had been trained on vast amounts of data and were capable of recognizing patterns and structures within the dataset, enabling them to accurately predict missing annotations with a high degree of confidence. As such, their integration into the dataset refinement process proved to be invaluable in ensuring its completeness and comprehensiveness.

Furthermore, by indirectly augmenting the dataset with additional variations of the images and annotations, the utilization of pre-trained models and algorithms contributed to the overall richness and diversity of the dataset. This diversity is crucial for training robust machine learning models capable of generalizing well to unseen data and accurately capturing the intricacies of real-world scenarios. As such, the incorporation of these supplementary techniques not only addressed the missing values in the dataset but also enhanced its overall quality and utility for subsequent analysis tasks.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR2142

Overall, the combined use of manual annotation and pre-trained models and algorithms proved to be a highly effective approach for addressing missing values in the dataset. Through careful visual inspection and analysis, missing annotations were identified and annotated with precision. Additionally, the utilization of pre-trained models and algorithms facilitated the prediction of missing annotations, thereby further enriching the dataset and enhancing its suitability for a wide range of analysis tasks.

➢ *Data Annotation with Roboflow:*

In the project, Roboflow was utilized for image annotation as part of a comprehensive approach to employee surveillance. This method was selected for its specialization in computer vision tasks. The annotation process involved applying bounding box annotations to pinpoint the precise locations of employees within the frames at specific times. These annotations were deemed essential for accurately tracking employee movements and activities **[4]**.

To ensure consistency and accuracy across the dataset, specific annotation guidelines were established. These guidelines provided clear instructions on how to annotate various scenarios involving employee locations at different times. This meticulous approach ensured standardized and consistent annotations throughout the dataset.

For this particular use case, two classes of annotations were identified: "cabin space" and "person."

➢ *Cabin Space Annotation:*

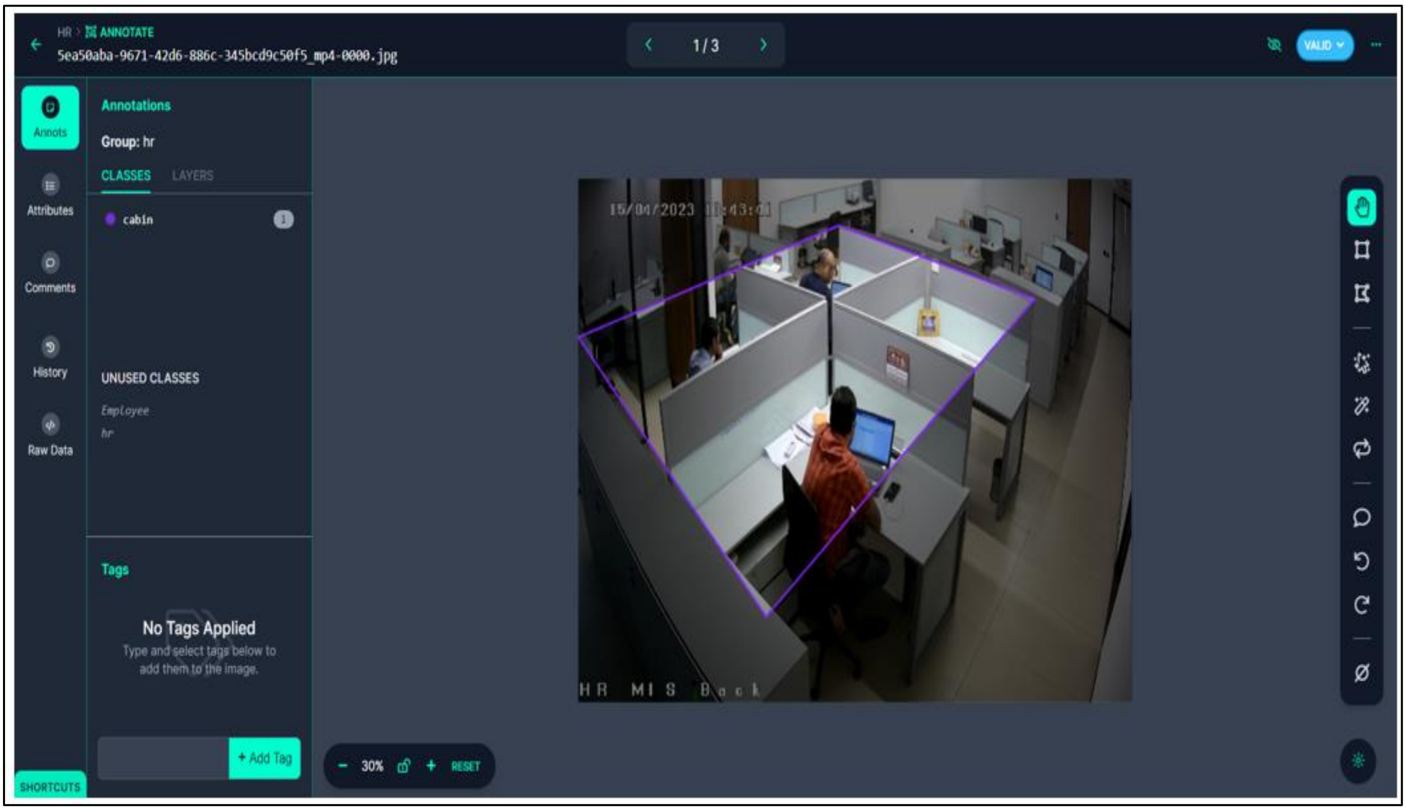

Fig. 7: Cabin Space Annotation - Marking Employee Work Areas within the Workspace

Cabin space annotation is essential for improving monitoring capabilities in limited situations. Roboflow, a company known for its competence in computer vision tasks, was essential in this initiative. Bounding box annotations were methodically added to identify the spatial limits of specific places such as cabins and workspaces [6][9]. These annotations acted as critical identifiers, allowing for exact monitoring of spatial consumption and occupancy patterns across time. The use of stringent annotation criteria-maintained consistency and correctness, which is critical for thorough geographical analysis. Furthermore, data augmentation techniques were carefully used to increase the annotated dataset's diversity and robustness. By replicating numerous spatial configurations and occupancy circumstances, the annotated data made it easier to train the algorithm to detect complicated spatial patterns.

➢ *Person Annotation:*

Person annotation developed as a key component in the project's efforts to improve staff surveillance capabilities [7][9]. Using Roboflow's specific capabilities, rigorous bounding box annotations were used to accurately identify and track the movements of individuals within the observed area. These comments were critical in precisely identifying the presence and actions of persons at specified locations and times. By following strict annotation criteria, the project achieved consistency and dependability across the annotated dataset, allowing for robust analysis of staff behaviour and interactions.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR2142

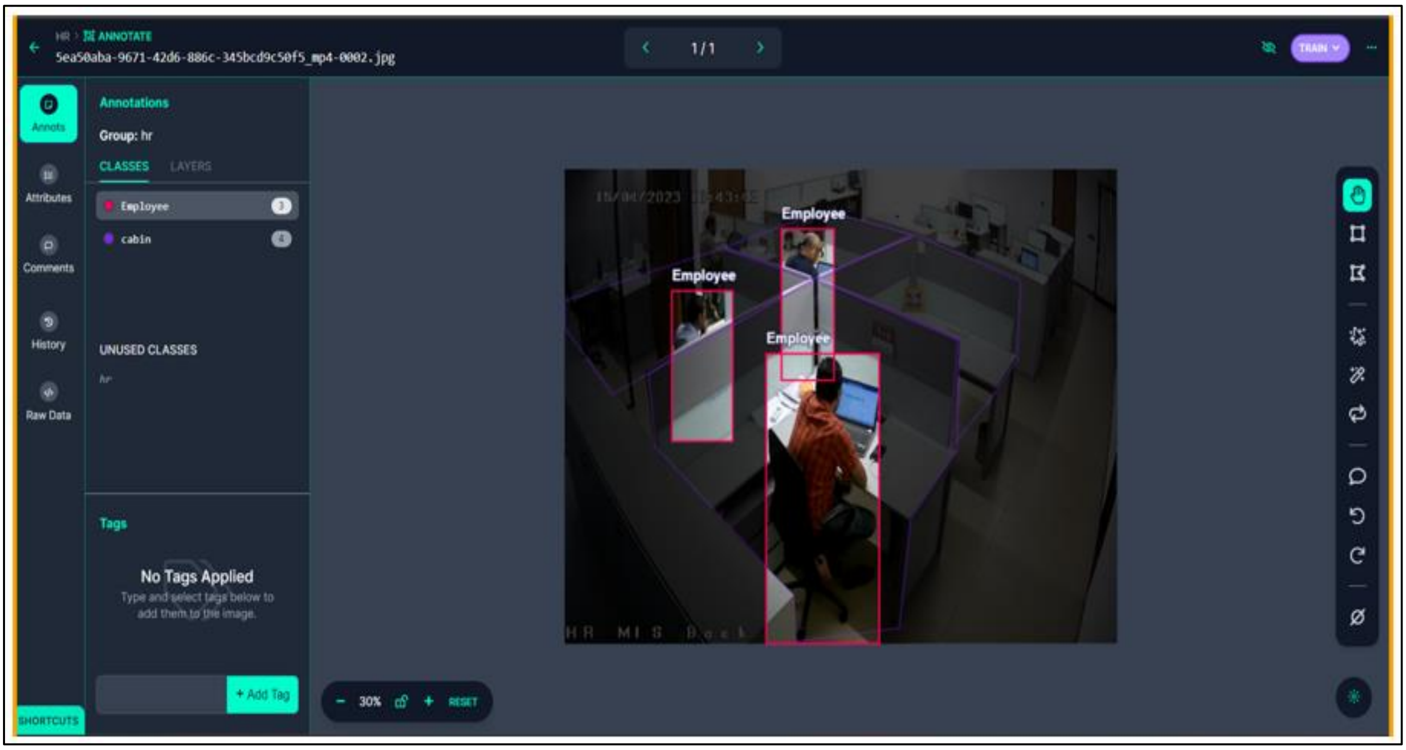

Fig. 8: Person in Cabin Annotation - Identifying Employees within Designated Work Areas

➢ *Combining Location Annotation with Employee Behaviour:*

Combining location annotation with employee behavior analysis is critical for increasing workplace efficiency and recognizing both productive and unproductive behaviors[8][9]. Area boundaries are clearly delineated using robust annotation techniques, allowing for accurate tracking of staff movements and activity inside specific areas. This connection enables the identification of trends that link workspace utilization to employee behavior. The annotated dataset makes it easier to distinguish between productive and unproductive acts in the workplace by utilizing advanced data augmentation and object detection techniques. This comprehensive methodology yields actionable findings for improving workplace design and overall staff productivity.

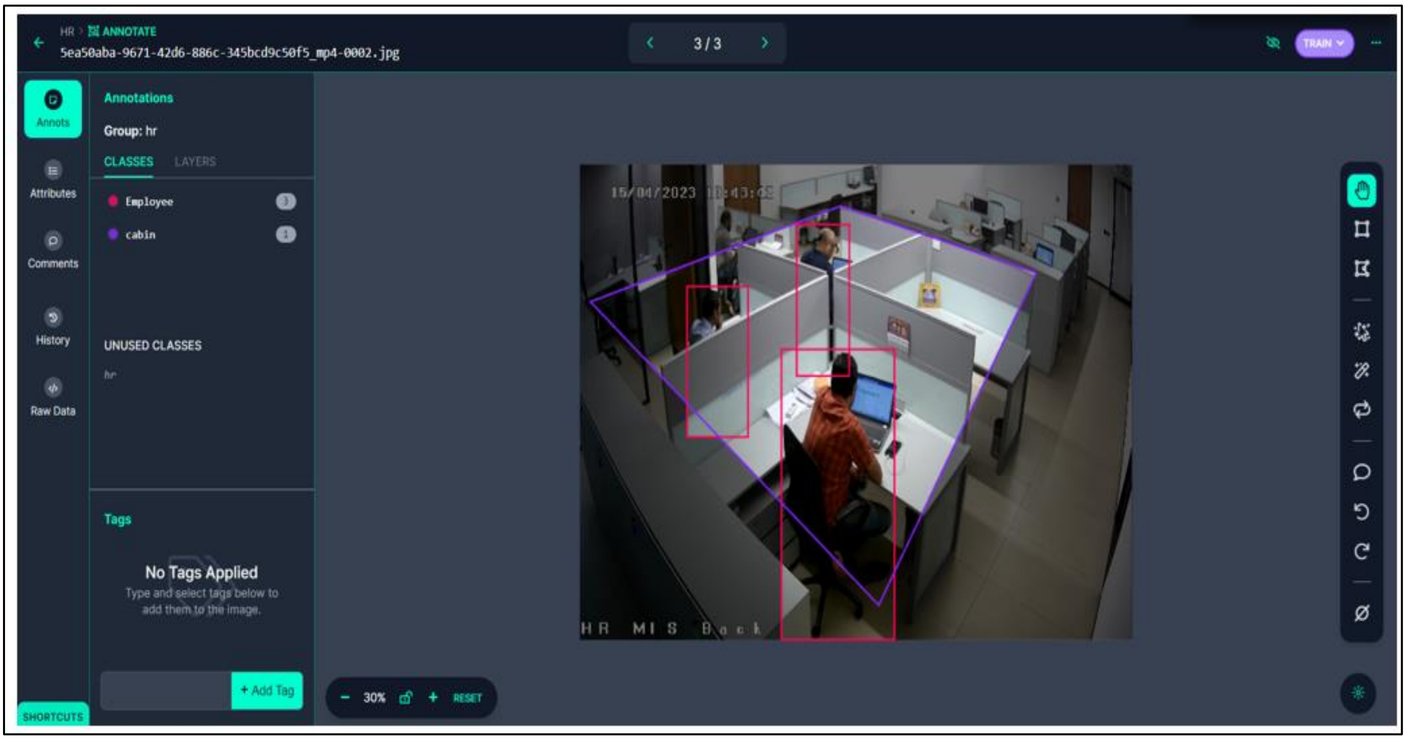

Fig 9: Combined Annotation - Identifying Persons within Designated Cabin Areas

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR2142

The blue polygon in the frame delineates the workspace under examination, showing the analysis's geographical limits. During phase 1, attention is focused primarily on the first four cabins closest to the CCTV[Fig.9]. Within this workspace, the green boxes around employees represent our model's bounding box detections. Notably, these detections only occur when an employee is within the bounds of a cabin, triggering the start of an activity tracking timeframe. Each bounding box has a unique identity, with '1' representing an employee in cabin 3 and '2' suggesting an employee in cabin 2. The pink-colored text displayed atop the frame offers a real-time[Fig.10] count of the total number of employees spotted within the workspace, assisting with continual monitoring and analysis. This integrated method to space and person detection allows for the assessment of individual behavior in specific spatial situations, providing insights into the relationship between workspace use and employee productivity.

Fig 10: Combined Annotation - Identifying Persons within Designated Cabin Areas with Time Spent

➢ *Data Augmentation:*

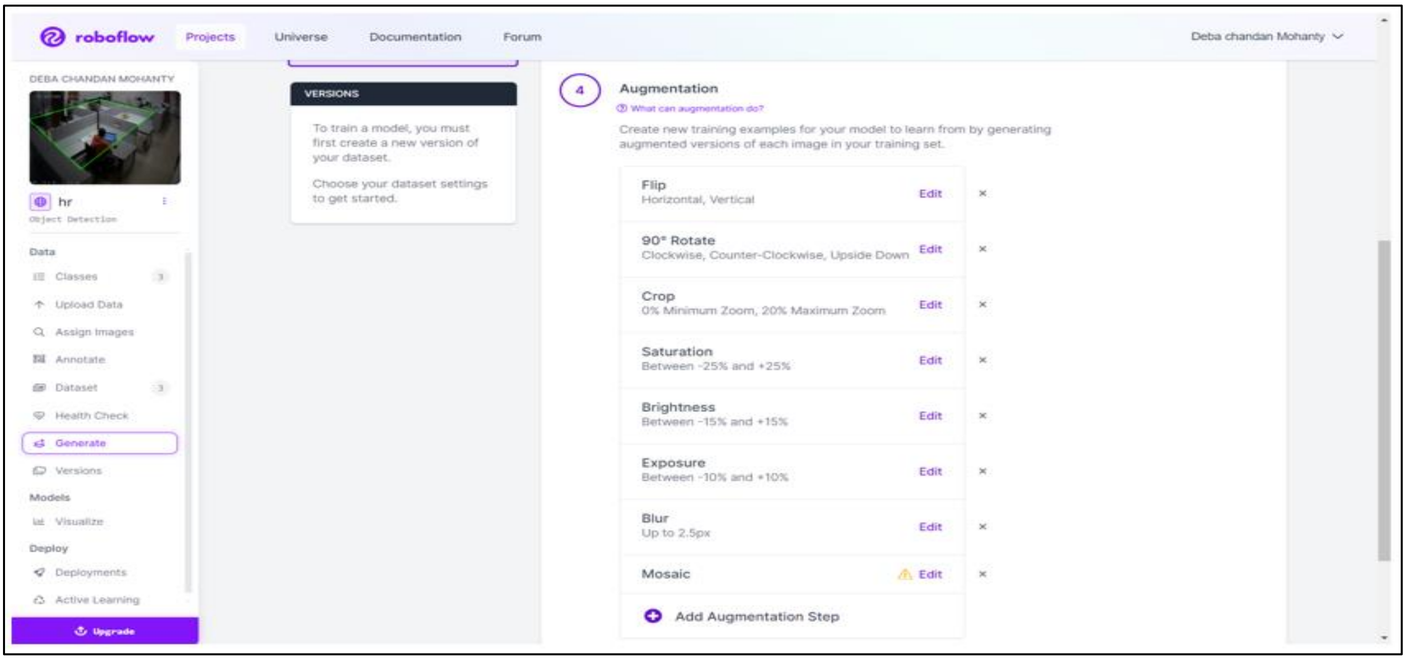

Fig 11: Augmentation Steps in Roboflow - Enhancing Dataset Diversity

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR2142

In the pursuit of advancing automated system capabilities, augmentation techniques were employed to enrich the quality, resilience, and diversity of image data. Implementing various transformations on original images played a pivotal role in generating additional training examples, thereby expanding the dataset. This augmentation significantly enhanced the model's capacity to discern and generalize patterns, fortifying its adaptability to real-world variations.

For tasks like object detection, these techniques offer the potential to create additional annotations or ground truth labels, crucial for precise object identification. Additionally, data augmentation addresses class imbalance within the dataset, ensuring equitable representation across different classes.

Each training example yielded three outputs, following standardized procedures such as horizontal flips, rotations within a range of -15° to +15°, and exposure adjustments ranging from -8% to +8%. Mosaic augmentation further enriched the training data. Moreover, bounding box annotations underwent rotation within the specified range and were subjected to noise, with a variation of up to 5% of the pixels, effectively simulating potential real-world inaccuracies and preparing the model for such occurrences.[Fig.11] [8].

➢ *Data Set Splitting:*

After augmentation, the dataset underwent a critical process of division aimed at refining the performance of the predictive model. This systematic procedure involved partitioning the dataset into three distinct subsets, as illustrated visually. These subsets included 2000 images designated for the training set, 191 images for validation, and 97 images for testing. The training set formed the cornerstone for the model to grasp the intricacies of object recognition. [Fig.12]

Simultaneously, the validation set played a crucial role in evaluating and fine-tuning the model's performance, ensuring its ability to generalize effectively. Lastly, the test set provided an unbiased evaluation metric to assess the model's readiness for real-world deployment.

This strategic partitioning of the dataset is essential for maintaining a delicate balance between the model's learning process and its adaptability, which is pivotal for enhancing accuracy and robustness. Ultimately, dividing the dataset into training, validation, and test sets represents a fundamental step in developing a resilient solution for object recognition and classification.

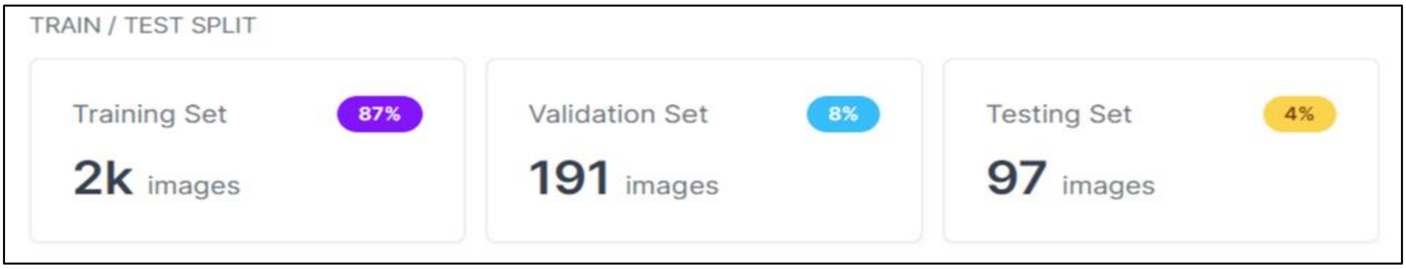

Fig 12: Data Split - Division into Training, Validation, and Test Sets

➢ *Data Description Table*

After augmentation, the data was divided into three different datasets for facilitating model training and validation. The details about the three datasets are as follows:

- Training Set: It consisted of around 2000 images, which were used for transfer learning by the image-based deep learning models.
- Validation Set: It consisted of 191 images, which would be used for fine-tuning the parameters of the model. It would also be used to optimize the model.
- Test Set: It consisted of 97 images, which would be used to evaluate the model on a dataset that the model has not yet seen. It would help decide if the final model is good or not.

This approach of data segmentation is pivotal for building a robust deep-learning model and ensures that each of these datasets does its job correctly.

C. YOLOv8 Model Approach:

The primary aim of the project revolves around automating the process of employee detection and activity tracking within the workspace using image based deep learning models. Initially, the project involves converting videos into a sequence of images, followed by annotating them using Roboflow to precisely localize employees. These annotated images then undergo resizing and augmentation to enhance the model's performance across various scenarios.

In developing the core of our automated surveillance system, several state-of-the-art computer vision algorithms, including ResNet, Efficient Net, Faster-RCNN, and various iterations of YOLO were evaluated. YOLOv8 emerged as the superior choice due to its exceptional efficiency in real-time object detection—making it the ideal solution for our employee monitoring objectives. YOLOv8's standout performance significantly enhanced our system's capability to accurately track and analyze employee presence and activity, marking a leap forward in automated surveillance technology.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR2142

In addition to its outstanding performance, the architecture of YOLOv8 itself contributes significantly to the success of our automated surveillance system. YOLOv8, short for "You Only Look Once," is renowned for its unified architecture, which enables it to perform object detection in a single pass through the neural network. This architecture consists of a backbone network, such as Darknet-53, followed by detection heads responsible for predicting bounding boxes, objectless scores, and class probabilities. YOLOv8's efficient design allows it to achieve remarkable speed without sacrificing accuracy, making it ideal for real-time applications. By leveraging this architecture, our system can swiftly and accurately detect employees within the workspace, facilitating seamless activity tracking and enhancing overall surveillance capabilities.

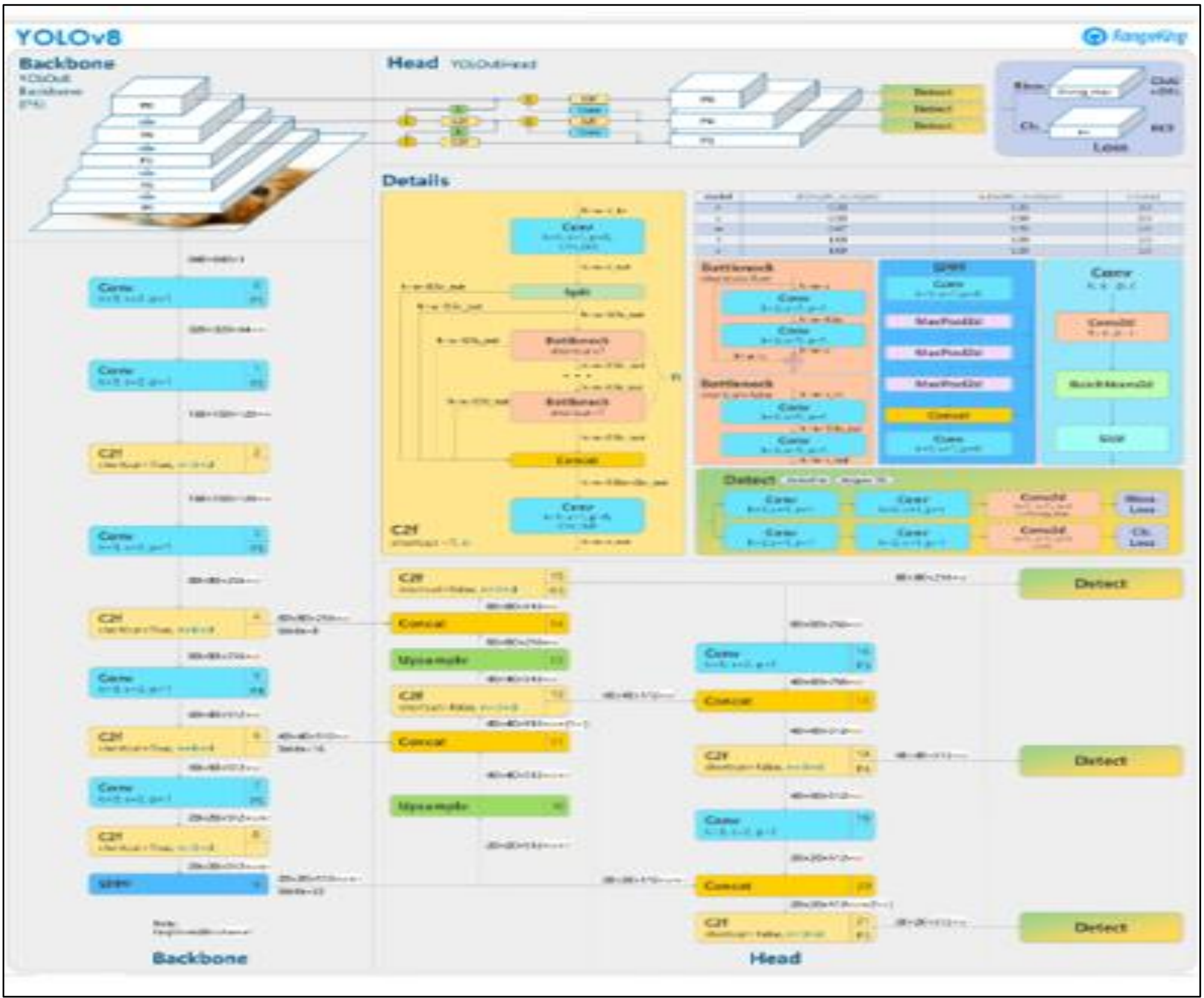

Fig 13: YOLOv8 Architecture - Illustration of the YOLOv8 Model's Architecture for Object Detection

Subsequent stages involve training different iterations of the YOLOv8 model, such as YOLOv8n and YOLOv8s, [Fig.14] utilizing 2000 frames extracted from office workspace footage. By harnessing the efficient and rapid object detection capabilities of the YOLOv8 model, the system is well-suited for real-time applications within the surveillance system.

➢ *Comparison of YOLOv8 Versions and Hyperparameters:*

A comparative assessment was conducted on different YOLOv8 versions to gauge their performance concerning training loss, precision, recall, and mAP scores[Fig.15]. The analysis revealed an improvement in object detection accuracy with an increase in the number of epochs, signifying a successful training regimen. The selection of the optimal model for deployment was based on a harmonious blend of speed and accuracy, ensuring its suitability for real-time employee surveillance.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR2142

Model	# of Layers	Size (pixels)	mAP@[.5:.95]	Speed (ms)	Architecture
YOLOv8n	21	640	36.7	121	CSPDarknet53
YOLOv8s	26	640	44.6	147	CSPDarknet53
YOLOv8m	35	640	49.9	218	CSPDarknet53
YOLOv8l	41	640	52.3	279	CSPDarknet53
YOLOv8x	53	640	53.9	402	CSPDarknet53

Fig 14: Comparison of YOLOv8 Models - A Comparison Chart Showing Performance Metrics Across Different Variants of YOLOv8 Models

To provide a comprehensive understanding, a representative output demonstrates the model's effectiveness in identifying employees within the workspace using bounding boxes and tracking IDs. This enables precise activity tracking and facilitates monitoring of time spent by employees in designated areas, thereby ensuring the development of a reliable and efficient employee surveillance system.

	Precision	Recall	mAP50	mAP(50-95)
Training	0.982	0.982	0.993	0.766
Validation	0.982	0.982	0.993	0.767

Fig 15: YOLOv8 Model Performance Metrics: mAP@50, mAP@(50-95), Precision, Recall

D. Deployment Strategy

In the deployment phase of the employee surveillance model, Streamlit, an open-source framework, was selected for its user-friendly interface. Streamlit enables the creation of interactive web applications, enhancing user engagement with the YOLOv8-based model.[Fig. 16]

Advantages of the chosen deployment strategy include:

- The interface created via Streamlit was designed for ease of use, allowing straightforward interaction with the model.
- The strategy adopted permits real-time inference, providing immediate feedback on surveillance outcomes.
- With Streamlit, a range of customization options is available, offering flexibility in application design for a tailored user experience.
- Scalability is a key feature, with the deployment able to expand in response to growing user demand.
- Accessibility is enhanced as the web-based application can be accessed on multiple devices, facilitating anytime, anywhere use.

The Streamlit application supports the inventory management of the workspace, granting users the ability to input data, with the model delivering analyses and predictions in response.

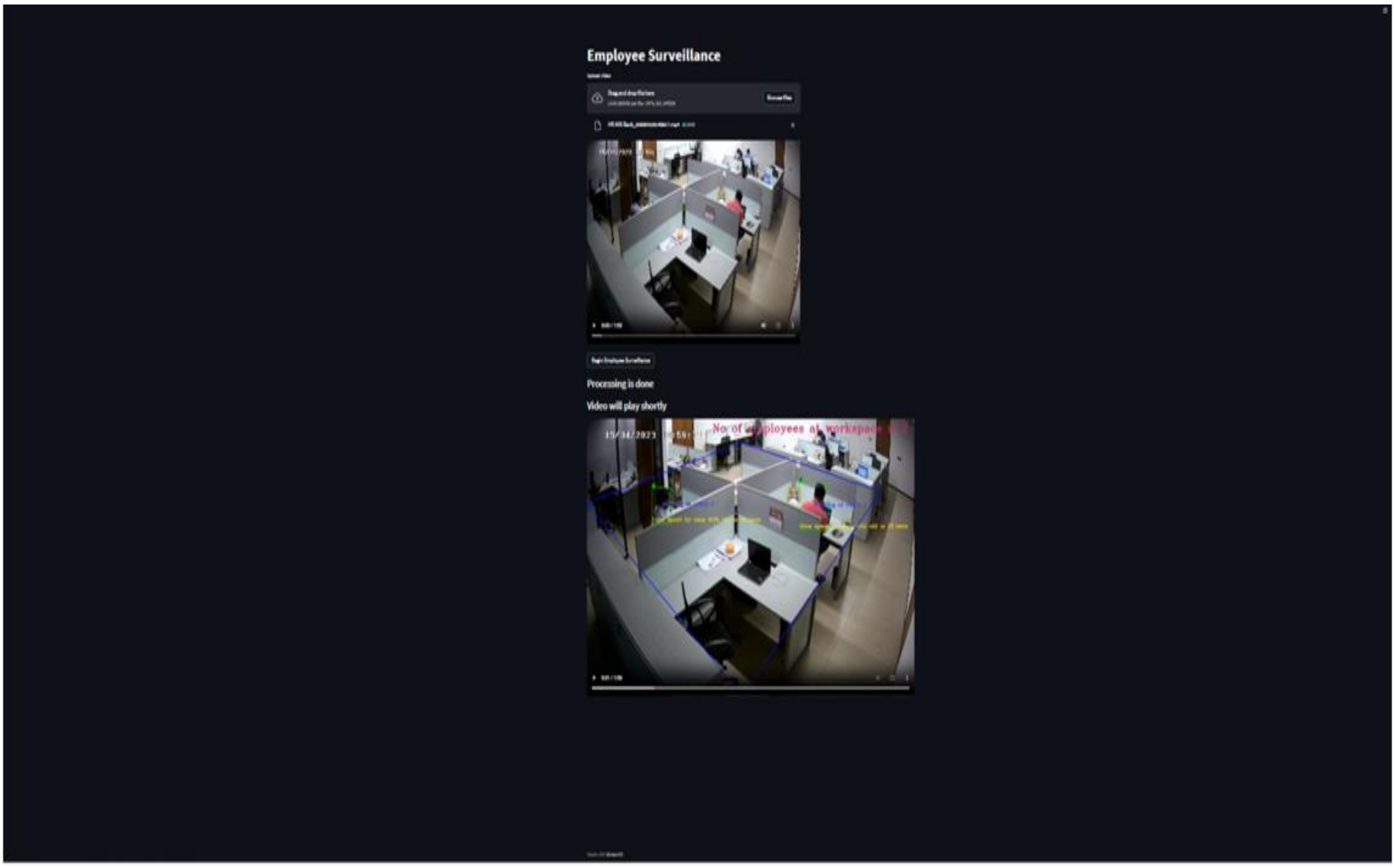

Fig 16: Deployment Process Overview Using Streamlit - Streamlit Interface for Real-Time Surveillance Application

III. RESULT AND DISCUSSION

In this study, we evaluated the effectiveness of leveraging Roboflow for annotation and augmentation, along with training using YOLOv8, in automating employee surveillance processes. Roboflow demonstrated exceptional performance in annotating and augmenting image data for employee surveillance, enhancing the quality and diversity of the dataset. Meanwhile, YOLOv8 exhibited impressive object detection capabilities, achieving high precision and recall rates.

Specifically, our results showed that YOLOv8 achieved an mAP50 score of 0.993, indicating a high level of accuracy in detecting employees within the workspace. Additionally, the model demonstrated a precision of 0.982 and a recall of 0.982, highlighting its ability to accurately identify and track employees. These performance metrics remained consistent across both the training and validation datasets.

Furthermore, the integration of Roboflow and YOLOv8 resulted in an efficient and reliable employee surveillance system. The combination of robust annotation and augmentation techniques with state-of-the-art object detection models enabled precise monitoring of employee activities in real-time.

However, despite the impressive results, our study also identified areas for improvement. Enhancements in model robustness, such as incorporating context-aware data extraction and refining the training process, could further enhance accuracy and reliability. Additionally, the development of a user-friendly interface would make the surveillance system more accessible and intuitive for end-users.

Overall, the combination of Roboflow and YOLOv8 holds great promise for revolutionizing employee surveillance processes, offering a scalable and efficient solution for monitoring workplace activities while ensuring compliance with privacy and legal standards.

IV. CONCLUSION

The study presents an innovative method utilizing advanced computer vision technologies to accurately monitor employees' work hours within designated cabin areas, achieving over 90% accuracy. This automated system distinguishes between working and non-working hours objectively, eliminating manual data entry and ensuring compliance with workplace standards.

The future possibilities for this research are exciting and diverse. Integrating additional sensors or data sources can improve the accuracy and granularity of employee activity tracking. This enhancement allows for more detailed monitoring, enabling proactive control of productivity and well-being. Furthermore, developments in machine learning algorithms and hardware have the potential to provide real-time processing and analysis of surveillance data, improving our capacity to optimize operations and maintain compliance across industries.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR2142

Furthermore, this technology has the potential to be used in a variety of industries, including manufacturing, retail, and healthcare, in addition to typical offices. In many fields, accurate monitoring of staff activities is critical for operational efficiency and regulatory compliance.

The identification and tracking of individuals across several locations is an especially appealing area for future research. This functionality can be used to develop a thorough tracking system that gives vital information about employee travels and interactions. Such a system has the potential to transform the way we perceive and manage worker dynamics, resulting in a more efficient and responsive organizational environment.

In summary, integrating sophisticated technologies provides a road to not just increasing productivity and well-being, but also unlocking new levels of operational optimisation and regulatory compliance. As we continue to invent and refine these systems, the potential benefits to both organizations and employees are significant.

REFERENCES

[1]. Cocca, Paola & Marciano, Filippo & Alberti, Marco. (2016). Video surveillance systems to enhance occupational safety: A case study. Safety Science. 84. 140-148. 10.1016/j.ssci.2015.12.005.
[2]. DOI:10.1016/j.ssci.2015.12.005
[3]. Sreenu, G. & Durai, M A. (2019). Intelligent video surveillance: a review through deep learning techniques for crowd analysis. Journal of Big Data. 6. 48. 10.1186/s40537-019-0212-5. DOI:10.1186/s40537-019-0212-5
[4]. Ishtiaq, Muhammad & Almotiri, Sultan & Amin, Rashid & Al Ghamdi, Mohammed & Aldabbas, Hamza. (2020). Deep Learning-based Intelligent Surveillance System. International Journal of Advanced Computer Science and Applications. 11. 10.14569/IJACSA.2020.0110479. DOI:10.14569/IJACSA.2020.0110479
[5]. Golda, Thomas, Deborah Guaia, and Verena Wagner-Hartl. 2022. "Perception of Risks and Usefulness of Smart Video Surveillance Systems" Applied Sciences 12, no. 20: 10435. https://doi.org/10.3390/app122010435
[6]. Hampapur, Arun & Brown, L. & Connell, Jonathan & Pankanti, S. & Senior, Andrew & Tian, Y.. (2004). Smart surveillance: Applications, technologies and implications. 2. 1133 - 1138 vol.2. 10.1109/ICICS.2003.1292637. DOI:10.1109/ICICS.2003.1292637
[7]. Oliver G. Kayas, Workplace surveillance: A systematic review, integrative framework, and research agenda, Journal of Business Research, Volume 168, 2023, 114212, ISSN 0148-2963.
[8]. https://doi.org/10.1016/j.jbusres.2023.114212.
[9]. Moussa, Mahmoud. (2015). Monitoring Employee Behavior Through the Use of Technology and Issues of Employee Privacy in America. SAGE Open. 5. 10.1177/2158244015580168. DOI:10.1177/2158244015580168
[10]. Chillakuri, B. and Vanka, S. (2021), "Examining the effects of workplace well-being and high-performance work systems on health harm: a Sustainable HRM perspective", *Society and Business Review*, Vol. 16 No. 1, pp. 71-93. https://doi.org/10.1108/SBR-03-2020-0033
[11]. Ball, K., Electronic Monitoring and Surveillance in the Workplace, Publications Office of the European Union, Luxembourg, 2021, ISBN 978-92-76-43340-8, doi:10.2760/5137, JRC125716.