Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004

JOURNAL OF INFORMATION SYSTEM AND TECHNOLOGY MANAGEMENT (JISTM)
www.jistm.com

APPLICATION OF AUTOMATED INTELLIGENT VIDEO SURVEILLANCE SYSTEM WITH DEEP LEARNING CAPABILITY FOR SAFETY MONITORING IN MANUFACTURING SECTOR

Marizuana Mat Daud[1], Mohamad Hanif Md Saad[2*], Nazhatul Hafizah Kamarudin[3], Mohamad Taha Ijab[4], Fazlina Mohd Ali[5]

[1] Institute of Visual Informatics, Universiti Kebangsaan Malaysia, Malaysia
Email: marizuana.daud@ukm.edu.my
[2] Faculty of Engineering & Built Environment, Universiti Kebangsaan, Malaysia
Email: hanifsaad@ukm.edu.my
[3] Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Malaysia
Email: nazhatulhafizah@ukm.edu.my
[4] Institute of Visual Informatics, Universiti Kebangsaan Malaysia, Malaysia
Email: taha@ukm.edu.my
[5] Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Malaysia
Email: fazlina.mohdali@ukm.edu.my
[*] Corresponding Author

Article Info:

Article history:
Received date: 15.09.2023
Revised date: 29.09.2023
Accepted date: 23.10.2023
Published date: 05.12.2023

To cite this document:

Mat Daud, M., Md Saad, M. H., Kamarudin, N. H., Ijab, M. T., & Ali, F. A. (2023). Application Of Automated Intelligent Video Surveillance System With Deep Learning Capability For Safety Monitoring In Manufacting Sector. *Journal of Information System and*

Abstract:

In many cases, occupational and accidental injuries arise from the neglect of employees and the failure of upper management to ensure the constant adherence to safety regulations. Nevertheless, occupational accidents can also result from a combination of factors such as the presence of potential hazards in the workplace and unsafe actions or procedures. Traditionally, to mitigate safety risks in the workplace, security personnel are employed to carry out safety checks and monitoring, a process that can be arduous and inefficient, particularly in large spaces like factories. The objective of this article is to assess previous studies pertaining to the application of intelligent visual surveillance through deep learning methods in enforcing safety regulations. The paper also discusses the use of deep learning in safety management across various sectors. Notably, as far as our knowledge extends, there currently exists no comprehensive survey or initial assessment of deep learning for safety management specifically within manufacturing facilities. This survey is intended to serve as a catalyst for future research into the implementation of intelligent visual surveillance through deep learning for safety management and regulatory purposes.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004

Technology Management, 8 (33), 42-56.

DOI: 10.35631/JISTM.833004

Keywords:

Intelligent visual surveillance, Deep Learning, Safety Risk, Safety Regulation, Manufacturing

Introduction

Engaging in any business activity or event inevitably carries various uncertainties and potential risks. Simply put, a risk involves the likelihood of a negative event occurring and the resulting consequences (Boholm et al., 2019). To manage the risks associated with a specific activity, one can take measures to prevent these potential consequences from happening. This approach is crucial across diverse business sectors, with a particular emphasis on manufacturing. Identifying inherent risks in a given activity is a common practice in these sectors, allowing for informed decisions to either minimize or alleviate these risks to an acceptable level. As highlighted by Major et al. (2018), industries must prioritize the safe operation of their facilities and employ robust risk management techniques when planning and executing their activities. For instance, the utilization of Personal Protective Equipment (PPE) plays a critical role in safeguarding workers and significantly reducing the risk of exposure to hazards, especially in industrial settings. PPE is typically used to shield individuals from workplace health and safety risks. Furthermore, the use of safety gear such as safety boots and hard hats is imperative in hazardous activities involving potential collisions with fixed objects like pipework, machinery, and scaffolding, as well as transport activities with the risk of falling materials, such as hoists, lifting equipment, and conveyors (Major et al., 2018). Consequently, it is evident that risks associated with events of this nature should be either prevented or eliminated.

It is worth noting that the substantial advancements in deep neural networks can be directly harnessed to enhance the intelligence of intelligent visual surveillance (IVS) systems, ensuring adherence to safety management and security requirements. Deep learning (DL) falls within the realm of machine learning (ML) and is a component of artificial intelligence (AI), as illustrated in Figure 1. The rapid proliferation of DL applications spans various domains, including natural language processing, image processing, robotics, medical applications, fault detection (Saufi et al., 2019), and video surveillance. Predictive analysis in video surveillance encompasses multiple modules such as object detection, recognition, action recognition, and the classification of identified actions into categories. For instance, it can identify activities like a person lying down, sitting, standing, or cycling for physical activity monitoring (Fanchamps et al., 2018), or detecting events like backward falls, chest pain, coughing, fainting, headaches, vomiting, or taking medication for patient monitoring (Cocca et al., 2016), as well as distinguishing between normal and anomalous situations for public safety (Sreenu and Saleem, 2019).

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004

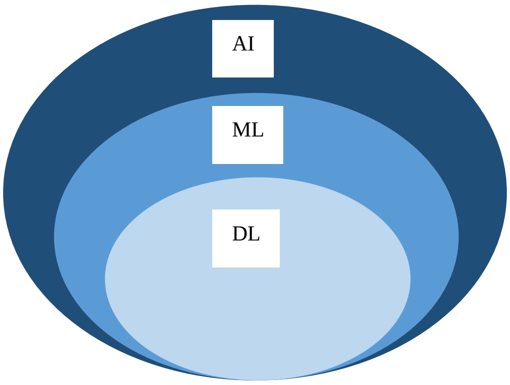


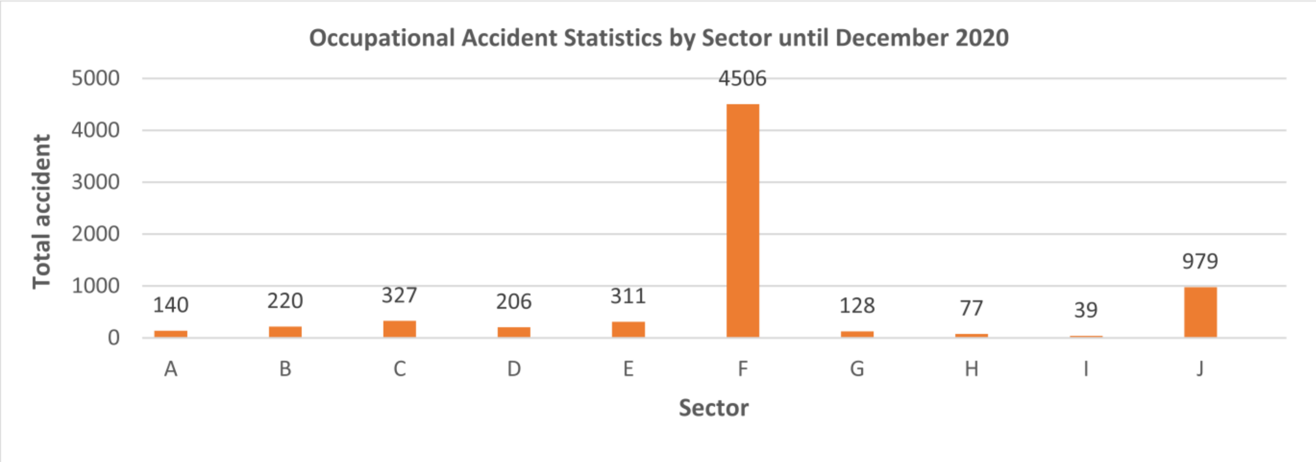
Figure 1: Sub-areas in Artificial Intelligence (AI)

Video surveillance has become ubiquitous in various sectors and locations, including manufacturing (Su et al., 2021), construction (Gul et al., 2020), hospitals (Cocca et al., 2016), public safety (McCall et al., 2021), government buildings (Kayas et al., 2019), railway stations (Chumuang et al., 2018), and ship industries (Patel et al., 2021). Most of these sectors prioritize safety and security within their premises. However, manual video surveillance requires security personnel to be constantly attentive, focused, and alert to any abnormal events, which can be both tedious and time-consuming.

Intelligent visual surveillance (IVS) represents a significant leap forward from traditional manual video surveillance, incorporating automated features like object detection (Patel et al., 2021), facial recognition (Wang et al., 2017), scene interpretation (Gul et al., 2020), anomaly detection (Cocca et al., 2016), risk assessment (Choi et al., 2019), and identification of suspicious activities (Mudgal et al., 2021), (Vallathan et al., 2021). These systems are primarily designed to autonomously identify and interpret objects and events captured by surveillance cameras, especially in scenarios where numerous camera networks cover extensive areas, overwhelming human capacity to process vast quantities of information simultaneously.

IVS is synonymous with enhancing safety enforcement and monitoring. However, there exists a scarcity of research papers and studies on the application of artificial intelligence (AI) for enhancing worker safety, particularly in medium to heavy industries such as manufacturing, where employees face substantial risks and accidents. In Malaysia, the manufacturing sector recorded a total of 4,294 accidents in 2020, encompassing 4,027 cases of non-permanent disability, 209 cases of permanent disability, and 58 fatalities, according to a report by the Department of Occupational Safety and Health (DOSH, 2021). Among all sectors, the manufacturing industry accounted for the highest incidence of accidents during the January to November 2020 period, at 63.26%, as illustrated in Figure 2.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004



A - Hotel and Restaurant
B - Utilities (Electricity, Gas, Water and Sanitary Service)
C - Finance, Insurance, Real Estate and Business Services
D - Construction
E - Transport, Storage and Communication
F - Manufacturing
G - Wholesale and Retail Trade
H - Public Services and Statutory Authorities
I - Mining and Quarrying
J - Agriculture, Forestry and Fishery

Figure 2: Occupational Accident Statistics by Sector in 2020

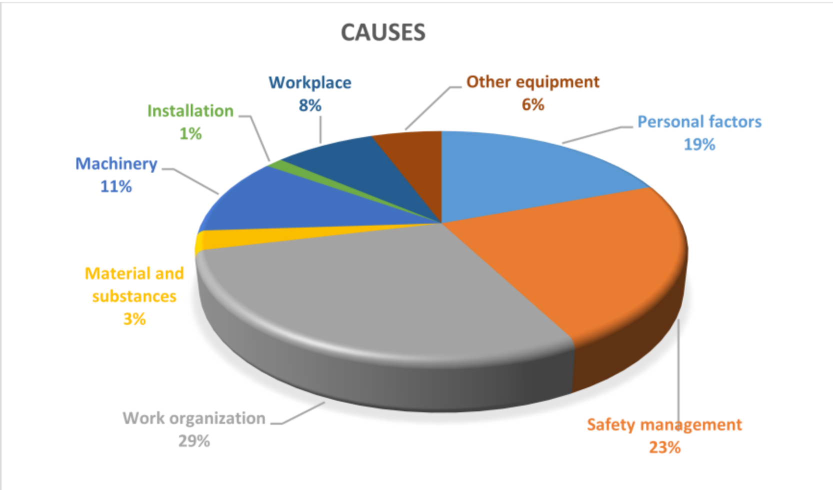


Figure 3: Distribution of the Accident Causes by Group

The rapid growth of the manufacturing industry, characterized by the influx of new workers and the adoption of novel technologies, machinery, and equipment, introduces new hazards to the workforce (Said et al., 2012). Additionally, the employment of new workers poses risks as they may not be acclimated to the workplace environment (Said et al., 2012). Accidents often occur due to improper work methods, inadequate training and communication, insufficient supervision, and deficiencies in qualifications and experience (Carillo et al., 2016). Figure 3 illustrates the causes of accidents in the Andalusian manufacturing sector investigated by the Labor Authority from 2004 to 2011. The pie chart highlights that work organization (29%) is the leading cause of accidents, followed by safety management (23%). Work organization factors encompass elements such as work methods, activity planning and execution, training, and equipment selection. Consequently, as an alternative approach to ensure effective

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004
supervision and management, automatic safety monitoring and supervision systems have begun to be implemented in manufacturing environments.

This paper's primary objective is to shed light on the suitability of employing IVS with deep learning methods in safety management, especially within the manufacturing sector. The paper is structured into two main sections: "Deep Learning in Safety Management" and "Deep Learning in Manufacturing Sector." Methods and outcomes from the state-of-art previous work will be discussed.

Deep Learning in Safety Management

Deep learning (DL) involves the utilization of multi-layer neural networks comprising multiple neurons to carry out various learning tasks, such as regression, classification, clustering, autoencoding, and more (Hatcher et al., 2018). The capacity to handle data in its raw format makes DL methods remarkably easy to adopt and implement, offering superior scalability compared to traditional machine-learning approaches (Hatcher et al., 2018). Conversely, conventional methods necessitate the transformation of input data from pixel values into a specific representation that models can comprehend and process, demanding considerably more effort than DL methods. Additionally, DL techniques excel in learning intricate patterns and features by leveraging multiple layers of representation (Hatcher et al., 2018). Its application has experienced significant growth and is now being integrated into real-time operations across various domains. Nevertheless, only a limited number of studies have explored the integration of DL into safety management alongside intelligent visual surveillance (IVS).

Object Detection

Choi et al (2019) introduced a method for recognizing perilous areas and evaluating the associated risks by employing a deep learning-based object detection approach. They utilized the YOLO model to train images containing floor openings, aerial lifts, fixed cranes, ladders, and scaffolds, as illustrated in Figure 4. These five hazardous zones were pinpointed within the shipbuilding industry. The algorithm demonstrated a noteworthy ability to identify these objects, achieving an average precision rate exceeding 90% and an average recall rate surpassing 83%.

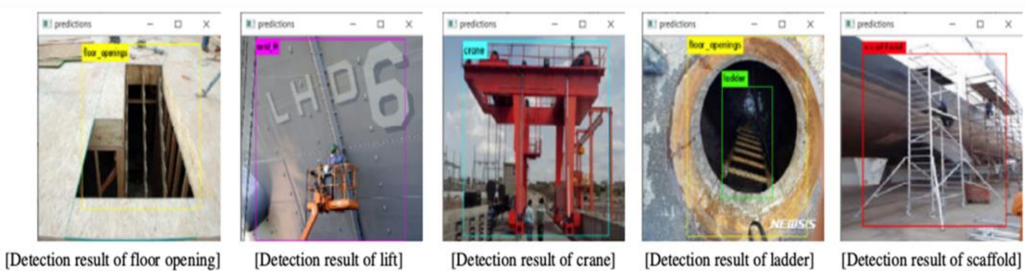


Figure 4: Detection Results of Hazardous Zone (Choi et al., 2019)

Fang et al (2018) introduced a method for identifying instances of non-hardhat usage in distant surveillance videos captured on construction sites. Their algorithm employed the Fast-RCNN technique to detect safety helmets. The study comprehensively examined various visual conditions within construction sites to assess how adaptable the methods were to the construction site environment. These visual conditions are visualized in Figure 5, and the algorithm proved its capability to identify instances of non-hardhat usage even in challenging

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004
conditions such as when there is a large visual range, which makes objects appear smaller, or when visibility is reduced due to misty rain or haze. The overall performance of the system achieved precision and recall rates exceeding 90%.

Nath et al. (2020) conducted a comparative analysis of three distinct approaches to verify the compliance of construction site workers with personal protective equipment (PPE) regulations. They employed the YOLO network to detect PPE, with differences in how the detection and classification were carried out: (a) worker, hat, and vest classification using machine learning, (b) simultaneous detection and classification using a single network, and (c) classification using CNN-based classifiers like VGG-16, ResNet-50, and Xception. Their proposal included the application of a Bayesian Framework to consolidate outputs from multiple classifiers. Sample results are depicted in Figure 6, where the classifier categorizes objects into worker (W), worker with hat (WH), worker with vest (WV), and worker with both hat and vest (WHV).

In a qualitative assessment, Approach-1 surpasses both Approach-2 and Approach-3 in terms of scalability, modularity, robustness, and adaptability. Furthermore, when it comes to processing speed, Approach-1 stands out as the fastest compared to the other methods. However, from a quantitative perspective, Approach-2 achieves the highest level of accuracy in detecting PPE attire, with a performance of 72.3% in mean average precision (mAP). Although Approach-1 and Approach-3 achieve mAP scores of 81.2% and 85.6%, respectively,

Figure 5: Detection Results of Non-Hardhat-Use in Various Visual Conditions (Fang et al., 2018)

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004

(a) Approach-v1 (b) Approach-v2 (c) Approach-v3

Figure 6: Detection Results of PPE at Construction Site in Three Different Approaches (Nath et al., 2020)

in the object detection component, the classifier component diminishes the mAP to 63.1% in Approach-2 and 67.9% in Approach-3. It's important to note that these methods were evaluated exclusively for hat and vest classes, whereas there are numerous other components and types of PPE available.

Su et al (2021) developed a data-driven Convolutional Neural Network (CNN) fire detection model aimed at enhancing fire safety practices on construction sites. Their study focused on training and testing a real-time construction fire detection (RFCD) model. The process begins with high-speed cameras capturing images, which are then transmitted to the central processing unit of the RFCD system. Subsequently, the image scene data are converted into a format recognizable by the machine within the fire detection module, allowing for scenario classification. Finally, the fire warning module conveys the results to the safety management personnel's receiving equipment (Su et al., 2021). This system can operate continuously for 24 hours throughout the entire lifecycle of a construction project, eliminating the need for managers and construction workers to undergo training. This research significantly improves fire safety management, especially in situations where traditional methods prove inefficient and the transmission of fire-related information to decision-makers is delayed. However, it's important to note that the proposed approach has limitations and is most effective in unobstructed scenes, as it becomes more challenging when buildings and objects obstruct the view at the construction site.

Object Recognition

Object recognition is essentially the subsequent step after object detection. Object detection involves identifying and localizing an object by drawing a bounding box around it. For instance, in face detection, the goal is to locate and delineate the face with a bounding box. Once the location information is obtained, the next step is face recognition, which entails identifying the individual within the bounding box. Face recognition has garnered substantial attention over the past three decades as a fundamental application in image analysis and pattern recognition (Adjabi et al., 2020).

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004
Wang et al (2017) introduced a method for real-world surveillance video face recognition using the VGG16 model. In the initial stage, a rough dataset is generated using the Haar face detector, and face tracking is accomplished using the KCF tracking method. Subsequently, VGG features are extracted from each face image, and the cosine distance between these images is calculated. If the cosine distance between two images falls below a specified threshold, denoted as T, the two images are connected by an edge (Wang et al., 2017). A graph is constructed by aggregating all these edges and images. Fine-tuning of the VGG face model is performed by replacing backpropagation with their collected dataset. The VGG model, after fine-tuning, achieves a recognition rate of 92.1%, whereas the VGG model without fine-tuning achieves a rate of 83.6%. Clearly, fine-tuned VGG outperforms the non-fine-tuned version. This fine-tuning process represents a fundamental technique in implementing pre-trained deep learning models, also known as transfer learning.

In a separate study, Taigman et al (2014) introduced DeepFace, a multistage approach that employs a generic 3D shape model for face alignment. The architecture of DeepFace is depicted in Figure 7. This network comprises over 120 million parameters within a nine-layer deep neural network, featuring locally connected layers without weight sharing, in contrast to standard convolutional layers. Training was conducted using the largest facial dataset, Labeled Face in the Wild (LFW), which includes over 4,000 identities and four million facial images. DeepFace was one of the pioneering works that achieved remarkably high accuracy on the LFW dataset using CNNs (Adjabi et al., 2020). This study catalyzed the research in face recognition, shifting the focus toward deep learning-based approaches. In just three years, accuracy rates witnessed significant improvements (Adjabi et al., 2020). Presently, face recognition systems perform well in relatively controlled settings. However, they encounter challenges in real-world surveillance systems, where variations in pose, lighting, facial expressions, and aging can substantially impede their performance in recognizing individuals in unconstrained scenarios (Taigman et al., 2014). Consequently, this field of study remains an active area of research.

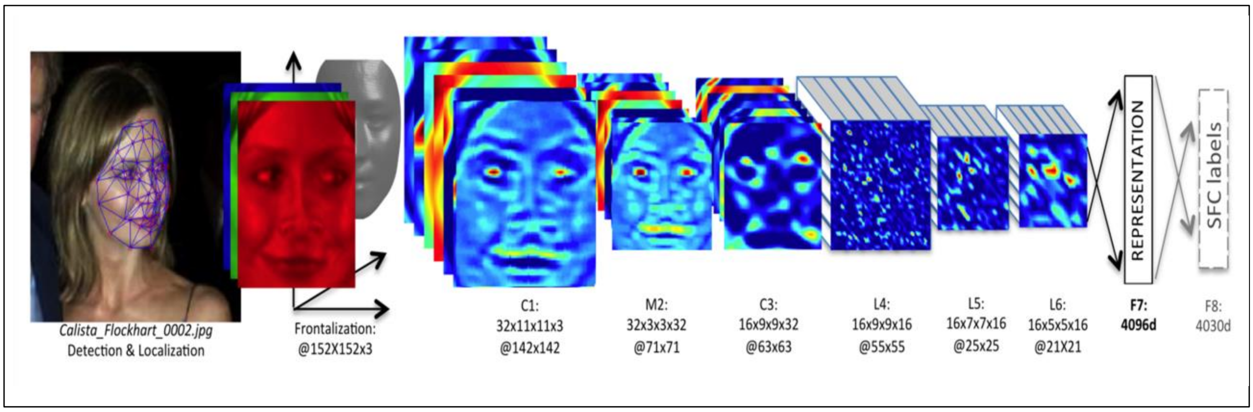


Figure 7: A Front-End of a Single Convolution-Pooling-Convolution Filtering on The Rectified Input, Followed by Three Locally Connected Layers And Two Fully Connected Layers. Colors Illustrate Feature Maps Produced at Each Layer (Taigman et al., 2014).

Object Re-identification

In video surveillance systems, the objective is to detect, recognize, and track individuals as they move within a single camera's view or even across multiple cameras in cases of extensive area monitoring. The process of person or object re-identification (ReID) involves determining

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004

the correct match of a given query person from a vast collection of images or videos obtained from non-overlapping cameras [28]. In scenarios involving wide areas with multiple non-overlapping cameras where the path individuals take between cameras is uncertain, appearance matching methods are essential for re-identifying individuals as they traverse the area (Bialkowski et al., 2012). In contrast, in single-camera setups, object tracking can be employed to re-identify individuals across consecutive frames (Bialkowski et al., 2012). Traditionally, most methods have focused on extracting appearance features such as colors, shapes, and textures. Nowadays, deep learning, particularly through Convolutional Neural Network (CNN) models, extracts features from objects, which has led to significant advancements in various computer vision applications, including activity recognition, video analysis, and medical data (Liu et al., 2016).

Object ReID plays a crucial role in numerous safety applications and commercial management systems. Khan and colleagues (2021) proposed a person re-identification (P-ReID) approach using deep encoders. Their framework comprises two main steps. Firstly, the input image is divided into upper and lower body segments, and these patches are then processed by the SqueezeNet for feature extraction. Subsequently, the features from both patches are fused into a single-dimensional vector with a size of 2000 dimensions and subjected to deep autoencoder-based feature learning. In the second step, the algorithm repeats the process from the first framework, and similarity scores are computed using cosine similarity, followed by the evaluation of various ranking scores. Visual results based on different scenarios for test samples are presented in Figure 8.

Figure 8. Visual Results of Test Samples: (A) The First Two Rows Represent The Probe And Reidentified Images Generated Over CUHK01 (B) The Third and Fourth Rows Show The Market1501 Query, And (C) Samples Outcome on Viper Database (Khan et al., 2021).

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004

In the field of transportation, Vehicle Re-Identification (Vehicle ReID) plays a vital role in intelligent transportation systems, urban computing, and intelligent surveillance. Its primary purpose is to swiftly identify, locate, and track target vehicles. Figueiredo and colleagues (2021) introduced the pioneering Motocycles Re-Identification (MoRe) dataset, which is the first extensive dataset designed for re-identifying motorcycles. This choice is significant because motorcycles are commonly encountered objects that have relevance in transportation, traffic management, and law enforcement.

In this proposed approach, we utilized a pretrained RetinaNet model from MS-COCO to detect both motorbikes and motorcyclists. Subsequently, the ResNet50 model is applied to perform the re-identification of motorcycles, encompassing both motorbikes and riders. Additionally, several techniques were incorporated and assessed within a robust baseline model, including label smoothing, warmup learning rates, last stride adjustments, BNNeck, and center loss. These techniques demonstrated noteworthy enhancements in various aspects of the model's performance. Furthermore, the study considered essential factors such as input image shape and the aspect ratio of input images, which were found to have a significant impact on the method's performance.

Deep Learning in Manufacturing Sector

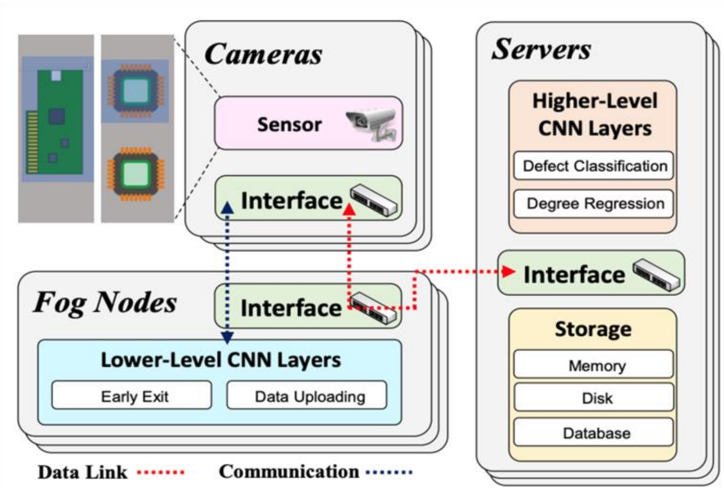


Figure 9: Overview of The Defect Detection System (Li et al., 2018).

Amidst the backdrop of the Fourth Industrial Revolution (IR4.0), smart manufacturing is rapidly emerging, driven by the integration of innovative technologies like the Internet of Things (IoT) sensors, Artificial Intelligence (AI), and Big Data. In this context, Li and their team (2018) proposed a Deep Learning (DL)-based approach for detecting product defects. The system's overarching architecture is depicted in Figure 9. Initially, images are captured and forwarded to local fog nodes to extract lower-level features and, whenever possible, generate rapid inference results to enhance efficiency and reduce data traffic (Li et al., 2018). If the fog nodes cannot make definitive decisions on their own, intermediate values from the CNN model inference are uploaded to remote cloud servers for further processing. These central servers possess the capability to complete the entire CNN inference, manage classification and regression tasks, and ultimately provide the final results. The deep model comprises four convolutional layers, two fully connected layers, and two end modules at the top, which consist of one classifier and one regressor. This work effectively adapts the CNN model to fog

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004
computing, enabling it to handle substantial data volumes. According to the study, their proposed approach proved to be highly efficient, performing nearly twice as fast as local computation, thanks to the use of fog computing.

In a different context, Shao et al (2017) introduced a deep belief network (DBN) designed to extract features from the frequency distribution of vibration signals, facilitating the characterization of the operating status of induction motors. A reliable diagnosis system for faulty induction motors relies on vibration signals as input, as they typically contain valuable information about the motor's operational status. The study discusses the comparison between DBN using time-domain signals and FFT-DBN. It was observed that FFT-DBN exhibited greater stability compared to non-FFT-DBN, primarily because DBN struggles to capture the correlations among the inputs, which have a significant impact on the classification task. Effective diagnosis of faulty induction motors is critical for enhancing reliability and reducing operational and maintenance costs in the manufacturing sector, thereby contributing to accident prevention.

Furthermore, Francis and their team (2019) developed a DL model geared towards predicting distortion by considering local heat transfer in Laser-Based Additive Manufacturing (LBAM) processes. Thermal images are employed as inputs to the deep learning model, as many quality issues in LBAM, such as geometric inaccuracies and distortion, can be attributed to the intense heating (thermal expansion) and subsequent cooling (thermal contraction) processes, as illustrated in Figure 10. The CNN in this context consists of two convolutional layers and two pooling layers before being flattened. This work aligns with the objectives of IR4.0, which aim to implement numerous sensors for continuous process control in the manufacturing sector.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004

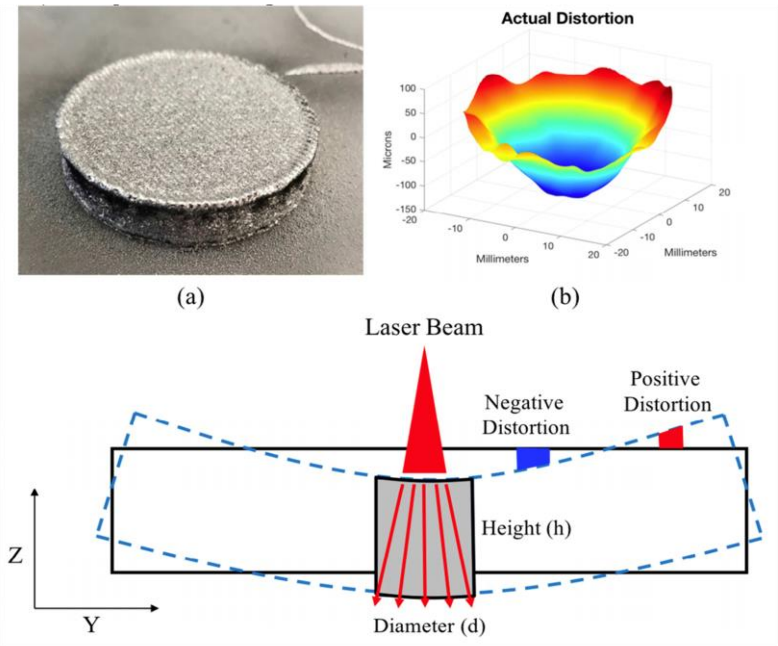


Fig. 10. (a) The Disk Fabricated for The Case Study, (B) The Distortion of The Disk is Given. The Shown Figure is a Smoothed Estimate of The Noisy Point Cloud Measurement of Distortion, and (c) Visualizes The Side View of The Part During Fabrication. The Black Line Indicates The Original Part Specification, While The Blue Dashed Line Shows The Distortion Part. As The Laser Beam Passes Over The Part, The Heat Transfer, Shown As Red Arrows, Occurs In A Local Area Shown In Gray (Francis et al., 2019).

Discussion on Challenges

The primary objective of this study is to ensure the sustained growth of the manufacturing industry, which hinges on maintaining a stable and highly secure environment. Despite the integration of technology in certain process designs, system disruptions due to human errors and hardware failures, as highlighted in our previous analysis, remain a challenge. Organizations must actively strive to reduce human errors and enhance the system's error tolerance to prevent and mitigate accidents and incidents. This can be achieved by adopting an anthropocentric approach, where cyber systems or technology serve as assistants to humans rather than fully controlling or dictating human actions. The incorporation of AI technology should complement human observation and intervention.

However, we have discussed the employment of deep learning in safety monitoring especially in manufacturing sector. There are several key challenges of deep learning integrated with IoT.

Data Quality and Quantity

Deep learning (DL) models require large volumes of high-quality data for training. In many manufacturing environments, obtaining such data can be challenging due to issues like data silos, data inconsistency, or insufficient data.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004

Data Privacy and Security

Manufacturing data often contains sensitive information related to processes and products. Ensuring data privacy and security while sharing and storing data for deep learning is a significant challenge.

Complexity

Manufacturing processes can be highly complex, involving various interconnected systems. Developing deep learning models that can understand and optimize these complex processes is a substantial challenge. The other issue, pertains to the insufficient availability of a substantial number of training samples within industrial settings, leading to reduced model accuracy and efficiency due to overfitting (Khalil et al., 2021). To address this issue, a solution involves the implementation of a Tensor-Train Deep Compression (TTDC) model, specifically designed for efficient feature learning from industrial data (Roopei et al., 2017). TTDC effectively compresses numerous factors within the deep learning model, thereby enhancing its processing speed and mitigating overfitting concerns.

Data Labelling

IoT generates vast amounts of data, making it challenging to label data at scale. Manually annotating all the data for DL models can be time-consuming and expensive, especially for specialized domains, can be costly due to the need for domain experts such as medical field. This cost can hinder the deployment of DL solutions in IoT. Labelling errors or noise in IoT data can adversely affect DL model performance. Ensuring data quality is crucial for accurate model training.

Conclusion

The paper conducts a comprehensive survey of intelligent video surveillance using deep learning approaches for safety management within the manufacturing sector, encompassing a wide range of applications. The review commences with a broad examination of video surveillance for safety management and subsequently delves into the implementation of deep learning in the manufacturing domain. The survey specifically focuses on three key components of the deep learning approach: object detection, object recognition, and object re-identification. It is evident that deep learning-based surveillance systems are extensively utilized across various sectors. However, it is noteworthy that safety management, particularly within the manufacturing sector, has received relatively limited attention in the realm of scientific research. The issues highlighted in this survey serve as a guiding framework for researchers in their future endeavors, shaping the direction of future research towards delivering more effective and efficient solutions for intelligent visual surveillance through deep learning.

Acknowledgment

This work was carried out with the financial support from the Universiti Kebangsaan Malaysia under the research grant GP-K010187.

References

Adjabi, I., Ouahabi, A., Benzaoui, A., & Taleb-Ahmed, A. (2020). Past, present, and future of face recognition: A review. Electronics, 9(8), 1188.

Boholm, M. (2019). How do Swedish Government agencies define risk? Journal of Risk Research, 22, 717-734.

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004
Bialkowski, A., Denman, S., Sridharan, S., Fookes, C., & Lucey, P. (2012). A database for person re-identification in multi-camera surveillance networks. In 2012 International Conference on Digital Image Computing Techniques and Applications (DICTA) (pp. 1-8). IEEE.

Carrillo-Castrillo, J. A., Rubio-Romero, J. C., Onieva, L., & López-Arquillos, A. (2016). The causes of severe accidents in the Andalusian manufacturing sector: the role of human factors in official accident investigations. Human Factors and Ergonomics in Manufacturing & Service Industries, 26(1), 68-83.

Choi, Y., Park, J. H., & Jang, B. (2019). A risk estimation approach based on deep learning in shipbuilding industry. In 2019 International Conference on Information and Communication Technology Convergence (ICTC) (pp. 1438-1441). IEEE.

Chumuang, N., Ketcham, M., & Yingthawornsuk, T. (2018). CCTV based surveillance system for railway station security. In 2018 International Conference on Digital Arts, Media and Technology (ICDAMT) (pp. 7-12). IEEE.

Cocca, P., Marciano, F., & Alberti, M. (2016). Video surveillance systems to enhance occupational safety: A case study. Safety Science, 84, 140-148.

Department of Occupational Safety and Health, Malaysia. (2023). Occupational Accident and Statistics. Retrieved from https://www.dosh.gov.my/index.php/statistic-v/occupational-accident-statistics

Fanchamps, M. H. J., Horemans, H. L. D., Ribbers, G. M., Stam, H. J., & Bussmann, J. B. J. (2018). The accuracy of the detection of body postures and movements using a physical activity monitor in people after a stroke. Sensors, 18(7), 2167.

Fang, Q., Li, H., Luo, X., Ding, L., Luo, H., Rose, T. M., & An, W. (2018). Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. Automation in Construction, 85, 1-9.

Figueiredo, A., Brayan, J., Reis, R. O., Prates, R., & Schwartz, W. R. (2021). MoRe: A Large-Scale Motorcycle Re-Identification Dataset. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 4034-4043).

Francis, J., & Bian, L. (2019). Deep learning for distortion prediction in laser-based additive manufacturing using big data. Manufacturing Letters, 20, 10-14.

Gul, M. A., Yousaf, M. H., Nawaz, S., Rehman, Z. U., & Kim, H. (2020). Patient Monitoring by Abnormal Human Activity Recognition Based on CNN Architecture. Electronics, 9(12), 1993.

Hatcher, W. G., & Yu, W. (2018). A survey of deep learning: platforms, applications and emerging research trends. IEEE Access, 6, 24411-24432.

Kayas, O. G., Hines, T., McLean, R., & Wright, G. H. (2019). Resisting government rendered surveillance in a local authority. Public Management Review, 21(8), 1170-1190.

Khalil, R. A., Saeed, N., Masood, M., Fard, Y. M., Alouini, M. S., & Al-Naffouri, T. Y. (2021). Deep learning in the industrial internet of things: Potentials, challenges, and emerging applications. IEEE Internet of Things Journal, 8(14), 11016-11040.

Khan, S. U., Hussain, T., Ullah, A., & Baik, S. W. (2021). Deep-ReID: deep features and autoencoder assisted image patching strategy for person re-identification in smart cities surveillance. Multimedia Tools and Applications, 1-22.

Li, L., Ota, K., & Dong, M. (2018). Deep learning for smart industry: Efficient manufacture inspection system with fog computing. IEEE Transactions on Industrial Informatics, 14(10), 4665-4673.

Liu, X., Liu, W., Ma, H., & Fu, H. (2016). Large-scale vehicle re-identification in urban surveillance videos. In 2016 IEEE International Conference on Multimedia and Expo (ICME) (pp. 1-6). IEEE.

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 8 Issue 33 (December 2023) PP. 42-56
DOI: 10.35631/JISTM.833004
Maior, C. B. S., Santana, J. M., Nascimento, L. M., , Macedo, J. B., Moura, M. C., Isis, D. L., & Droguett, E. L. (2018). Personal protective equipment detection in industrial facilities using camera video streaming. In Safety and Reliability - Safe Societies in a Changing World, Haugen et al (Eds), Taylor & Francis Group (pp. 2863-2868).

McCall, M. K., Skutsch, M. M., & Honey-Roses, J. (2021). Surveillance in the COVID-19 Normal: Tracking, Tracing, and Snooping–Trade-Offs in Safety and Autonomy in the E-City. International Journal of E-Planning Research (IJEPR), 10(2), 27-44.

Mudgal, M., Punj, D., & Pillai, A. (2021). Suspicious Action Detection in Intelligent Surveillance System Using Action Attribute Modelling. Journal of Web Engineering, 129-146.

Nath, N. D., Behzadan, A. H., & Paal, S. G. (2020). Deep learning for site safety: Real-time detection of personal protective equipment. Automation in Construction, 112, 103085.

Patel, S., Patel, N., Deshpande, S., & Siddiqui, A. (2021). Ship Intrusion Detection using Custom Object Detection System with YOLO Algorithm. International Research Journal of Engineering and Technology (IRJET), 8(1), 669-677.

Pramanik, A., Sarkar, S., & Maiti, J. (2021). A real-time video surveillance system for traffic pre-events detection. Accident Analysis & Prevention, 154, 106019.

Resheed, A. (2018). An Integrated Framework for CCTV Infrastructures Deployment in KSA: Towards an Automated Surveillance. Doctoral thesis, Northumbria University.

Roopaei, M., Rad, P., & Jamshidi, M. (2017). Deep Learning Control for Complex and Large Scale Cloud Systems. Intelligent Automation & Soft Computing, 23(3), 389-391.

Said, S. M., Halim, Z. A., & Said, F. (2012). Workplace injuries in Malaysian manufacturing industries. Journal of Occupational Safety and Health, 9(1), 21-32.

Saufi, S. R., Ahmad, Z. A., Leong, M. S., & Lim, M. H. (2019). Challenges and opportunities of deep learning models for machinery fault detection and diagnosis: A review. IEEE Access, 7, 122644-122662.

Shao, S. Y., Sun, W. J., Yan, R. Q., Wang, P., & Gao, R. X. (2017). A deep learning approach for fault diagnosis of induction motors in manufacturing. Chinese Journal of Mechanical Engineering, 30(6), 1347-1356.

Sreenu, G., & Durai, M. S. (2019). Intelligent video surveillance: a review through deep learning techniques for crowd analysis. Journal of Big Data, 6(1), 1-27.

Su, Y., Mao, C., Jiang, R., Liu, G., & Wang, J. (2021). Data-Driven Fire Safety Management at Building Construction Sites: Leveraging CNN. Journal of Management in Engineering, 37(2), 04020108.

Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1701-1708).

Vallathan, G., John, A., Thirumalai, C., Mohan, S., & Lin, J. C. (2021). Suspicious activity detection using deep learning in secure assisted living IoT environments. The Journal of Supercomputing, 77(4), 3242-3260.

Wang, Y., Bao, T., Ding, C., & Zhu, M. (2017). Face recognition in real-world surveillance videos with deep learning method. In 2017 2nd International Conference on Image, Vision and Computing (ICIVC) (pp. 239-243). IEEE.