University of Salford MANCHESTER

A deep learning approach towards railway safety risk assessment

Alawad, H, Kaewunruen, S and An, M

http://dx.doi.org/10.1109/ACCESS.2020.2997946

Title	A deep learning approach towards railway safety risk assessment
Authors	Alawad, H, Kaewunruen, S and An, M
Type	Article
URL	This version is available at: http://usir.salford.ac.uk/id/eprint/57175/
Published Date	2020

A Deep Learning Approach Towards Railway Safety Risk Assessment

Hamad Alawad[1], Sakdirat Kaewunruen[1], and Min An[2]

[1]The Birmingham Centre for Railway Research and Education, University of Birmingham, Birmingham B15 2TT, UK
[2]The School of Science, Engineering and Environment, The University of Salford, Manchester, M5 4WT, UK

Corresponding author: Sakdirat Kaewunruen (s.kaewunruen@bham.ac.uk).

ABSTRACT Railway stations are essential aspects of railway systems, and they play a vital role in public daily life. Various types of AI technology have been utilised in many fields to ensure the safety of people and their assets. In this paper, we propose a novel framework that uses computer vision and pattern recognition to perform risk management in railway systems in which a convolutional neural network (CNN) is applied as a supervised machine learning model to identify risks. However, risk management in railway stations is challenging because stations feature dynamic and complex conditions. Despite extensive efforts by industry associations and researchers to reduce the number of accidents and injuries in this field, such incidents still occur. The proposed model offers a beneficial method for obtaining more accurate motion data, and it detects adverse conditions as soon as possible by capturing fall, slip and trip (FST) events in the stations that represent high-risk outcomes. The framework of the presented method is generalisable to a wide range of locations and to additional types of risks.

INDEX TERMS Artificial intelligence, Convolutional neural network, Deep learning, Railway station, Risk management

I. INTRODUCTION

Railway station environments are dynamic, and this dynamicity varies according to size and location. A variety of passengers transit the station, including families, old and disabled individuals and groups. Some stations are crowded at peak times because of the limited space, and increases in demand due to operational delays, design or layout deficiencies or management shortages can increase the risk of fall, slip and trip (FST) events.

FSTs are a leading cause of injury. In particular, falls due to slipping are statistically the main cause of accidents on crossways in built environments and railway stations [1]. The consequences of FSTs are not limited to the individual who suffered the accident, who may be seriously injured; FSTs can also affect railway operations, causing delay and disturbing the flow of people. Platforms which offer access to trains and escalators are hazards that form hot spots for FSTs. According to the RSSB Annual Report on Public Safety (2015/2016),

over the last five years, the highest percentage of injuries from slips, trips and falls in stations occurred on stairs (38%), with platforms holding the second most likely spot (27%) [2]. Generally, the magnitude of falls worldwide rises with age: it has been reported that the proportion of 32–42% elderly adults (aged 70 years or older) fall each year from 5 to 7 times[3].Some factors previously presented as the most crucial in FST events in the station include intoxication, security, hurrying, station design, staff skills and training [4]. Other challenges, such as weather conditions, congestion, cultural differences, insufficient maintenance and unwanted events, may cause panic and FSTs [5], [6]. Much of the unsafe behaviour exhibited by passengers, employees and the public can be described via the theory of behaviour-based safety (BBS), which has been demonstrated to be an effective tool for promoting safety [7]. The BBS includes three steps: observation, feedback and training. However, in railway

stations, these observations are based on humans; thus, they also involve human error. Moreover, covering all points in the organisation is costly, time-consuming, or impossible [8], [9].

Worldwide, FSTs are a severe problem that lead to substantial numbers of injuries and have endemic societal economic consequences that affect people of all ages. That is, despite logical, well-conceived attempts to diminish the number of casualties, these approaches have had mixed success [10]. FSTs are classified as causes of unintentional injury in many activities, both occupational and leisure-related, and their causes include loss of balance, which may result in falls to the ground or to lower levels. The factors to be considered include footwear, flooring conditions and visibility level as well as other external factors, such as crowding [11], [12]. Moreover, station design and layout factors , including corridors and entrances, egress routes and escalators, play an important role in safety and security and in preventing FSTs; however, the designs of some older stations include narrow areas [13]. Large numbers of people in these limited areas may lead to crowding over a wide-range of areas, such as in railway stations or entertainment venues (e.g., sports matches or music festivals), which raises the risk level of FSTs. The flow of people may also be affected by obstructions, which can result in pushing, falling or, in the worst case, trampling, which may increase the number of incidents. The FSTs in crowded situations can have serious consequences; historically, many people have died or suffered serious injury during events such as religious pilgrimages. Such risks increase when the railway industry's growth level is inadequate to serve the market demand for train travel. In addition, such risks also increase for older passengers, for travellers carrying large luggage items and when intoxication is considered [14]. FSTs are associated with many aspects of accidents, such as human-to-station environment interactions, including infrastructure and trains. Many causes are attributable to such risks and they lead to many different consequences (see Fig. 1). Moreover, the nature and patterns of FSTs and their active control measures require more research [12].

Currently in railway stations, detecting such risks relies on CCTV or staff observations; however, this approach has the potential for human error and may not result in a timely response, which can exacerbate the consequences. Furthermore, accurate station area detection includes platforms, escalators and tunnels; the images can include the full range of the station and thus provide the potential for timely responses. Technological growth has helped to extend and improve protection, especially CCTV systems. In recent years, automated video surveillance has enhanced public safety awareness and led to innovative research in a wide range of fields, including disaster management, crime prevention and security, assistance for people with disabilities, productivity enhancements and monitoring critical infrastructure [15].

In railway stations, CCTV and analogue cameras aid in accurately detecting station areas, including platforms, escalators and tunnels. However, the human behind the screens is the core of the system, which leads to possible human errors. In practice, however, a greater number of cameras in such areas leads to a loss of the monitors' ability to gain an overview of events in real-time. In fact, in accidents, the CCTV function primarily as evidence; thus, the process involves working with historical records of events more than with real-time event detection. Artificial intelligence (computer vision-based) techniques have been suggested as a possible solution to these issues; at minimum, they could function as an important assisting element to overcome some of the limitations of the conventional methods of risk management at stations and improve the safety system. This paper proposes a monitoring method that uses computer vision to automatically and rapidly identify risks in stations by recognising unsafe actions, providing support for decision-makers in real-time and reducing the potential consequences of unwanted events.

A vision-based approach can be considered the most suitable for crowded critical locations such as railway stations. Computer vision technology has demonstrated its potential for practical, cost-effective, rapid visual data collection, and vision-based approaches have been adopted in many fields, such as construction, safety and quality management [16]–[18] and productivity management [19]–[21]. The railway industry has already seen benefits from such methods, including railway track-gauge irregularities [22], railway maintenance [23] and trespassing detection [24]. Similarly, computer-based image recognition has been applied to detect and recognise railway infrastructure and changes in the surrounding environment [25]. Deep learning techniques have also been proposed to evaluate rail quality using track geometry [26]. Such research focuses on how to utilise technology such as the convolutional neural network (CNN) to analyse big data collected by railway systems to build risk-recognition frameworks—in the case of FSTs—risk in the railway stations, Fig. 2. The railway industry creates big data that have potential value for improving the system. This massive data can be utilised to provide suitable solutions for safety and security risks. The goal is to tackle the changing risks that face a sector via image data. The data can cover a wide variety of aspects and take many forms, such as spatial-temporal data, videos or images and data fusion. The data used for monitoring can be collected at fixed points or be installed on moving trains or other vehicles, such as drones [27], [28]. Moreover, these data-gathering systems and their configuration can integrate with the Internet of Things (IoT), which is a framework suitable for big data technology, smart stations, smart cities and smart maintenance [29]–[33].

This remainder of this paper is organised as follows: Section II reviews the related works. Section III provides background information about deep learning and risk management in the railway industry. Section IV presents the concepts of using

deep learning for risk management decision making in railway stations. Section V presents the framework of the case study model. Section VI discusses the application of the CNN case study model in railway stations, Section VII provides a data analysis, and finally, a discussion and conclusions are given in Section VIII.
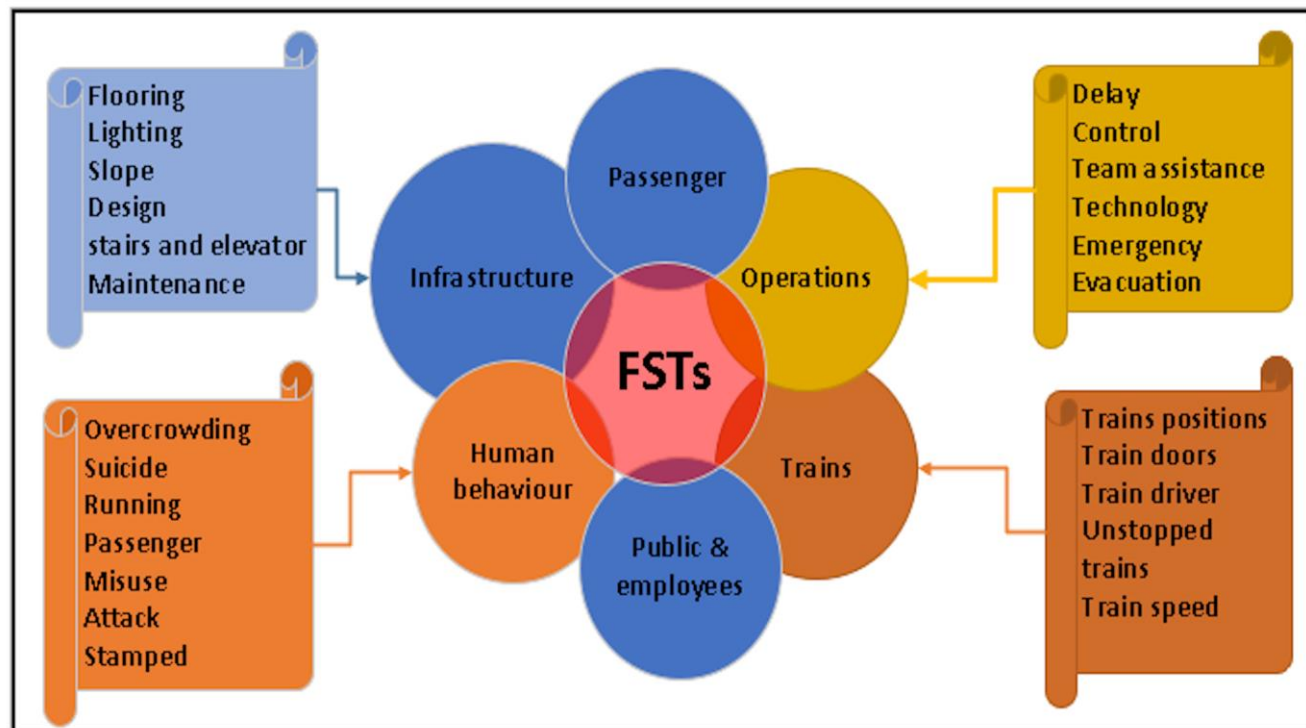

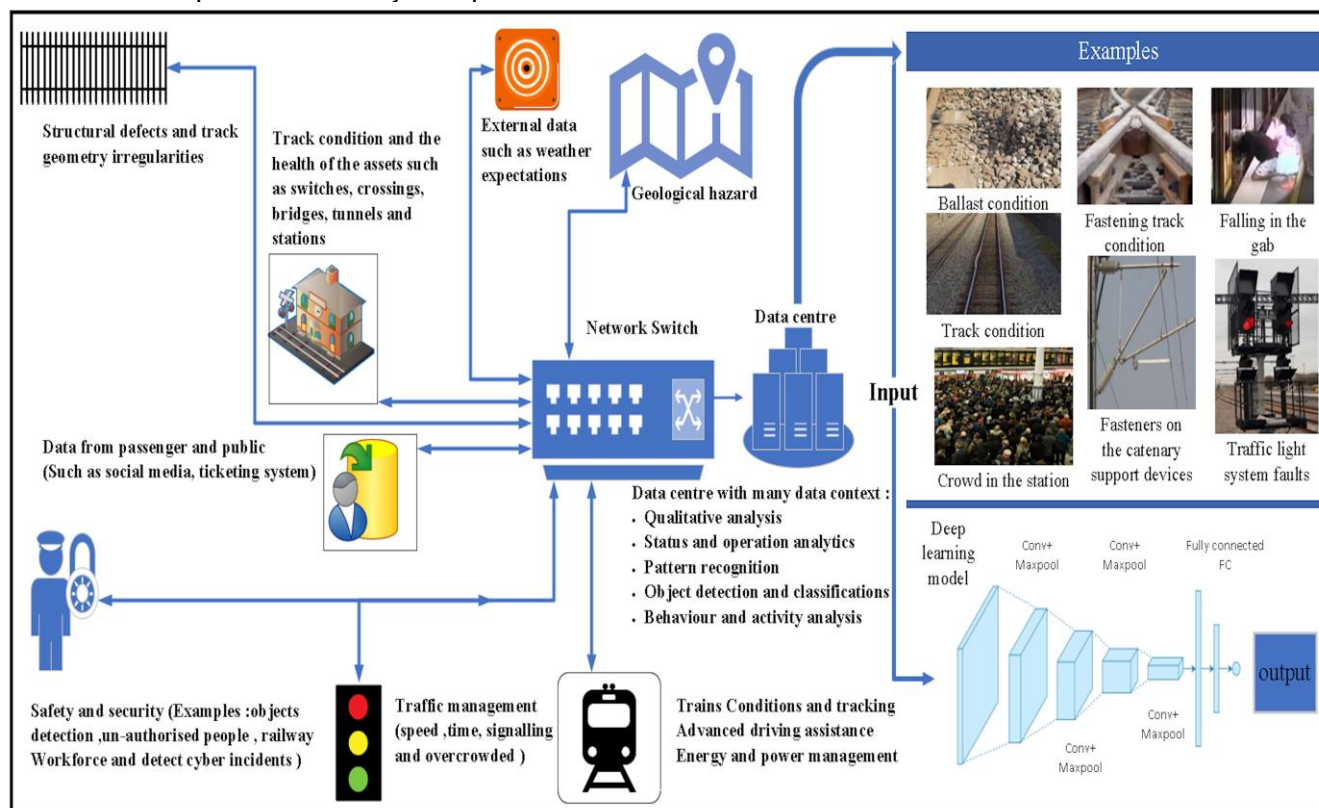**FIGURE 1.** The overlap between FSTs and system aspects.

FIGURE 2. The framework for applying a CNN to big data in the railway system to maintain risks.

II. RELATED WORKS

In this section, we attempt to assign the previous works to various subsections; however, their topics are scattered across many fields and include a variety of perspectives. CNNs have been widely applied in a wide range of visual computing applications, including signal processing [34], [35], speech recognition [36], medical imaging [37]–[42], object detection [43]-[47], face recognition [48]–[51], robot control [52], [53], autonomous driving (AD) and control [53]–[55] crash detection, risk estimation and traffic monitoring [56], [57]. Some models have even been implemented on mobile devices, such as Google's FaceNet [58] and Facebook's DeepFace [59], which are used for face recognition [60], [61]. Other studies have different perspectives, such as energy efficiency and data availability [62]–[64] and deep learning technologies for civil engineering applications, infrastructure monitoring and pedestrian detection and tracking [52], [65]–[68]. From an occupational safety perspective at a steel plant, images and deep learning have been utilised to detect oil spills [69] and to augment safety in the construction industry [70]. Furthermore, the method used to detect and track humans underwater [71] has also been applied to the automatic detection of unsafe actions in on-site videos [8] and for transport security using X-ray security images [72].

While the main areas initially involved computer science and related technologies, researchers have been applying deep learning techniques in their own fields. In the railway industry, the main concerns of this research are railway operations and safety (risk management). Features generated from CCTV images or other cameras in stations are fed into deep learning models so that they can learn from passengers' actions over a period of time. The goal is to train the CNN to automatically extract feature sequences that represent unsafe acts from videos, detect the presence of such sequences, and then initiate actions to mitigate the possible risks. Depth sensors such as Kinect™ or multiple cameras have been used to detect and record unsafe actions by extracting 3D skeletal models of humans [73]. Additionally, machine learning techniques combined with various processing methods have been applied [74]–[77]. The studied technologies include multisensor fusion-based approaches [78], accelerometer-based approaches [79], [80], smartphone-based approaches [81], [82], vision-based approaches [83], [84] and systems based on video data. Such systems can assist in detecting falls by monitored individuals at their homes [85]. Moreover, a previous study aimed at protecting and detecting falls showed that several major categories of sensing equipment have been used (see Fig. 3) [76], [84], [86]–[94].
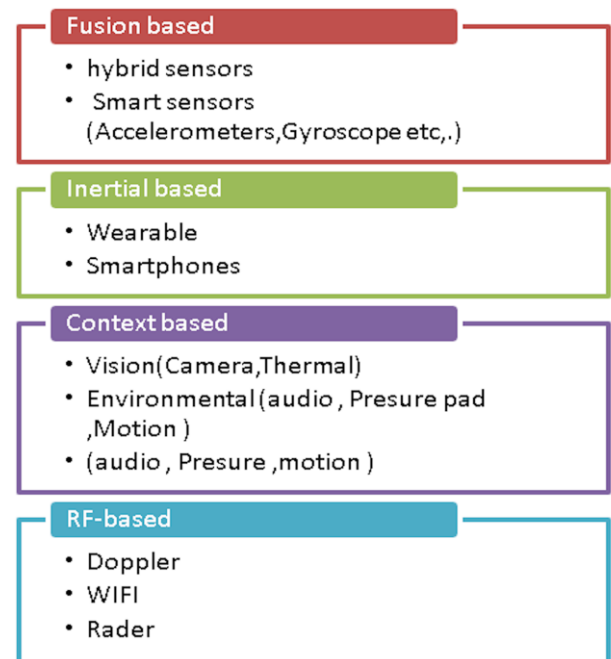

FIGURE 3. Fall detection and prevention based on sensor class perspectives.

The next subsection presents a review of the previous works in some fields related to this study, which involves detecting proposed risks in the study framework in a railway station by applying a CNN.

A. RELATED WORKS IN THE CONSTRUCTION FIELD

In construction and other fields, unsafe human behaviour is an important root cause of accidents [95], [96]. To identify common unsafe actions, stereo cameras were used to collect motion data and construct a 3D skeleton model; then, pattern recognition was applied to manage worker safety in the construction field [8] and to detect problems occurring on the site. Several defect management systems based on image matching have been suggested [97]. For less operational constraints, two smartphones have been used as stereo cameras to acquire motion data and extract 3D human skeletons to track people working in construction fields [98]. Real-time machine learning models with CNN frameworks have been proposed to detect whether workers are wearing safety equipment, such as hats and vests, from images/videos [99] and to detect ground objects [100]. CNNs have also been used to detect safety guardrails [101], objects on roof construction sites [102], workers who fail to wear hard hats [103], [104], falls from heights [105], [106], to maintain safe distances among objects for safety to prevent accidents [107] and unsafe behaviours [73]. Additionally, to estimate risk and reduce accidents, deep learning has been recommended in the shipbuilding Industry [108] and for ship bridge-collision assessment [109]. CNNs have been utilised for automated

detection of employees near heavy equipment at construction locations [110], detection of construction vehicles [111], [112] and recognition of structural damage [65]. In fact, for advanced safety performance, computer vision combined with deep learning has been recommended because such approaches can automatically classify unsafe behaviour and conditions on construction sites [70].

B. RELATED WORKS IN CRACK DETECTION

Crack detection has been classified in previous studies into two general method types: image-based crack detection and crack detection based on machine learning. An image-based crack detection method was suggested to automate crack detection for safety and cost-effective bridge maintenance [113]. Additionally, the authors proposed automating the processes of bridge monitoring and maintenance for safe transportation infrastructure and compared the effectiveness of four crack-detection algorithms (wavelet, Fourier transforms, Sobel, and Canny [114]) for detecting healthy concrete surfaces [115], bridge damage [116] and corrosion detection [117]. Moreover, with the goal of automating concrete bridge decks inspections, a principal component analysis (PCA) algorithm was applied to mitigate the dimensionality problem of feature vectors to extract significant crack features from a database of bridge images using a random forest algorithm [122] to perform crack detection on 3D asphalt surfaces [123]. Due to their high performance and promising results, convolutional neural networks (CNNs) have been utilised in visual computing in many studies in the field [124]–[126] and for floor area detection [127].

Due to their high performance and promising results, convolutional neural has led to extreme weather, while demand causes the industry to raise capacity and increase the number of trains in the system. Nevertheless, learning machine algorithms can estimate the exact abnormalities by monitoring rail tracks [27] to perform risk assessment of rail failure [128], diagnose track circuit faults [129] provide early and precise detection methods that are essential for avoiding risks [130] and provide information for decision support [131]. It has been shown that video camera inspection is a flexible, effective and automatic method for monitoring rail tracks. Running rolling stock can provide high-resolution images from different angles regarding their surroundings, including tracks and other assets. This data enhances the machine learning and enables high-performance predictions of abnormal changes or unwanted events [132], [133]. Moreover, the use of vision allows for more frequent infrastructure inspections and reduces human errors [134], helping to avoid maintenance train collisions [135] and monitoring to ensure passenger safety at stations [136]. Using a robot for railway tunnel detection reduces worker risk and improves the detection efficiency [137], [138]. Additionally, computer vision has been analysed for use in autonomous emergency train stops [139].

[118]. In addition, automatic concrete crack detection in tunnels using deep fully convolutional networks was proposed in [66]. To achieve automated detection and reduce the computational cost of detecting large concrete surface cracks, a method by percolation-based image processing was proposed in [119]. Tunnel crack features extracted based on detecting pixel intensity were classified by a support vector machine SVM algorithm to determine whether cracks were present in pre-processed images [120]. For safety inspection and structure health and reliability, an automated method based on a backpropagation neural network (BNN) was developed for crack detection [121]. For road crack detection, to deal with crack intensity inhomogeneity by capturing and utilising some unique crack characteristics, an automated method was suggested that extracts crack features based on a discriminative integral channel and then classifies the features

C. RELATED WORKS IN RAILWAY SYSTEMS

Technology such as computer vision will play an essential role in railway system networks and provide effective methods to solve various problems. The vast distances and long tracks in many areas of the world and the growth of complexity pose challenges to maintenance and in fulfilling safety, security and quality; in addition, there are cost restrictions, time-consumption and reliability issues.

Deep learning methods have been suggested for addressing many obstacles in the railway industry, such as poor or missing data; such methods are expected to improve prediction accuracy, optimise timing, reveal the types of maintenance that should be performed to rail infrastructure [33] and to perform object detection for railway traffic [140].

Of the many applications that have been applied to CNN, in this subsection, we present those that are specifically related to railways. Such studies have been widely reported in the recent literature and use many data sources; they cover management, maintenance, safety and operations [141]. Image-processing approaches for implementing automatic detection have been suggested for monitoring railway infrastructure [128], rail track maintenance [133], railway track inspections and train component inspections [142]–[152] such as the rolling bearings of trains [153].

CNNs have also been utilised to perform railcar safety inspection [154], determine the area of the rails ahead [155] detecting objects ahead [156], detect multiple catenary systems and support components [157]–[159], tracking joints [160] and detecting track defects [161], [162].

Sydney trains conducted condition monitoring for inspections and prevention of overhead wiring teardowns using laser and computer vision technologies [163]. Similarly, deep learning has been implemented to conduct traffic signal detection [164], [165], predict train delays [166], detect rail fastener defects and ballast history [167]–[169], detect cracks in and the shape and location of bolts [170], inspect railway ties [134], predict safety risks in communication-based train

control systems (CBTCs) [171] and to perform subgrade status inspections [172].

A CNN can be used to estimate crowd density at railway stations [173],to detect intrusions in track areas, such as pedestrians or large livestock via images captured in railway areas [174], to monitor railway construction [152] and for intrusion detection at railway crossings [175]. From the security side, the method been used for detecting violent crowd flows [176], protect the critical infrastructure [177], and identifying tools wielding by attackers such as knives, guns and Explosives [178].

A railway system contains a wealth of data, and visual processing technology can play an essential role in the industry's future. The most up-to-date applications were reviewed in [179].

III. DEEP LEARNING AND RISK MANAGEMENT

As one type of machine learning in AI, deep learning (DL) has been suggested as a method for risk management in railway stations. Accordingly, in this paper, we address some risks by utilising vision data from many points in the system, including both still frames and motion video. Currently, face recognition plays an important role in computer vision and has many applications, such as in autonomous vehicles, human-computer interactions, video surveillance, robotics, health care, medical imaging and homecare technology. Improvements in IT have enabled vision sensors to be installed in railway environments. For example, CCTV cameras widely used and rely on numerous cameras sensors; these cameras are intended to avoid and manage safety and accidents in railway environments.

In this study, we explore DL by utilising a convolutional neural network (CNN) to detect passenger falls. FTSs are common accidents in stations; their causes are sometimes related to human factors such as people running or to factors such as damage to floors (wet or muddy conditions) or a lack of lighting or poor steps design. FTS risks are correlated with other risks, such as overcrowding or emergency evacuation. In some cases, passenger falls can lead to overcrowding and panic; passengers can fall into the gap between the train and the platform, onto the track, or even under trains, and such incidents may escape notice by the train driver or station workers. A team in the platform may not notice a passenger trapped in the doors or people who are very close to the train or children—who might be at increased risk. CCTV cameras in stations can capture a vast amount of data, and such data is typically archived for some period before being deleted. The recorded data can be utilised by the police as evidence in criminal cases, and the system data can be utilised for monitoring all station operations; however, the outcomes currently depends on employees whose job is to watch video screens all the time. However, CCTV management systems (in control rooms) are passive: they provide only a limited ability to maintain safety in stations. When an emergency situation occurs, it is very challenging to identify and manage the emergency immediately.

Human error in such cases can be high, and the locations of monitored cameras may not fully cover all station areas. Accordingly, it is necessary to systematically observe the risks and any related factors relating to passengers in the station and raise a notification concerning any potential emergency condition in a timely manner.
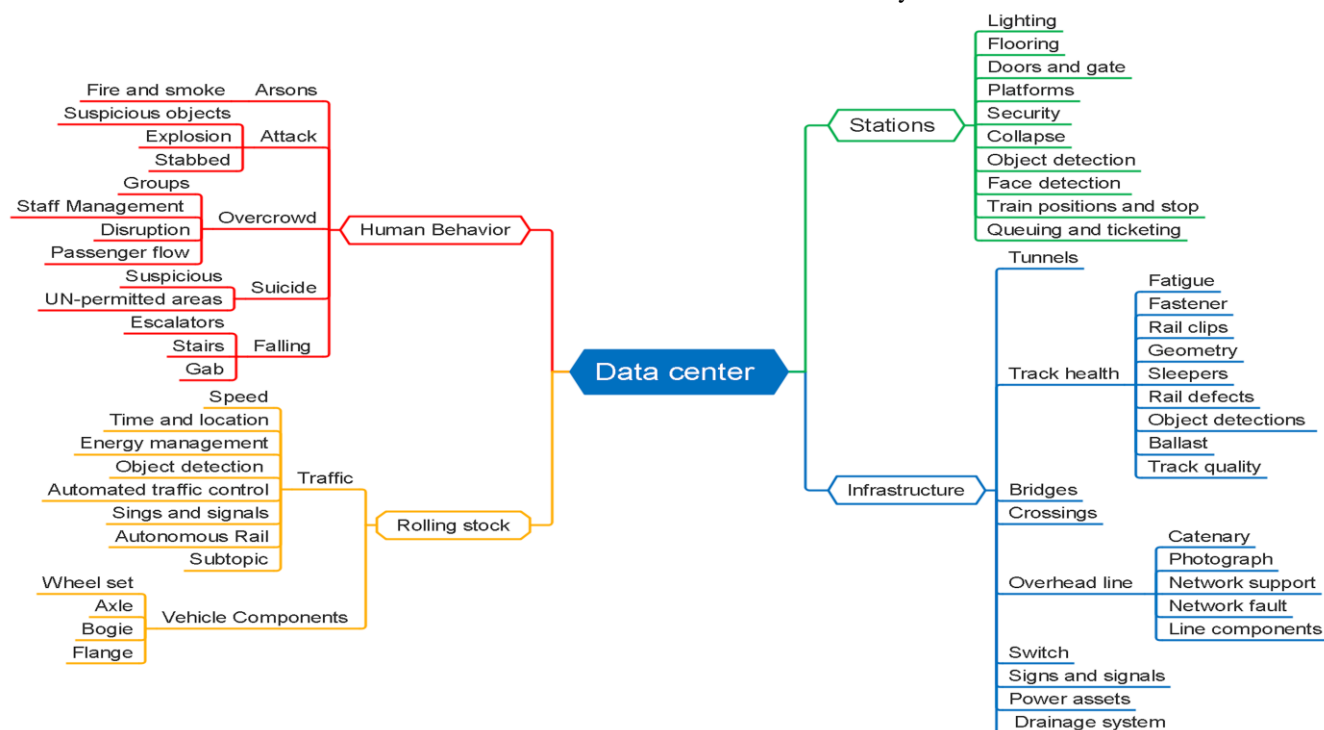

FIGURE 4. Map of Utilising CNN in railway stations shows examples of risk sources to be managed.

Multiple cameras can cover all station areas, such as platforms, tunnels and tracks, while image-processing technology can detect real-time risks and then take actions such as notifying the train driver, the central control room (CCR) and station staff with the information, including the location, time and any alarm message. The captured images can be input to a smart system, which can be trained to recognize any pattern differences and can learn over time. This approach reduces the risk of human error and increases the reliability of real-time predictions. It is expected that utilising a smart method such as a CNN would be able to identify passenger falls, running, overcrowding, or any behaviour or conditions that look suspicious. Some current techniques are effective in detecting suspicious behaviour is in real life, such as Hitachi video analysis [180]. Moreover, a thermal camera has been used to detect human body temperature and used to detect changes in emotion [181].

Video surveillance can play many roles in industry security and safety by utilising advanced detection algorithms and identifying risks in early stages, such as suicide, traffic flow, criminal activity, trespassers, smoke and fires. Advanced methods can detect objects and conduct video analytics to assist emergency responses and support decision-makers. It is expected that these detection techniques would aid in developing emergency response plans and communications schemes, which are critical in reducing risks from emergency events in railway stations. Additionally, the new technologies can contribute to measures for ensuring passenger egress and transit at critical station locations, such as tunnels and access points, for emergency responders. Moreover, advanced analytical video surveillance can cover a range of risks, such as collisions, derailments and intrusions from adjacent areas into unauthorised station locations such as a track [182] while managing other subsystems in real-time with minimum manpower and high efficiency (see Fig. 4). In the literature, it has been noted that achievements in deep learning can enable vision and video processing, classification, image captioning, segmentation, object detection, recognition of human actions from the video, picture recovery, security, observation and so on [136], [183]–[185]. Applying new technology, including image processing, computer vision and machine learning, will provide both direct and indirect benefits, such as improvements in safety and security, such as detecting problems at early stages, resulting in time and cost savings for the long term and lead to automatic many processes in the railway system.

IV. DEEP LEARNING FOR DECISION MAKING IN RISK MANAGEMENT AT RAILWAY STATIONS

DL is a subset of machine learning, which depends on employing nonlinear algorithms to match data. There are many methods that employ this technique, but they
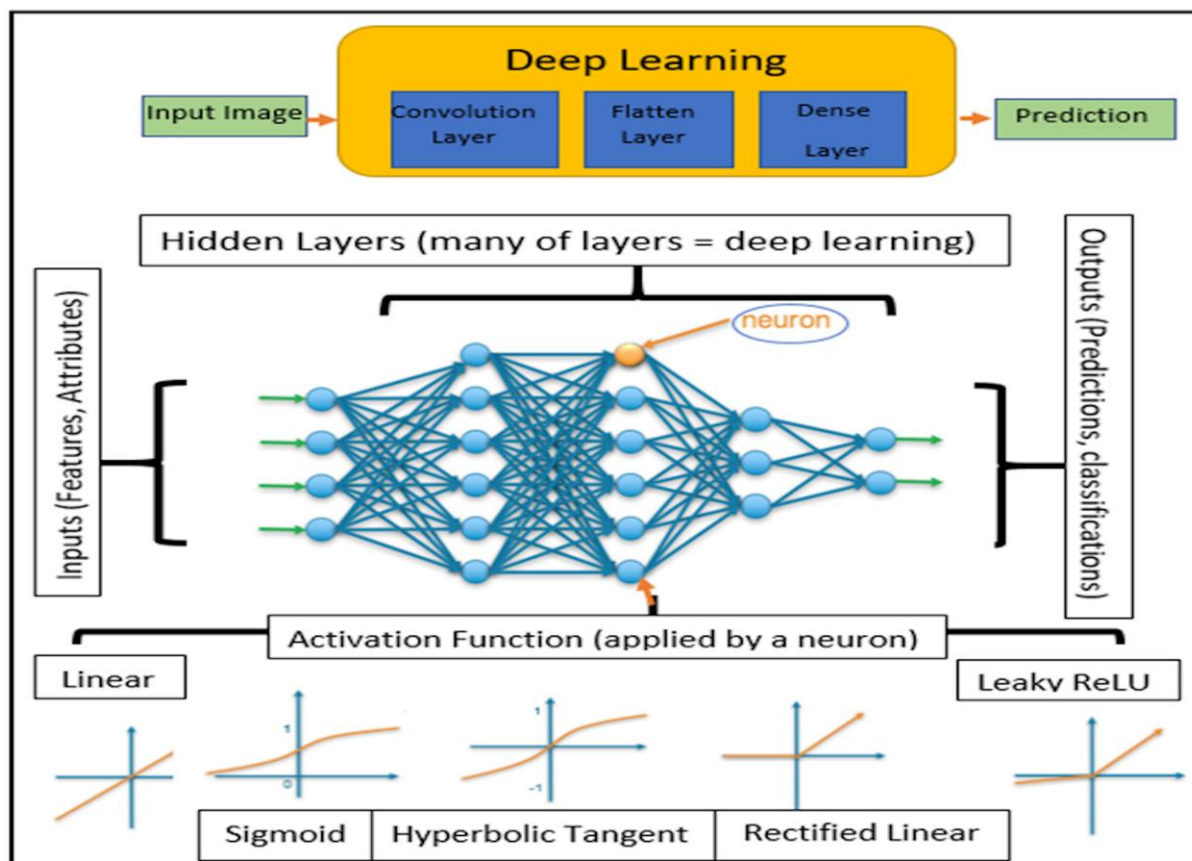
FIGURE 5. Deep learning structure [188].

generally share some commonalities, such as the way each layer receives the output from the former layer as inputs, as shown in Fig. 5. Advancements in hardware and increased data availability have contributed to the ability to effectively train deep CNN networks to identify features not only from static images but also from videos [186], [187]. In addition, including a set of convolution layers in an NN framework has revolutionized image processing. The convolution operation can be defined as follows:

$$s(t) = \int x(a)w(t-a)da,$$

where x and w is the kernel, which is an adaptive filter that the network learns [189].

Video identification is challenging role compared to static images due to the complexities involved in capturing continuous spatial and temporal data [190]. In the past few years, DL has gained enormous power for object detection and tracking. Some object detection algorithms include the region-based convolutional neural networks (RCNN), Faster-RCNN, the single shot detector (SSD) and you only look once (YOLO). Among these, Faster-RCNN and SSD achieve higher accuracy, while YOLO offers is more advantageous speed is given preference over accuracy [191].

Many CNNs are configured to use a graphics processing unit (GPU) as a specialised type of electronic circuit, that can swiftly manipulate and convert memory to accelerate the creation of images in a frame buffer [192].

Currently, most machine learning efforts rely on DL techniques, which connects the layers of an artificial neural network (ANN) to systematically identify patterns in the data that affect decision making. DL is a powerful method of machine learning; however, it requires large amounts of training data to be efficient. Such systems make it possible to make decisions without human input; moreover, the system can learn continuously. For instance, self-driving cars are able to make timely decisions about speed and direction from information captured in real-time from their surroundings.

Offering a decision making algorithm to enhance railway station safety and risk management would be a significant improvement in the use of CCTV data, passenger smartphones, ticketing systems, or other related subsystems in stations. In the initial phase of such applications, we can use DL to support the decision makers; later, in the more advanced phases, we can rely on AI as a highly accurate decision maker. In other words, individuals and AI technologies can cooperate to manage various decision-making challenges (uncertainty, complexity and equivocality) [193], [194]. Based on CCTV systems in stations, which can be updated and utilised to capture video frames and collect data reflecting human actions and motion, the resulting data contain spatial and temporal information from many locations in the station, such as platforms. Then, unsafe acts can be detected using a deep learning method, which is mainly based on a set of algorithms that attempt to model high-level abstractions in the data. The model is trained from multiple frames and the spatial features they contain. For a more comprehensive application, we compare the traditional risk management process to a CNN model process to present the steps of the two systems in parallel (see Fig. 6).

Both outcomes will support the decision-maker and reduce uncertainties to a low level in complex systems. The process improvement will support many field activities such as maintenance, passenger crowding. System reviews will enhance actions, add alternatives and redesign the processes regarding predictions and advanced analytics and—importantly, training the model. The cycle of control, continuous improvement and incorporating lessons learned is an essential part of a safety system; thus, this innovative approach fits well into that process, as shown in see Fig. 7.

V. MODEL FRAMEWORK

Railway station monitoring is vital to guarantee that people and the rail system are safe and secure. A monitoring failure can result not only in significant impacts to train delays and maintenance costs but also to passenger safety at the station and to society and the economy.
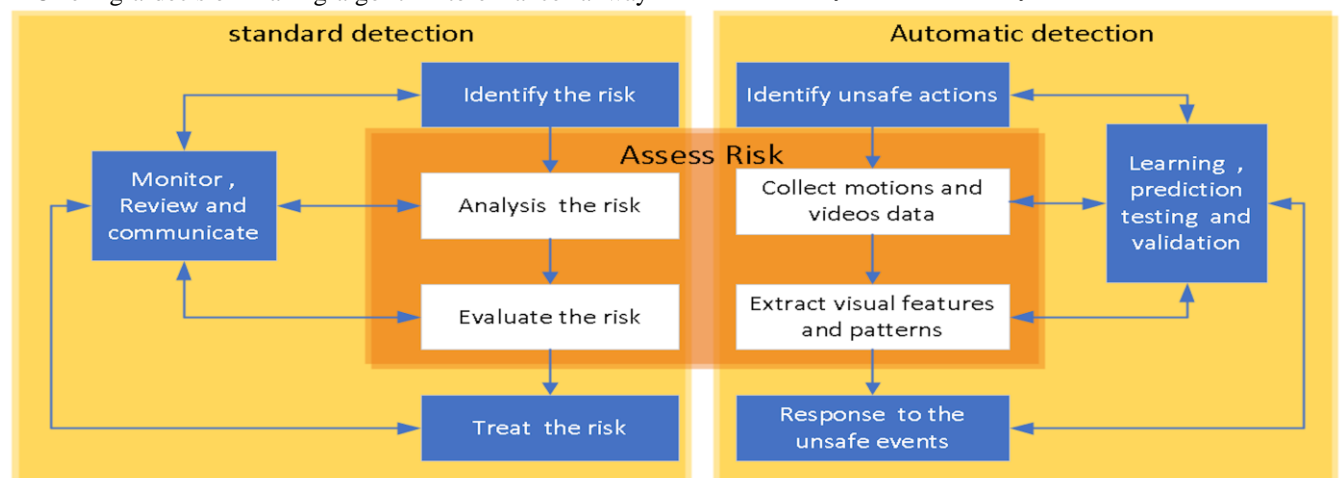

FIGURE 6. Diagram of the risk management steps process compared to the proposed CNN method.

IEEE *Access*
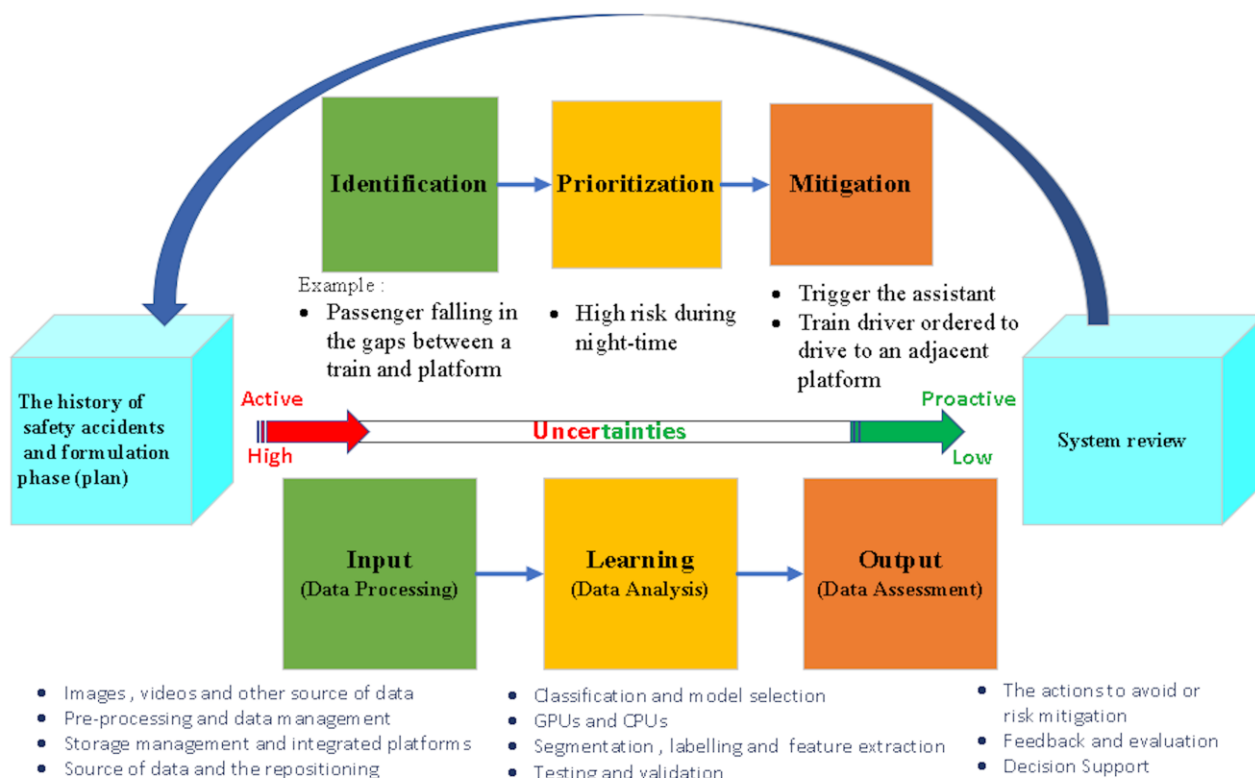
FIGURE 7. The safety cycle transferred to the approach with new techniques.
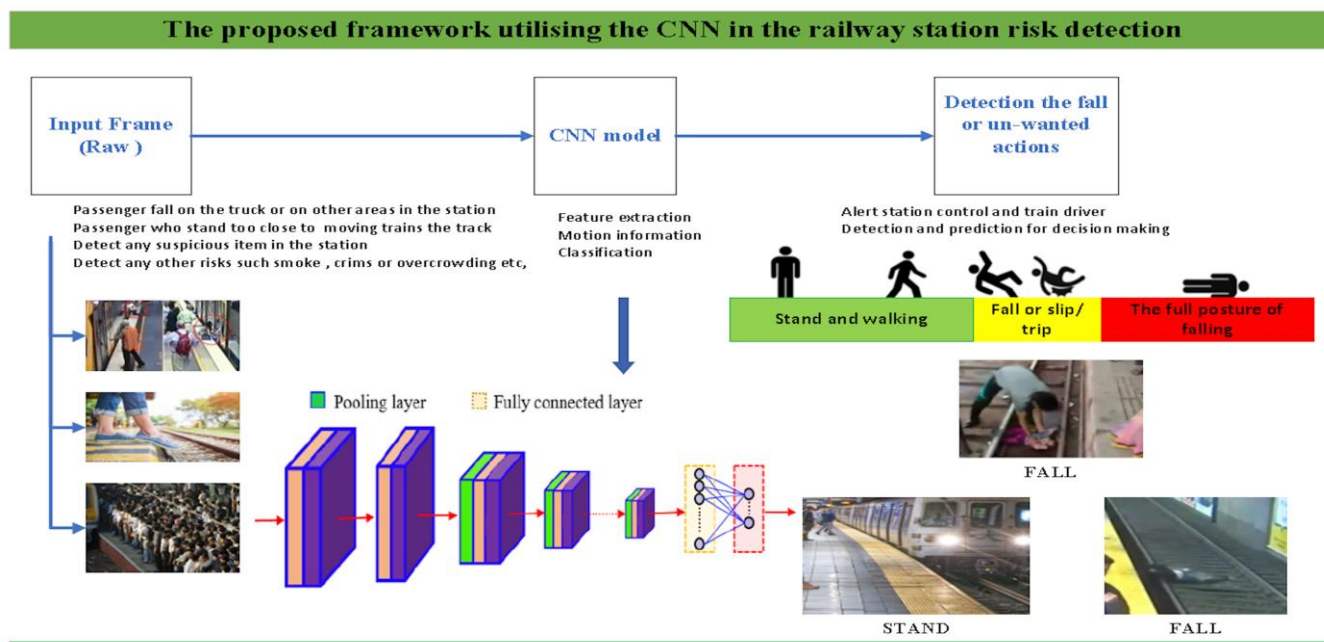
FIGURE 8. Flowchart of the offered approach (Framework).

In the case model adopted in this paper, the goal is to manage the risk of falls by detecting and analysing passengers automatically among the enormous amount of data from CCTV cameras. The outcomes illustrate the practicality and efficiency of the proposed approach. This model relies on image-detection methods and introduces a risk management framework that uses a CNN to analyse the images or videos to detect risks. The proposed framework is depicted in Fig. 8. Video images can be used to identify deficiencies, such as interruptions to passenger flow that cause falls, which leads to

overcrowding. In addition, they can be utilised to discover unwanted events that occur in the station.

VI. THE CNN CASE MODEL

The goal of the proposed model is to detect falls based on a CNN. To implement such a method, the system needs to be able to characterise the complex motions of passengers and address more than one passenger fall at the same time. When a fall is detected, the system should present the significant information to railway station control, such as the time and location. The difference between a CNN and ordinary neural networks is that each neuron in a CNN is locally connected to only a few neurons in the previous layer; not to all neurons, as is the case in ordinary neural networks. This enables CNNs to be used to construct deeper networks and, consequently, learn more complex features [195]. Furthermore, CNNs have demonstrated high performance and are relatively easy to train. A basic CNN can be characterised as having two layers: a convolutional base layer that extracts features from an image and a classifier (a fully connected layer) to classify the image based on the detected features. Each frame undergoes a the data acquisition phase that supplies the system with the digitized data from such images. These data may include many events or statuses and can be acquired from both internal and external networks, such as traffic and/or track conditions and weather. Then, manipulation or data mining processes such as feature selection, extraction and standardisation can be applied to process the raw data for analysis. The data can contain many layers, including the acquisition time and location. Next, an appropriate model and deep learning technique are used to perform feature detection and make predictions along with the actions and triggers to be activated when a threshold is breached. The goal is to create a proactive system that can avoid or mitigate unwanted events. The history of events and scenarios from many points in the system will improve the prediction accuracy, and the model is trained from past activities, as shown in Fig. 9.

CNNs have become the main tool used for recent innovations in the comprehension of images [196], videos [136], [197] and audio signals [183], [184].

In this study, we used the Keras library, which is an open source neural network library written in Python that supports easy and fast prototyping. Furthermore, it maintains the CNN and runs seamlessly on both CPUs and GPUs. Keras is compatible with other Python code and can use raw images as inputs to the CNN model, which extracts features. A summary of the experimental configuration is shown in Table I.

TABLE I
KERAS LIBRARY PACKAGES EXAMPLE INSTALLED DETAILS RESULTS AFTER RUNNING THE MODEL

Layer (type)	Output Shape	Param #
Conv2d_1 (conv2D)	(None,62,62,64)	1792
Max_pooling2d_1 (MaxPooling2)	(None,31,31,64)	0
Conv2d_1 (conv2D)	(None,29,29,128)	73856
Max_pooling2d_1 (MaxPooling2)	(None,14,14,128)	0
Flatten_1(Flatten)	(None,25088)	0
Dense_1(Dense)	(None,128)	3211392
Dense_2(Dense)	(None,64)	8256
Dense_3(Dense)	(None,32)	2080
Dropout_1(dropout)	(None,32)	0
Dense_4(Dense)	(None,32)	1056
Dropout_2(dropout)	(None,32)	0
Dense_5(Dense)	(None,1)	33

Total params: 3,298,465
Trainable params: 3,298,465
Non-trainable params :0

In this part of the study, we employed available processors to execute the framework; however, for large data, employing more powerful CPUs and GPUs is recommended.

We build a model layer by layer using the sequential model type was selected, which is the simplest way to build a model in Keras. Next, to deal with input images as 2D matrices, we selected Conv2D layers with 64 nodes in each layer. A 3×3 filter matrix was used for the convolution kernel (see Fig. 10). A CNN structure includes convolutional layers that are the major building blocks; these layers learn the features that are suitable for differentiating between a 'falling' image and a 'not falling' image. Each convolutional layer employs a set of kernels that apply a convolution operation based on the outputs of the preceding layers.
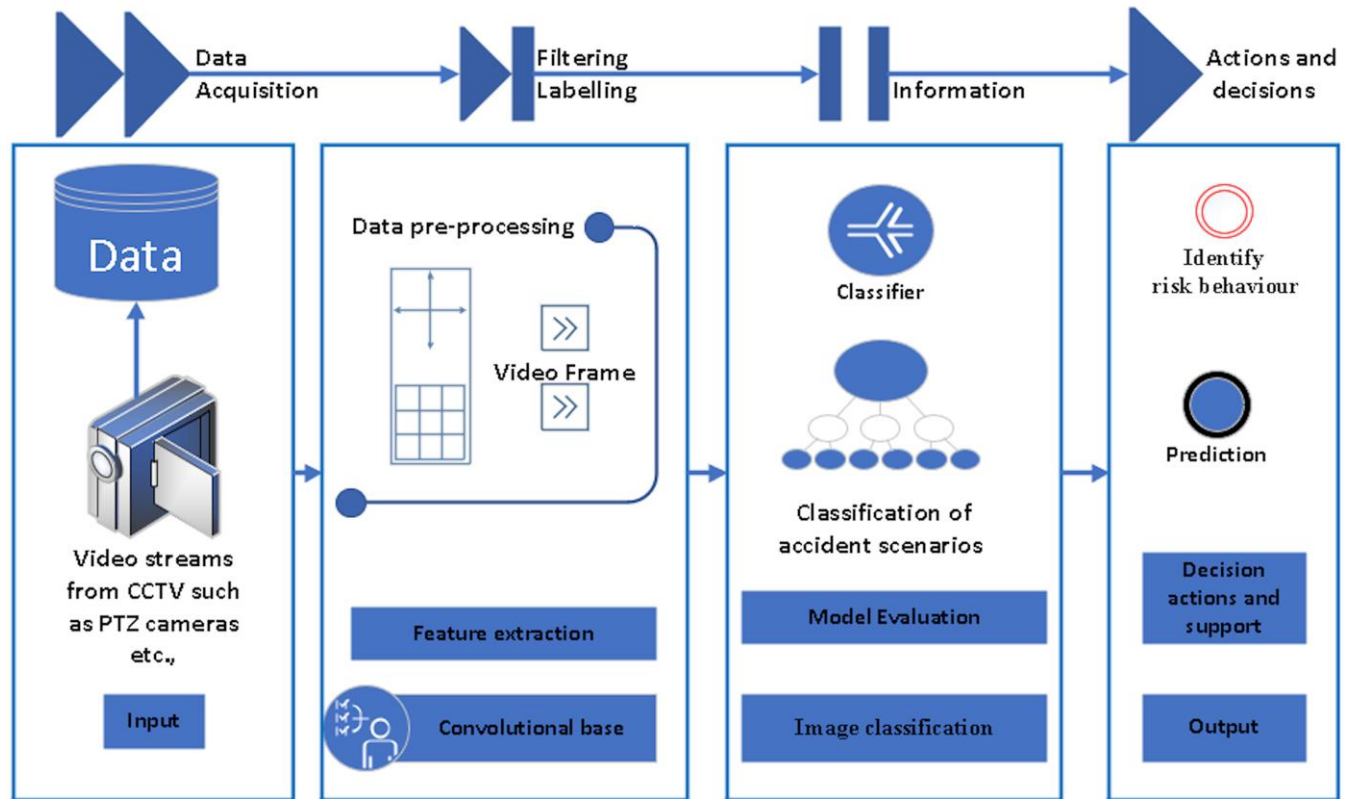
FIGURE 9. Architecture of a CNN-based model.

We adopted the rectified linear unit (ReLU) as an activation function because ReLU has previously been shown to work well in neural networks.

For the output layer, we selected a dense layer, which is a standard type of layer used in various neural network architectures. To connect the convolution and dense layers, a flattening layer is preferable. In addition, we used dropout layers between the, various layers to avoid data overfitting [197], [198], [199].

VII. DATA AND ANALYSIS

A. EXPERIMENTS

1) DATASETS

The objective of the CNN in this study is to take input image data of passengers in a station and classify each image into one of two classes: either 'falling' or 'not falling'. The dataset was divided into separate frames with known labels (falling or not falling), which were then used as training data for the classifier. To construct the dataset, different activities and complex falling events from different locations were selected from many open source sequences, such as falls on stairs or in the gap between the train and platform, as shown in Fig. 11.

We implemented the proposed method after training to predict risk states in a railway system (at the station) and evaluated the performance of the model. For all the experiments, we used one computer equipped with an Intel Core i7 CPU, 64 GB of memory and an NVIDIA GeForce MX 150 GPU.

We gathered data consisting of both still frames and videos from open sources. Finding such data is challenging both because of privacy concerns and lack of availability for many reasons, such as that data is deleted from data centres periodically and the difficulty of finding and collecting such data.
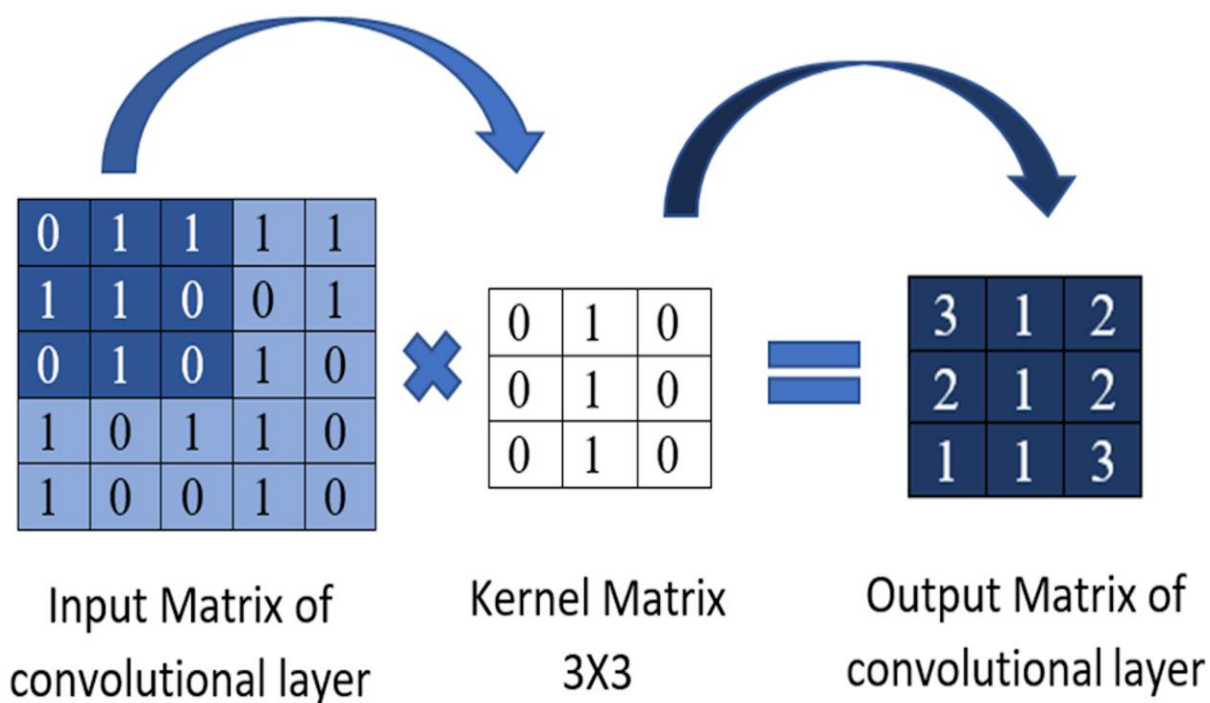
FIGURE 10. An example of a convolutional layer operation.

FIGURE 11. (a) Sample falling passenger images. (b) Sample passenger does not fall images.

The data must clustered, classified and labelled; the images show some risky passenger behaviours (see Fig. 12) and the collected videos and images cover many countries. The data raise significant concerns that should be considered in the future station design and in daily operations, for example:

- People standing in risky positions near the gap between the train and the platform and close to moving trains.
- Some people trespass into the track areas and can be found in restricted areas.
- Some passengers cross the tracks to take shortcuts between locations.
- Some passengers are pressed against the train and their clothes become trapped in train doors.
- Elderly people fall on escalators and other passengers misuse the escalators.
- Children and those susceptible to fainting falling into the gap between the train and the platform.
- Impacts from technology and lifestyles, such as taking selfies behind the trains

- Wheelchairs falling down stairs and escalators.
- Passengers standing in restricted areas.

However, the limited data available for each dataset are not sufficient for training deep learning models. Thus, we augmented the collected data with the Le2i dataset built by Charfi et al. [200], which covers many falling positions, and our model performs only binary classification: falling or not falling. We split the dataset into training and test datasets.

2) PRE-PROCESSING AND PREDICTIONS
Despite the data limitations, the data (both images and videos) collected from the web require intensive cleaning. The variety of sources imposes many constraints, images with poor quality and obstructed vision (to the point that the targets cannot be seen) must be removed. After being trained, the model is applied to a test dataset, in which the images have not been seen before, to classify the risk of falling.
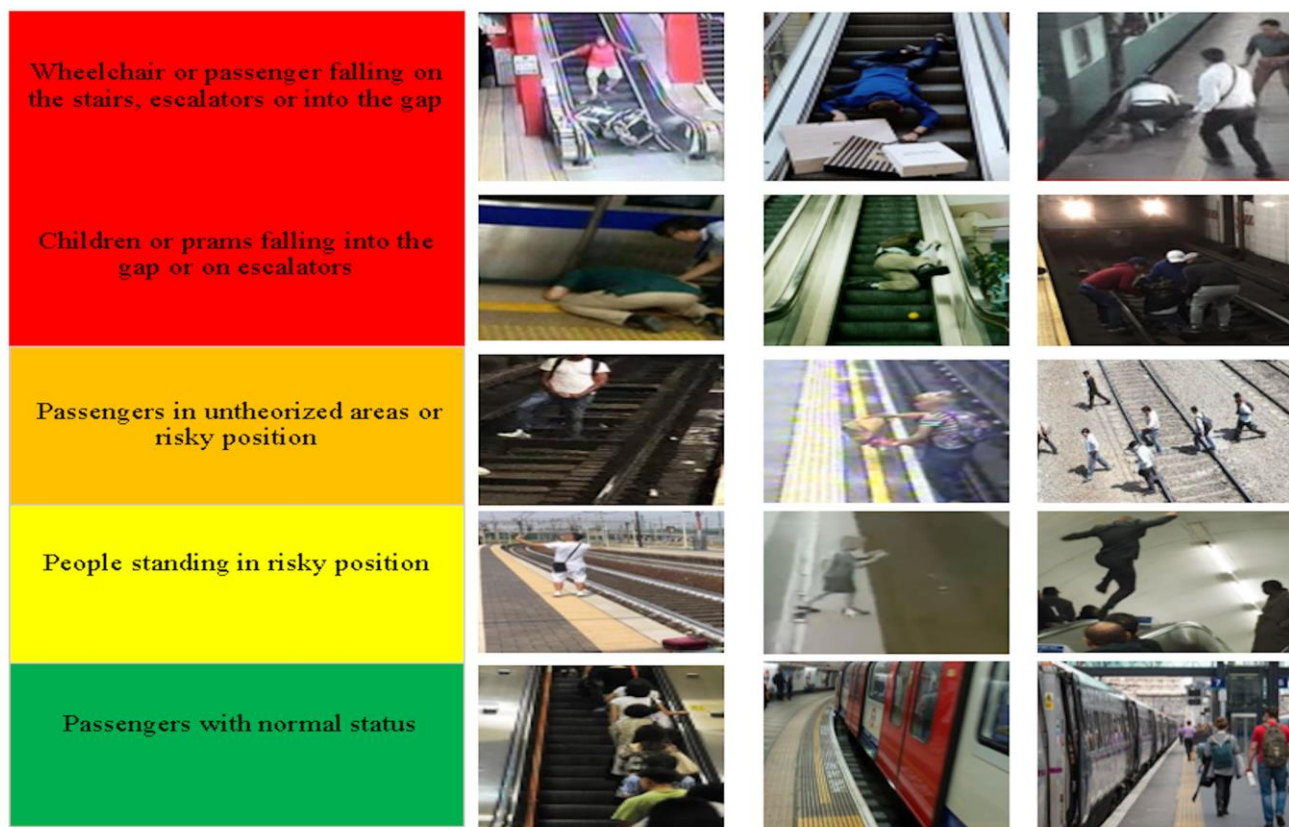
FIGURE 12. Examples of clusters of passengers' behaviours.

Using randomly selected open source images, we divided the data into three sets (training, testing and prediction).The CNN training outcome of results in an accuracy increase with each model training iteration; thus, the model performance validation data eventually reach an acceptable level as the error decreases, as shown in. After training rule, the prediction

ability of the model was evaluated on the test sample (see Fig. 14).

3) THE EVALUATION
During the testing process, performance indicators can be calculated from the trained model output. We selected indicators such as accuracy, precision, recall and F1-score and

the receiver operating characteristic curve (ROC) for this study. For Predictions, we are focused on identifying the fall risk. Hence, we sample data present the fall and the behaviour of falling occurrence which cover unsafe people positions. For an estimate the risk and to identify the best classes, different datasets cases are studied. Falling and not (case1), and three categories, fall, not falling (normal or safe station) and unsafe behaviour (case 2) see Table 2 and Fig. 13.

TABLE 2
DETAILS RESULTS OF THE EXPERIMENTS MODELS

Model ID	Number of classes	Classes labels	Accuracy	Number of Parameters	Dataset
CNN-0	3	Fall Not fall Risky position	81.90	3,301,635	1782
CNN-1		Fall Not fall Risky position	71.71	3,301,635	1826
CNN-2		Passenger Fall Risky behaviour Safe station	77.90	3,301,635	1781
CNN-3	2	Fall	80.77	3,303,681	2820
CNN-4		Not fall	**82.2**	3,298,465	1980
CNN-5			75.00	3,298,465	2885

The prediction model classifies instances of passenger behaviour using a two-class prediction (case1) show the high results. When the prediction is positive and the ground-truth value is also positive, the prediction is called a true positive (TP). Similarly, false positive (FP), true negative (TN) and false negative (FN) values can be calculated [141]. These four values can be presented as a 2×2 contingency table, called confusion matrix, as depicted in Fig. 15.

The ROC is a metric that reflects both the sensitivity and specificity of continuous variables and reveals the relationship between them. The ROC curve of the case study results is shown in Fig. 16.

From the perspective of the ROC curve, the model performs effectively in making falling predictions.

The lack of risk-class images in accident cases at stations means that an uneven number of pictures exists between the risk and non-risk groups. Thus, the data make it challenging to model safety vs. risk when training a deep learning model utilising the available images. In this study, we used 80% of the dataset for training, 10% for validation and the remaining 10% for model testing.
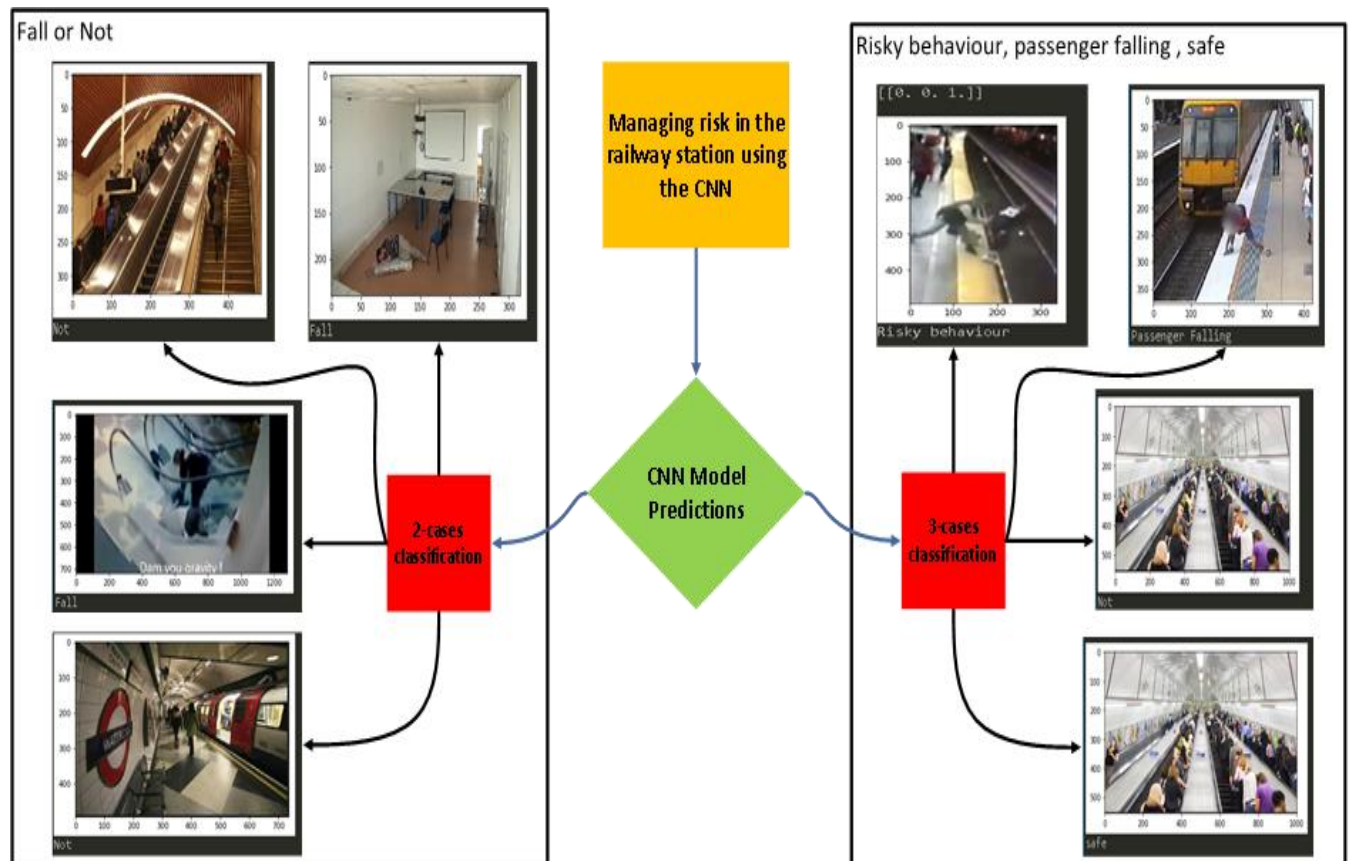

FIGURE 13. Examples of clusters of passengers' behaviours Predictions.

VIII. DISCUSSION AND CONCLUSIONS

In this study, we proposed a process for moving from conventional risk management to applying innovative technology to risk management; such an approach can improve safety and security throughout the entire railway industry paradigm. Many benefits can be gained from deep learning in risk management, such as the following.

- A real-time ability can be gained to help avoid risks
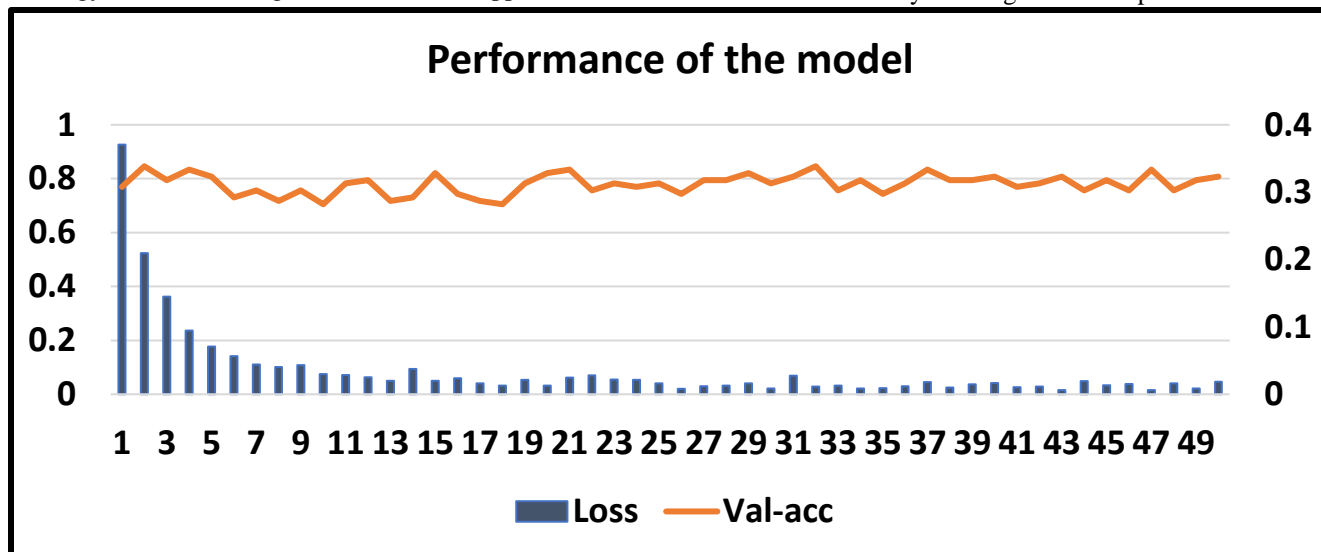
FIGURE 14. The trained model's accuracy and the model loss.

FIGURE 15. The classification performance evaluation results shown as a confusion matrix.

- Many subsystems in the field can be integrated, including maintenance, security, traffic and passenger models, to form actions that consider multiple aspects.
- Lessons and experience can be integrated into the learning process and automated effectively via machine learning, which is critical for safety systems.
- The effectiveness of operations in stations and other areas linked to railway activities can be improved.
- Time and costs can be saved while improving accuracy to enable long-term quality improvements
- Both passenger and workforce experiences can be improved, which reflect on the overall market image.
- Data gathering can be enhanced to more fully utilise effective connections between assets and people.

FIGURE 16. ROC curve of the model.

We propose an efficient railway system technique framework based on a CNN and applied DL algorithms to foster detection of unwanted events in railway stations. We ado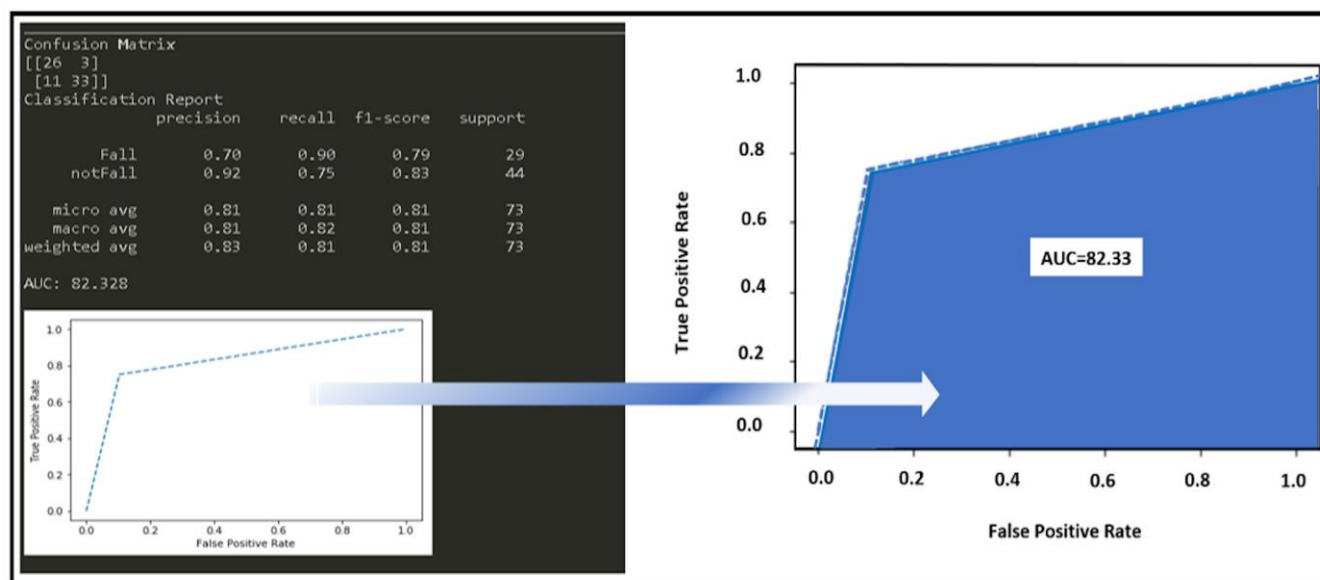pted a CNN to extract events such as passenger falls, which may occur on stairs, escalators, or platforms. Different scenarios were anticipated, such as suicide or falling under moving or stopped trains. The fall event detection process can alert the station control centre and then action can be taken to better clarify the situation, which might be an attack, crime, or intoxicated passenger incident. Timely detection will mitigate the risks to other passengers, lead to more rapid responses in emergency or evacuation situations and decrease the potential hospitalisation time. We presented datasets from open sources; however, compiling additional datasets containing training examples would improve the accuracy and cover a wider range of station risks. CNN-based methods require a large pool of labelled training data, and collecting and labelling such data is a complex task. Nevertheless, automatic detection can help with timely maintenance and risk control, and the results can be used as feedback to train the model to obtain improved accuracy. The results with the proposed model confirm that increasing the depth of a deep network can lead to better performances in terms of accuracy. Finding accident data (such as falls in stations) is challenging for many reasons, such as the lack of available data and passenger privacy concerns. Using methods such as computer vision techniques will improve timely risk management, detection and safety and ultimately affect risk management in railway systems. The proposed method could be generalised to detect other risks, such as people running, overcrowding, suspicious item detection or other complex activities in addition to fall patterns in stations. The CNN approach provides real-time, accurate visual monitoring of the risks in railway stations to assist safety or risk management operations, which are reflected in passenger services. The method is more suitable for real world conditions and is cost-effective (enabling, for example, 24-hour monitoring of CCTV cameras with the intention of identifying potential acts of vandalism). Image processing has been shown to be a promising technology that has the ability to improve station safety, manage risks, reduce dwell times and reduce the number of operators at stations. Moreover, image processing techniques are useful for detecting congestion, assessing flow, accessing dangerous zones, identifying people moving in forbidden areas and notifying train drivers about foreign objects ahead [19], [20].

Our results demonstrate that the proposed CNN model can automatically extract and classify risky behaviours (i.e., falling on the platform) with a high level of accuracy. The method carries high confidence that all the objects in a data sequence are detected and recognized. Nevertheless, this CNN model should be improved and implemented to automatically detect risk actions related to human behaviour or asset conditions both during normal operations and in any unanticipated conditions. Such models can lead to intervention by management or execute high-level automated actions; these can directly modify behaviours and mitigate risks or reduce the consequences of accidents. Moreover, the results can be used to provide designers, operators and decision-makers with direct visual outcomes and to allow them to learn how to deliver operations more safely. Additionally, the results indicate that the process can achieve efficient railway system detection under numerous conditions, including aspects such as:

- Safety and security
- Infrastructure and assets
- Maintenance and traffic management operations
- Quality and reliability

- Operations, passengers, train drivers, workforce management and so on,

The development of specialized algorithms for the model can overcome errors and improve response time. By capitalizing of existing CCTV systems, the costs are expected to be reduced over the long term and improve system efficiency by considering the locations and coverage of the cameras. The model offers other benefits to stations that are worth further research, such as predictive maintenance, emergency plans, people counting, train positioning and security. However, the data availability and quality remain a challenge because this technology depend heavily on large amounts of high-quality data. Finally, it is time to invest in AI to benefit railway systems, making them safer for staff, customers and the public.

ACKNOWLEDGMENT

REFERENCES

[1] Office of Rail and Road. "Annual statistical release - rail safety statistics." http://orr.gov.uk/__data/assets/pdf_file/0007/22876/rail-safety-statistics-2015-16.pdf (accessed Mar. 5, 2019).

[2] Rail Safety and Standards Board (RSSB). "Annual safety performance report." http://www.hwa.uk.com/site/wp-content/uploads/2017/09/CD6.8-RSSB-Annual-Report-on-Public-Safety-2015-2016.pdf (accessed Oct. 26, 2019).

[3] World Health Organization. "WHO global report on falls prevention in older age.",Community Health., https://www.who.int/violence_injury_prevention/publications/other_injury/falls_prevention.pdf?ua=1.(accessed Apr. 04, 2020).

[4] P. E. Waterson, V. L. Kendrick, R. A. Haslam, T. Jun, and B. Ryan, "Probing deeper into the risks of slips, trips and falls for an ageing rail passenger population: applying a systems approach," *IET Intell. Transp. Syst.*, vol. 10, no. 1, pp. 25–31, Feb. 2016, doi: 10.1049/iet-its.2015.0047.

[5] R. Marks and J. P. Allegrante, "Falls-prevention programs for older ambulatory community dwellers: From public health research to health promotion policy," *Sozial Präventivmedizin/Social Preventive Medicine*, vol. 49, no. 3, pp. 171–178, Jun. 2004, doi: 10.1007/s00038-004-3040-z.

[6] D. Helbing, I. Farkas, and T. Vicsek, "Simulating dynamical features of escape panic," *Nature*, vol. 407, no. 6803, pp. 487–490, Sep. 2000, doi: 10.1038/35035023.

[7] W. Fang, P. E. D. Love, H. Luo, and L. Ding, "Computer vision for behaviour-based safety in construction: A review and future directions," *Adv. Eng. Inform.*, vol. 43, p. 100980, Jan. 2020, doi: 10.1016/j.aei.2019.100980.

[8] S. Han and S. Lee, "A vision-based motion capture and recognition framework for behavior-based safety management," *Automat. Construction*, vol. 35, pp. 131–141, Nov. 2013, doi: 10.1016/j.autcon.2013.05.001.

[9] S. Han, S. Lee, and F. Peña-Mora, "Vision-based detection of unsafe actions of a construction worker: Case study of ladder climbing," *J. Comput. Civil Eng.*, vol. 27, no. 6, pp. 635–644, Nov. 2013, doi: 10.1061/(asce)cp.1943-5487.0000279.

[10] W.-R. Chang, S. Leclercq, T. E. Lockhart, and R. Haslam, "State of science: Occupational slips, trips and falls on the same level," *Ergonomics*, vol. 59, no. 7, pp. 861–883, 2016, doi: 10.1080/00140139.2016.1157214.

[11] R. Haslam and D. Stubbs, *Understanding and Preventing Falls: An Ergonomics Approach*: CRC Press,Boca Raton,Fla.;London:Taylor and Francis, pp. 209–234.2005.

[12] R. Haslam and V. Filingeri, "Slips, trips and falls in crowds," in *Proc. 20th Congr. Int. Ergonom. Assoc. (IEA 2018)*, Cham, Switzerland, Aug. 5, 2019, pp. 752–758.

[13] S. Y. Z. Au, G. Jenny, A. D. Livingston, and H. Roger, "Assessing spectator safety in seated areas at a football stadium," in *Contemporary Ergonomics 2004*, P. T. McCabe, Ed., Boca Raton, FL: CRC Press, 2014, pp. 23–27.

[14] V. L. Kendrick, "The user experience of crowds," Ph.D. dissertation, Design School,Loughborough Univ., Loughborough, UK, 2013.

[15] I. Ahmad, Z. He, M. Liao, F. Pereira, and M. T. Sun, "Special issue on video surveillance," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 8, pp. 1001–1005, Aug. 2008, doi: 10.1109/tcsvt.2008.929646.

[16] J. Teizer and P. A. Vela, "Personnel tracking on construction sites using video cameras," *Adv. Eng. Inform.*, vol. 23, no. 4, pp. 452–462, Oct. 2009, doi: 10.1016/j.aei.2009.06.011.

[17] J. Yang, O. Arif, P. A. Vela, J. Teizer, and Z. Shi, "Tracking multiple workers on construction sites using video cameras," *Adv. Eng. Inform.*, vol. 24, no. 4, pp. 428–434, Nov. 2010, doi: 10.1016/j.aei.2010.06.008.

[18] Z. Zhu and I. Brilakis, "Concrete column recognition in images and videos," *J. Comput. Civil Eng.*, vol. 24, no. 6, pp. 478–487, Nov. 2010, doi: 10.1061/(asce)cp.1943-5487.0000053.

[19] I. P. T. Weerasinghe and J. Y. Ruwanpura, "Automated data acquisition system to assess construction worker performance," in *Construction Res. Congr. 2009 Building Sustain. Future*, Seattle, Washington, 2009, pp. 61–70.

[20] J. Gong, C. H. Caldas, and C. Gordon, "Learning and classifying actions of construction workers and equipment using Bag-of-Video-Feature-Words and Bayesian network models," *Adv. Eng. Inform.*, vol. 25, no. 4, pp. 771–782, 2011/10 2011, doi: 10.1016/j.aei.2011.06.002.

[21] C. Chen, Z. Zhu, and A. Hammad, "Automated excavators activity recognition and productivity analysis from construction site surveillance videos," *Automat. Construction*, vol. 110, p. 103045, Feb. 2020, doi: 10.1016/j.autcon.2019.103045.

[22] S. Zheng, X. Chai, X. An, and L. Li, "Railway track gauge inspection method based on computer vision," in *2012 IEEE Int. Conf. Mechatronics Automat.*, Chengdu, China, Aug. 5–8, 2012, pp. 1292–1296.

[23] R. K. W. Vithanage, C. S. Harrison, and A. K. M. DeSilva, "Importance and applications of robotic and autonomous systems (RAS) in railway maintenance sector: A review," *Computers*, vol. 8, no. 3, p. 56, Jul. 2019, doi: 10.3390/computers8030056.

[24] A. Zaman, B. Ren, and X. Liu, "Artificial intelligence-aided automated detection of railroad trespassing," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2673, no. 7, pp. 25–37, May 2019, doi: 10.1177/0361198119846468.

[25] Y. Wu, Y. Qin, Z. Wang, X. Ma, and Z. Cao, "Densely pyramidal residual network for UAV-based railway images dehazing," *Neurocomputing*, vol. 371, pp. 124–136, Jan. 2020, doi: 10.1016/j.neucom.2019.06.076.

[26] S. Ma, L. Gao, X. Liu, and J. Lin, "Deep learning for track quality evaluation of high-speed railway based on vehicle-body vibration prediction," *IEEE Access*, vol. 7, pp. 185099–185107, Mar. 2019, doi: 10.1109/access.2019.2960537.

[27] A. K. Singh, A. Swarup, A. Agarwal, and D. Singh, "Vision based rail track extraction and monitoring through drone imagery," *ICT Express*, vol. 5, no. 4, pp. 250–255, Dec. 2019, doi: 10.1016/j.icte.2017.11.010.

[28] M. Banić, A. Miltenović, M. Pavlović, and I. Ćirić, "Intelligent machine vision based railway infrastructure inspection and monitoring using UAV," *Facta Universitatis Ser. Mech. Eng.*, vol. 17, no. 3, pp. 357–364, Nov. 2019, doi: 10.22190/fume190507041b.

[29] H. Alawad and S. Kaewunruen, "Wireless sensor networks: Toward smarter railway stations," *Infrastructures*, vol. 3, no. 3, p. 24, Jul. 2018, doi: 10.3390/infrastructures3030024.

[30] J. Liu, T. Li, P. Xie, S. Du, F. Teng, and X. Yang, "Urban big data fusion based on deep learning: An overview," *Inf. Fusion*, vol. 53, pp. 123–133, Jan. 2020, doi: 10.1016/j.inffus.2019.06.016.

[31] P. McMahon, T. Zhang, and R. Dwight, "Requirements for big data adoption for railway asset management," *IEEE Access*, vol. 8, pp. 15543–15564, Jan. 2020, doi: 10.1109/access.2020.2967436.

[32] P. Fraga-Lamas, T. M. Fernández-Caramés, and L. Castedo, "Towards the internet of smart trains: A review on industrial iot-connected railways," *Sensors*, vol. 17, no. 6, p. 1457, Jun. 2017, doi: 10.3390/s17061457.

[33] M. Takikawa, "Innovation in railway maintenance utilizing information and communication technology (smart maintenance initiative)," *Japan Railway & Transit Review*, no. 67, pp. 22–35, Mar. 2016. [Online].Available: http://www.ejrcf.or.jp/jrtr/jrtr67/pdf/22-35.pdf [34] D. Xie, L. Zhang, and L. Bai, "Deep learning in visual computing and signal processing," *Appl. Comput. Intell. Soft Comput.*, vol. 2017, p. 1320780, Feb. 2017, doi: 10.1155/2017/1320780.

[35] D. Yu and L. Deng, "Deep learning and its applications to signal and information processing [exploratory DSP]," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 145–154, Jan. 2011, doi: 10.1109/msp.2010.939038.

[36] O. Abdel-Hamid, A. Mohamed, H. Jiang, and G. Penn, "Applying Convolutional Neural Networks concepts to hybrid NN-HMM model for speech recognition," in *2012 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Kyoto, Japan, Mar. 25–30, 2012, pp. 4277–4280.

[37] H. C. Shin *et al.*, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016, doi: 10.1109/TMI.2016.2528162.

[38] B. H. Menze *et al.*, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015, doi: 10.1109/TMI.2014.2377694.

[39] Y. Pan *et al.*, "Brain tumor grading based on Neural Networks and Convolutional Neural Networks," in *2015 37th Annu. Int. Conf. IEEE Eng. Medicine Biol. Soc. (EMBC)*, Milan, Italy, Aug. 25–29, 2015, pp. 699–702.

[40] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale convolutional neural networks for lung nodule classification," in *Inf. Process. Med. Imaging*, S. Ourselin, D. C. Alexander, C.-F. Westin, and M. J. Cardoso, Eds., Cham, Switzerland: Springer International Publishing, 2015, pp. 588–599.

[41] G. Luo, R. An, K. Wang, S. Dong, and H. Zhang, "A deep learning network for right ventricle segmentation in short-axis MRI," in *2016 Comput. Cardiology Conf. (CinC)*, Vancouver, BC, Sep. 11–14, 2016, pp. 485–488.

[42] F. Altaf, S. M. S. Islam, N. Akhtar, and N. K. Janjua, "Going deep in medical image analysis: Concepts, methods, challenges, and future directions," *IEEE Access*, vol. 7, pp. 99540–99572, Feb. 2019, doi: 10.1109/access.2019.2929365.

[43] L. Zhou and X. Gu, "Embedding topological features into convolutional neural network salient object detection," Neural Networks, vol. 121, pp. 308–318, Jan. 2020, doi: 10.1016/j.neunet.2019.09.009.

[44] H. R. Roth *et al.*, "Improving computer-aided detection using convolutional neural networks and random view aggregation," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1170–1181, May 2016, doi: 10.1109/tmi.2015.2482920.

[45] X. Wang, L. Zhang, L. Lin, Z. Liang, and W. Zuo, "Deep joint task learning for generic object extraction," in *Advances Neural Inf. Process. Syst.*, Montreal, Canada, Dec. 8–11, 2014, pp. 523–531.

[46] J. Liu *et al.*, "Colitis detection on abdominal CT scans by rich feature hierarchies," in *Med. Imag. 2016: Comput. Aided Diagnosis*, San Diego, CA, Mar. 24, 2016, p. 97851N.

[47] H. Fan, H. Su, and L. Guibas, "A point set generation network for 3D object reconstruction from a single image," in *2017 IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, Honolulu, HI, Jul. 21–26, 2017, pp. 2463–2471.

[48] D. Chen, X. Cao, F. Wen, and J. Sun, "Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification," in *2013 IEEE Conf. Comput. Vision Pattern Recognit.*, Portland, OR, Jun. 23–28, 2013, pp. 3025–3032.

[49] X. Cao, D. Wipf, F. Wen, G. Duan, and J. Sun, "A practical transfer learning algorithm for face verification," in *2013 IEEE Int. Conf. Comput. Vision*, Sydney, Australia, Dec. 1–8, 2013, pp. 3208–3215.

[50] T. Berg and P. N. Belhumeur, "Tom-vs-pete classifiers and identity-preserving alignment for face verification," in *Proc. Brit. Mach. Vision Conf.*, Surrey, Sep. 3–7, 2012, pp. 129.121–129.111.

[51] A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, *Computer Vision – ECCV 2012: 12th European Conference on Computer Vision*. Florence, Italy: Springer, 2012.

[52] X. Chen and J. Guhl, "Industrial robot control with object recognition based on deep learning," in Procedia CIRP, 2018, vol. 76, pp. 149–154, doi: 10.1016/j.procir.2018.01.021.

[53] S. Kuutti, R. Bowden, Y. Jin, P. Barber, and S. Fallah, "A survey of deep learning applications to autonomous vehicle control," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–22, Dec. 2020, doi: 10.1109/tits.2019.2962338.

[54] J. Kim and J. Canny, "Interpretable learning for self-driving cars by visualizing causal attention " in *Proc. IEEE Int. Conf. Comput. Vision*, 2017, pp. 2942–2950.

[55] S. Shalev-Shwartz, S. Shammah, and A. Shashua, "On a formal model of safe and scalable self-driving cars," 2017, *arXiv preprint arXiv:1708.06374*.

[56] T. Huang, S. Wang, and A. Sharma, "Highway crash detection and risk estimation using deep learning," *Accident Anal. Prevention*, vol. 135, p. 105392, Feb. 2020, doi: 10.1016/j.aap.2019.105392.

[57] B. Pradhan and M. I. Sameen, "Applications of deep learning in severity prediction of traffic accidents," in *Laser Scanning Systems in Highway and Safety Assessment*, Advances in Science, Technology & Innovation (IEREK Interdisciplinary Series for Sustainable Development) ,Cham, Switzerland: Springer,

2020, pp. 129–139, 03 April 2019, doi.org/10.1007/978-3-030-10374-3_11

[58] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *2015 IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, Boston, MA, Jun. 7–12, 2015, pp. 815–823.

[59] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in *2014 IEEE Conf. Comput. Vision Pattern Recognit.*, Columbus, OH, Jun. 23–28, 2014, pp. 1701–1708.

[60] T. Soyata, R. Muraleedharan, C. Funai, M. Kwon, and W. Heinzelman, "Cloud-Vision: Real-time face recognition using a mobile-cloudlet-cloud acceleration architecture," in *2012 IEEE Symp. Comput. Commun. (ISCC)*, Cappadocia, Turkey, Jul. 1–4, 2012, pp. 000059–000066.

[61] H. J. Hsu and K. T. Chen, "Face recognition on drones: Issues and limitations," in *Proc. First Workshop Micro Aerial Vehicle Netw. Syst. Appl. Civilian Use*, Florence, Italy, May, 2015, pp. 39–44.

[62] C. Wang, W. Lu, M. R. Narayanan, S. J. Redmond, and N. H. Lovell, "Low-power technologies for wearable telecare and telehealth systems: A review," *Biomed. Eng. Lett.*, vol. 5, no. 1, pp. 1–9, Mar. 2015, doi: 10.1007/s13534-015-0174-2.

[63] L. Ren, W. Shi, Z. Yu, and Z. Liu, "Real-time energy-efficient fall detection based on SSR energy efficiency strategy," *Int. J. Sensor Networks*, vol. 20, no. 4, pp. 243–251, Jan. 2016, doi: 10.1504/ijsnet.2016.076726.

[64] S. S. Khan and J. Hoey, "Review of fall detection techniques: A data availability perspective," *Med. Eng. Phys.*, vol. 39, pp. 12–22, Jan. 2017, doi: 10.1016/j.medengphy.2016.10.014.

[65] Y. Gao and K. M. Mosalam, "Deep transfer learning for image-based structural damage recognition," *Comput. Aided Civil Infrastructure Eng.*, vol. 33, no. 9, pp. 748–768, Apr. 2018, doi: 10.1111/mice.12363.

[66] B. F. Spencer, V. Hoskere, and Y. Narazaki, "Advances in computer vision-based civil infrastructure inspection and monitoring," *Engineering*, vol. 5, no. 2, pp. 199–222, Apr. 2019, doi: 10.1016/j.eng.2018.11.030.

[67] Y. Ren *et al.*, "Image-based concrete crack detection in tunnels using deep fully convolutional networks," *Construction Building Materials*, vol. 234, p. 117367, Feb. 2020, doi: 10.1016/j.conbuildmat.2019.117367.

[68] A. Brunetti, D. Buongiorno, G. F. Trotta, and V. Bevilacqua, "Computer vision and deep learning techniques for pedestrian detection and tracking: A survey," *Neurocomputing*, vol. 300, pp. 17–33, Jul. 2018, doi: 10.1016/j.neucom.2018.01.092.

[69] A. Pramanik, S. Sarkar, and J. Maiti, "Oil spill detection using image processing technique: An occupational safety perspective of a steel plant," in *Emerging Technologies in Data Mining and Information Security*, A. Abraham, P. Dutta, J. K. Mandal, A. Bhattacharya, and S. Dutta, Eds., Singapore: Springer Singapore, 2019, pp. 247–257.

[70] W. Fang *et al.*, "Computer vision applications in construction safety assurance," *Automat. Construction*, vol. 110, p. 103013, Feb. 2020, doi: 10.1016/j.autcon.2019.103013.

[71] M. Venkata and S. Nishant, "Detecting and tracking of humans in an underwater environment using deep learning algorithms," Master´s thesis, Comput. Sci., Blekinge Institute of Technology, Karlskrona, Sweden, 2019.

[72] S. Akcay and T. Breckon, "Towards automatic threat detection: A survey of advances of deep learning within X-ray security imaging," 2020, *arXiv preprint arXiv:2001.01293*.

[73] L. Ding, W. Fang, H. Luo, P. E. D. Love, B. Zhong, and O. Xi, "A deep hybrid learning model to detect unsafe

behavior: Integrating convolution neural networks and long short-term memory," *Automat. Construction*, vol. 86, pp. 118–124, Feb. 2018, doi: 10.1016/j.autcon.2017.11.002.

[74] J. Y. Hwang, J. M. Kang, Y. W. Jang, and H. C. Kim, "Development of novel algorithm and real-time monitoring ambulatory system using Bluetooth module for fall detection in the elderly," in *26th Annu. Int. Conf. IEEE Eng. Medicine Biol. Soc.*, San Francisco, CA, Sep. 1–5, 2004, pp. 2204–2207.

[75] Q. Li, J. A. Stankovic, M. A. Hanson, A. T. Barth, J. Lach, and G. Zhou, "Accurate, fast fall detection using gyroscopes and accelerometer-derived posture information," in *2009 Sixth Int. Workshop Wearable Implantable Body Sensor Networks*, Berkeley, CA, Jun. 3–5, 2009, pp. 138–143.

[76] A. Shahzad and K. Kim, "FallDroid: An automated smart-phone-based fall detection system using multiple kernel learning," *IEEE Trans. Ind. Inform.*, vol. 15, no. 1, pp. 35–44, Jan. 2019, doi: 10.1109/tii.2018.2839749.

[77] T. de Quadros, A. E. Lazzaretti, and F. K. Schneider, "A movement decomposition and machine learning-based fall detection system using wrist wearable device," *IEEE Sensors J.*, vol. 18, no. 12, pp. 5082–5089, Jun. 2018, doi: 10.1109/jsen.2018.2829815.

[78] G. Koshmak, A. Loutfi, and M. Linden, "Challenges and issues in multisensor fusion approach for fall detection: Review paper," *J. Sensors*, vol. 2016, p. 6931789 Dec. 2016, doi: 10.1155/2016/6931789.

[79] F. Bagalà *et al.*, "Evaluation of accelerometer-based fall detection algorithms on real-world falls," *PLoS One*, vol. 7, no. 5, pp. e37062, 2012, doi: 10.1371/journal.pone.0037062.

[80] A. K. Bourke *et al.*, "Evaluation of waist-mounted tri-axial accelerometer based fall-detection algorithms during scripted and continuous unscripted activities," *J. Biomechanics*, vol. 43, no. 15, pp. 3051–3057, Nov. 2010, doi: 10.1016/j.jbiomech.2010.07.005.

[81] M. A. Habib, M. S. Mohktar, S. B. Kamaruzzaman, K. S. Lim, T. M. Pin, and F. Ibrahim, "Smartphone-based solutions for fall detection and prevention: challenges and open issues," *Sensors*, vol. 14, no. 4, pp. 7181–7208, Apr. 2014, doi: 10.3390/s140407181.

[82] G. A. Koshmak, M. Linden, and A. Loutfi, "Evaluation of the android-based fall detection system with physiological data monitoring," in *2013 35th Annu. Int. Conf. IEEE Eng. Medicine Biol. Soc. (EMBC)*, Osaka, Japan, Jul. 3–7, 2013, pp. 1164–1168.

[83] Z. Zhang, C. Conly, and V. Athitsos, "A survey on vision-based fall detection," in *Proc. 8th ACM Int. Conf. Pervasive Technologies Related Assistive Environ.*, Corfu, Greece, Jul., 2015, p. Article 46.

[84] L. Ren and Y. Peng, "Research of fall detection and fall prevention technologies: A systematic review," *IEEE Access*, vol. 7, pp. 77702–77722, Jun. 2019, doi: 10.1109/access.2019.2922708.

[85] U. Asif *et al.*, "Privacy preserving human fall detection using video data," in *Proc. Mach. Learn. Res.*, Vancouver, Canada, Dec. 8–Dec. 14, 2019, pp. 1–12.

[86] A. Shahzad, S. Ko, S. Lee, J. A. Lee, and K. Kim, "Quantitative assessment of balance impairment for fall-risk estimation using wearable triaxial accelerometer," *IEEE Sens. J.*, vol. 17, no. 20, pp. 6743–6751, Oct. 2017, doi: 10.1109/jsen.2017.2749446.

[87] D. Yacchirema, J. S. de Puga, C. Palau, and M. Esteve, "Fall detection system for elderly people using IoT and Big Data," Procedia Comput. Sci., vol. 130, pp. 603–610, 2018, doi: 10.1016/j.procs.2018.04.110.

[88] Y. Su, D. Liu, and Y. Wu, "A multi-sensor based pre-impact fall detection system with a hierarchical classifier," in *2016 9th Int. Congr. Image Signal Process.*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2997946, IEEE Access

IEEE *Access*

Author Name: A Deep Learning Approach Towards Railway Safety Risk Management

BioMed. Eng. Inform. (CISP-BMEI), Datong, China, Oct. 15–17, 2016, pp. 1727–1731.

[89] P. C. Fino, C. W. Frames, and T. E. Lockhart, "Classifying step and spin turns using wireless gyroscopes and implications for fall risk assessments," *Sensors*, vol. 15, no. 5, pp. 10676–10685, May 2015, doi: 10.3390/s150510676.

[90] W. Lu, C. Wang, M. C. Stevens, S. J. Redmond, and N. H. Lovell, "Low-power operation of a barometric pressure sensor for use in an automatic fall detector," in *2016 38th Annu. Int. Conf. IEEE Eng. Medicine Biol. Soc. (EMBC)*, Orlando, FL, Aug. 16–20, 2016, pp. 2010–2013.

[91] X. Xi, M. Tang, S. M. Miran, and Z. Luo, "Evaluation of feature extraction and recognition for activity monitoring and fall detection based on wearable sEMG sensors," *Sensors*, vol. 17, no. 6, p. 1229, May 2017, doi: 10.3390/s17061229.

[92] G. Rescio, A. Leone, and P. Siciliano, "Supervised machine learning scheme for electromyography-based pre-fall detection system," *Expert Syst. Appl.*, vol. 100, pp. 95–105, Jun. 2018, doi: 10.1016/j.eswa.2018.01.047.

[93] J. Sun, Z. Wang, L. Chen, B. Wang, C. Ji, and S. Tao, "A plantar inclinometer based approach to fall detection in open environments," in *Emerging Trends and Advanced Technologies for Computational Intelligence: Extended and Selected Results from the Science and Information Conference 2015*, L. Chen, S. Kapoor, and R. Bhatia, Eds., Cham, Switzerland: Springer International Publishing, 2016, pp. 1–13.

[94] J. Sun, Z. Wang, B. Pei, S. Tao, and L. Chen, "Fall detection using plantar inclinometer sensor," in *2015 IEEE 12th Int. Conf. Ubiquitous Intell. Comput. 2015 IEEE 12th Int. Conf. Autonomic Trusted Comput. 2015 IEEE 15th Int. Conf. Scalable Comput. Commun. Associated Workshops (UIC-ATC-ScalCom)*, Beijing, China, Aug. 10–14, 2015, pp. 1692–1697.

[95] H. Li, M. Lu, S. C. Hsu, M. Gray, and T. Huang, "Proactive behavior-based safety management for construction safety improvement," *Saf. Sci.*, vol. 75, pp. 107–117, Jun. 2015, doi: 10.1016/j.ssci.2015.01.013.

[96] P. E. D. Love, P. Teo, and J. Morrison, "Unearthing the nature and interplay of quality and safety in construction projects: An empirical study," *Saf. Sci.*, vol. 103, pp. 270–279, Mar. 2018, doi: 10.1016/j.ssci.2017.11.026.

[97] O. S. Kwon, C. S. Park, and C. R. Lim, "A defect management system for reinforced concrete work utilizing BIM, image-matching and augmented reality," *Automat. Construction*, vol. 46, pp. 74–81, Oct. 2014, doi: 10.1016/j.autcon.2014.05.005.

[98] M. Liu, S. Han, and S. Lee, "Tracking-based 3D human skeleton extraction from stereo video camera toward an on-site safety and ergonomic analysis," *Construction Innov.*, vol. 16, no. 3, pp. 348–367, Jul. 2016, doi: 10.1108/ci-10-2015-0054.

[99] N. D. Nath, A. H. Behzadan, and S. G. Paal, "Deep learning for site safety: Real-time detection of personal protective equipment," *Automat. Construction*, vol. 112, p. 103085, Apr. 2020, doi: 10.1016/j.autcon.2020.103085.

[100] L. Meng *et al.*, "Real-time detection of ground objects based on unmanned aerial vehicle remote sensing with deep learning: Application in excavator detection for pipeline safety," *Remote Sens.*, vol. 12, no. 1, p. 182, Jan. 2020, doi: 10.3390/rs12010182.

[101] Z. Kolar, H. Chen, and X. Luo, "Transfer learning and deep convolutional neural networks for safety guardrail detection in 2D images," *Automat. Construction*, vol. 89, pp. 58–70, May 2018, doi: 10.1016/j.autcon.2018.01.003.

[102] M. Siddula, F. Dai, Y. Ye, and J. Fan, "Unsupervised feature learning for objects of interest detection in cluttered construction roof site images," *Procedia Eng.*,

vol. 145, pp. 428–435, May 2016, doi: 10.1016/j.proeng.2016.04.010.

[103] Q. Fang *et al.*, "Detecting non-hardhat-use by a deep learning method from far-field surveillance videos," *Automat. Construction*, vol. 85, pp. 1–9, Jan. 2018, doi: 10.1016/j.autcon.2017.09.018.

[104] J. Wu, N. Cai, W. Chen, H. Wang, and G. Wang, "Automatic detection of hardhats worn by construction personnel: A deep learning approach and benchmark dataset," *Automat. Construction*, vol. 106, p. 102894, Oct. 2019, doi: 10.1016/j.autcon.2019.102894.

[105] W. Fang, L. Ding, H. Luo, and P. E. D. Love, "Falls from heights: A computer vision-based approach for safety harness detection," *Automat. Construction*, vol. 91, pp. 53–61, Jul. 2018, doi: 10.1016/j.autcon.2018.02.018.

[106] W. Fang *et al.*, "A deep learning-based approach for mitigating falls from height with computer vision: Convolutional neural network," *Adv. Eng. Inform.*, vol. 39, pp. 170–177, Jan. 2019, doi: 10.1016/j.aei.2018.12.005.

[107] D. Kim, M. Liu, S. Lee, and V. R. Kamat, "Remote proximity monitoring between mobile construction resources using camera-mounted UAVs," *Automat. Construction*, vol. 99, pp. 168–182, Mar. 2019, doi: 10.1016/j.autcon.2018.12.014.

[108] Y. Choi, J. Park, and B. Jang, "A risk estimation approach based on deep learning in shipbuilding industry," in *2019 Int. Conf. Inf. Commun. Technol. Convergence (ICTC)*, Jeju Island, Korea, Oct. 16–18, 2019, pp. 1438–1441.

[109] X. W. Ye, T. Jin, and P. P. Ang, "Computer vision-based monitoring of ship navigation for bridge collision risk assessment," in *Machine Vision and Navigation*, O. Sergiyenko, W. Flores-Fuentes, and P. Mercorelli, Eds., Cham, Switzerland: Springer International Publishing, 2020, pp. 787–807.

[110] W. Fang, L. Ding, B. Zhong, P. E. D. Love, and H. Luo, "Automated detection of workers and heavy equipment on construction sites: A convolutional neural network approach," *Adv. Eng. Inform.*, vol. 37, pp. 139–149, Aug. 2018, doi: 10.1016/j.aei.2018.05.003.

[111] S. Arabi, A. Haghighat, and A. Sharma, "A deep-learning-based computer vision solution for construction vehicle detection," *Comput. Aided Civil Infrastructure Eng.*, pp. 1–15, Jan. 2020, doi: 10.1111/mice.12530.

[112] H. Kim, H. Kim, Y. W. Hong, and H. Byun, "Detecting construction equipment using a region-based fully convolutional network and transfer learning," *J. Comput. Civil Eng.*, vol. 32, no. 2, p. 04017082, Mar. 2018, doi: 10.1061/(asce)cp.1943-5487.0000731.

[113] X. Tong, J. Guo, Y. Ling, and Z. Yin, "A new image-based method for concrete bridge bottom crack detection," in *2011 Int. Conf. Image Anal. Signal Process.*, Hubei, China, Oct. 21–23, 2011, pp. 568–571.

[114] I. Abdel-Qader, O. Abudayyeh, and M. E. Kelly, "Analysis of edge-detection techniques for crack identification in bridges," *J. Comput. Civil Eng.*, vol. 17, no. 4, pp. 255–263, Oct. 2003, doi: 10.1061/(asce)0887-3801(2003)17:4(255).

[115] P. Hüthwohl and I. Brilakis, "Detecting healthy concrete surfaces," *Adv. Eng. Inform.*, vol. 37, pp. 150–162, Aug. 2018, doi: 10.1016/j.aei.2018.05.004.

[116] D. Isailović, V. Stojanovic, M. Trapp, R. Richter, R. Hajdin, and J. Döllner, "Bridge damage: Detection, IFC-based semantic enrichment and visualization," *Automat. Construction*, vol. 112, p. 103088, Apr. 2020, doi: 10.1016/j.autcon.2020.103088.

[117] W. T. Nash, C. J. Powell, T. Drummond, and N. Birbilis, "Automated corrosion detection using crowdsourced training for deep learning," *Corrosion*, vol. 76, no. 2, pp. 135–141, Dec. 2019, doi: 10.5006/3397.

[118] I. Abdel-Qader, S. Pashaie-Rad, O. Abudayyeh, and S. Yehia, "PCA-based algorithm for unsupervised bridge

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2997946, IEEE Access

IEEE Access

Author Names: A Deep Learning Approach Towards Railway Safety Risk Management

crack detection," *Advances Eng. Softw.*, vol. 37, no. 12, pp. 771–778, Dec. 2006, doi: 10.1016/j.advengsoft.2006.06.002.

[119] T. Yamaguchi and S. Hashimoto, "Fast crack detection method for large-size concrete surface images using percolation-based image processing," *Mach. Vision Appl.*, vol. 21, no. 5, pp. 797–809, Feb. 2010, doi: 10.1007/s00138-009-0189-8.

[120] Z. Liu, S. A. Suandi, T. Ohashi, and T. Ejima, "Tunnel crack detection and classification system based on image processing," in *Proc. Vol. 4664, Mach. Vision Appl. Ind. Inspection X*, San Jose, CA, Mar. 8, 2002, pp. 145–152.

[121] H. Moon and J. Kim, "Intelligent crack detecting algorithm on the concrete crack image using neural network," in *28th Int. Symp. Automat. Robot. Construction ISARC 2011*, Seoul, Korea, Jun. 29–Jul. 2, 2011, pp. 1461–1467.

[122] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, "Automatic road crack detection using random structured forests," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 12, pp. 3434–3445, Dec. 2016, doi: 10.1109/tits.2016.2552248.

[123] A. Zhang *et al.*, "Deep learning–based fully automated pavement crack detection on 3D asphalt surfaces with an improved cracknet," *J. Comput. Civil Eng.*, vol. 32, no. 5, p. 04018041, Sep. 2018, doi: 10.1061/(asce)cp.1943-5487.0000775.

[124] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," in *2015 IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Milan, Italy, Jul. 26–31, 2015, pp. 4959–4962.

[125] M. Ukai, "Technology for measurement and diagnosis. Development of image processing technique for detection of tunnel wall deformation using continuously scanned image," *Quart. Rep. RTRI*, vol. 41, no. 3, pp. 120–126, Sep. 2000, doi: 10.2219/rtriqr.41.120.

[126] L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu, "Road crack detection using deep convolutional neural network," in *2016 IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, AZ, 25–28 Sept., 2016, pp. 3708–3712.

[127] C. J. Lin, Y. C. Li, and C. L. Lee, "Using fully convolutional networks for floor area detection," *Sensors Mater.*, vol. 32, no. 1, pp. 159–170, Jan. 2020, doi: 10.18494/sam.2020.2577.

[128] A. Jamshidi *et al.*, "A big data analysis approach for rail failure risk assessment," *Risk Anal.*, vol. 37, no. 8, pp. 1495–1507, May 2017, doi: 10.1111/risa.12836.

[129] T. de Bruin, K. Verbert, and R. Babuska, "Railway track circuit fault diagnosis using recurrent neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 3, pp. 523–533, Mar. 2017, doi: 10.1109/tnnls.2016.2551940.

[130] C. Chellaswamy, M. Krishnasamy, L. Balaji, A. Dhanalakshmi, and R. Ramesh, "Optimized railway track health monitoring system based on dynamic differential evolution algorithm," *Measurement*, vol. 152, p. 107332, Feb. 2020, doi: 10.1016/j.measurement.2019.107332.

[131] A. Jamshidi *et al.*, "A decision support approach for condition-based maintenance of rails based on big data analysis," *Transp. Res. Part C Emerg. Technol.*, vol. 95, pp. 185–206, Oct. 2018, doi: 10.1016/j.trc.2018.07.007.

[132] L. Zhuang, L. Wang, Z. Zhang, and K. L. Tsui, "Automated vision inspection of rail surface cracks: A double-layer data-driven framework," *Transp. Res. Part C Emerg. Technol.*, vol. 92, pp. 258–277, Jul. 2018, doi: 10.1016/j.trc.2018.05.007.

[133] M. C. Nakhaee, D. Hiemstra, M. Stoelinga, and M. van Noort, "The recent applications of machine learning in rail track maintenance: A survey," in *Reliability, Safety, and Security of Railway Systems. Modelling, Analysis, Verification, and Certification*, S. Collart-Dutilleul, T. Lecomte, and A. Romanovsky, Eds., Cham, Switzerland: Springer International Publishing, 2019, pp. 91–105.

[134] X. Gibert, V. M. Patel, and R. Chellappa, "Deep multitask learning for railway track inspection," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 1, pp. 153–164, Jan. 2017, doi: 10.1109/tits.2016.2568758.

[135] F. Maire and A. Bigdeli, "Obstacle-free range determination for rail track maintenance vehicles," in *2010 11th Int. Conf. Control Automat. Robot. Vision*, Singapore, Dec. 7–10, 2010, pp. 2172–2178.

[136] C. Vondrick, H. Pirsiavash, and A. Torralba, "Generating videos with scene dynamics," in *Advances Neural Inf. Process. Syst.*, Barcelona, Spain, Dec. 5–Dec. 10, 2016, pp. 613–621.

[137] R. Montero, J. G. Victores, S. Martínez, A. Jardón, and C. Balaguer, "Past, present and future of robotic tunnel inspection," *Automat. Construction*, vol. 59, pp. 99–112, Nov. 2015, doi: 10.1016/j.autcon.2015.02.003.

[138] K. Makantasis, E. Protopapadakis, A. Doulamis, N. Doulamis, and C. Loupos, "Deep convolutional neural networks for efficient vision based tunnel inspection," in *2015 IEEE Int. Conf. Intell. Comput. Commun. Process. (ICCP)*, Cluj-Napoca, Romania, Sep. 3–5, 2015, pp. 335–342.

[139] M. Etxeberria-Garcia, M. Labayen, M. Zamalloa, and N. Arana-Arexolaleiba, "Application of computer vision and deep learning in the railway domain for autonomous train stop operation," in *2020 IEEE/SICE Int. Symp. Syst. Integration (SII)*, Honolulu, HI, Jan. 12–15, 2020, pp. 943–948.

[140] T. Ye, X. Zhang, Y. Zhang, and J. Liu, "Railway traffic object detection using differential feature fusion convolution neural network," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–13, Feb. 2020, doi: 10.1109/tits.2020.2969993.

[141] H. Alawad, S. Kaewunruen, and M. An, "Learning from accidents: Machine learning for safety at railway stations," *IEEE Access*, vol. 8, pp. 633–648, Jan. 2020, doi: 10.1109/access.2019.2962072.

[142] N. Taheri, F. M. Nejad, and H. Zakeri, "A brief overview and new knowledge based system for rail direct fastening evaluation using digital image processing," *Archives Comput. Methods Eng.*, Feb. 2019, doi: 10.1007/s11831-019-09325-z.

[143] L. Liu, F. Zhou, and Y. He, "Automated visual inspection system for bogie block key under complex freight train environment," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 1, pp. 2–14, Jan. 2016, doi: 10.1109/tim.2015.2479101.

[144] H. Yin, Z. Liu, Z. Xu, and L. Gao, "An automatic visual monitoring system for expansion displacement of switch rail," *IEEE Trans. Instrum. Meas.*, Jul. 2019, doi: 10.1109/tim.2019.2927547.

[145] S. Liu, L. Yu, and D. Zhang, "An efficient method for high-speed railway dropper fault detection based on depthwise separable convolution," *IEEE Access*, vol. 7, pp. 135678–135688, Sep. 2019, doi: 10.1109/access.2019.2942079.

[146] Z. Liu, Y. Lyu, L. Wang, and Z. Han, "Detection approach based on an improved faster RCNN for brace sleeve screws in high-speed railways," *IEEE Trans. Instrum. Meas.*, Sep. 2019, doi: 10.1109/tim.2019.2941292.

[147] F. Zhou, Y. Song, L. Liu, and D. Zheng, "Automated visual inspection of target parts for train safety based on deep learning," *IET Intell. Transp. Syst.*, vol. 12, no. 6, pp. 550–555, Aug. 2018, doi: 10.1049/iet-its.2016.0338.

[148] Y. Zhang, M. Liu, Y. Chen, H. Zhang, and Y. Guo, "Real-time vision based system of fault detection for freight trains," *IEEE Trans. Instrum. Meas.*, Nov. 2019, doi: 10.1109/tim.2019.2955799.

[149] L. Chang, Z. Liu, and Y. Shen, "On-line detection of pantograph offset based on deep learning," in *2018 IEEE 3rd Optoelectronics Global Conf.*, Shenzhen, China, Sep. 4–7, 2018, pp. 159–164.

[150] R. L. Rocha *et al.*, "Image inspection of railcar structural components: An approach through deep learning and discrete fourier transform," in *Anais VII Symp. Knowl. Discovery Mining Learn.*, Porto Alegre, Brazil, Nov., 2019, pp. 33–40.

[151] Y. Shen, Z. Liu, and G. Zhang, "PAC interaction inspection using real-time contact point tracking," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 10, pp. 4051–4064, Oct. 2019, doi: 10.1109/tim.2018.2884039.

[152] R. Guo, R. Liu, N. Li, and W. Liu, "DV3+HED+: a DCNN-based framework to monitor temporary works and ESAs in railway construction project using VHR satellite images," *J. Appl. Remote Sens.*, vol. 14, no. 1, p. 014507, Jan. 2020, doi: 10.1117/1.jrs.14.014507.

[153] H. Liu, D. Yao, J. Yang, and X. Li, "Lightweight convolutional neural network and its application in rolling bearing fault diagnosis under variable working conditions," *Sensors*, vol. 19, no. 22, p. 4827, Nov. 2019, doi: 10.3390/s19224827.

[154] J. R. Edwards *et al.*, "Development of machine vision technology for railcar safety appliance inspection," in *Proc. Int. Heavy Haul Conf. Specialist Tech. Session-High Tech Heavy Haul, Kiruna, Sweden*, Kiruna, Sweden, 2007, pp. 745–752.

[155] Z. Wang, X. Wu, G. Yu, and M. Li, "Efficient rail area detection using convolutional neural network," *IEEE Access*, vol. 6, pp. 77656–77664, Nov. 2018, doi: 10.1109/access.2018.2883704.

[156] T. Ye, B. Wang, P. Song, and J. Li, "Automatic railway traffic object detection system using feature fusion refine neural network under shunting mode," *Sensors*, vol. 18, no. 6, p. 1916, Jun. 2018, doi: 10.3390/s18061916.

[157] C. Huang and Y. Zeng, "The fault diagnosis of catenary system based on the deep learning method in the railway industry," in *Proc. 5th Int. Conf. Multimedia Image Process.*, New York, NY, Jan. 10, 2020, pp. 135–140.

[158] W. Liu, Z. Liu, A. Nunez, and Z. Han, "Unified deep learning architecture for the detection of all catenary support components," *IEEE Access*, vol. 8, pp. 17049–17059, Jan. 2020, doi: 10.1109/access.2020.2967831.

[159] Y. Chen, B. Song, X. Du, and N. Guizani, "The enhancement of catenary image with low visibility based on multi-feature fusion network in railway industry," *Comput. Commun.*, vol. 152, pp. 200–205, Feb. 2020, doi: 10.1016/j.comcom.2020.01.040.

[160] M. Molodova, M. Oregui, A. Núñez, Z. Li, and R. Dollevoet, "Health condition monitoring of insulated joints based on axle box acceleration measurements," *Eng. Struct.*, vol. 123, pp. 225–235, Sep. 2016, doi: 10.1016/j.engstruct.2016.05.018.

[161] S. Mittal and D. Rao, "Vision based railway track monitoring using deep learning," 2017, *arXiv preprint arXiv:1711.06423.*

[162] M. Nieniewski, "Morphological detection and extraction of rail surface defects," *IEEE Trans. Instrum. Meas.*, p. 1, Feb. 2020, doi: 10.1109/tim.2020.2975454.

[163] S. Doyle, C. Bastucescu, and T. Vale, "Pantograph condition monitoring system for automated maintenance inspections and prevention of overhead wiring tear downs," in *CORE 2016 Maintaining Momentum Conf. Railway Excellence*, Melbourne, Victoria, May 16–18, 2016, p. 178.

[164] S. Ritika, S. Mittal, and D. Rao, "Railway track specific traffic signal selection using deep learning," 2017, *arXiv preprint arXiv:1712.06107.*

[165] E. Choodowicz, P. Lisiecki, and P. Lech, "Hybrid algorithm for the detection and recognition of railway signs," in *Progress in Computer Recognition Systems*, R. Burduk, M. Kurzynski, and M. Wozniak, Eds., Cham, Switzerland: Springer International Publishing, 2020, pp. 337–347.

[166] P. Huang, C. Wen, L. Fu, Q. Peng, and Y. Tang, "A deep learning approach for multi-attribute data: A study of train delay prediction in railway systems," *Inf. Sci.*, vol. 516, pp. 234–253, Apr. 2020, doi: 10.1016/j.ins.2019.12.053.

[167] Q. Han *et al.*, "A rail fastener tightness detection approach using multi-source visual sensor," *Sensors*, vol. 20, no. 5, p. 1367, Mar. 2020, doi: 10.3390/s20051367.

[168] X. Wei, Z. Yang, Y. Liu, D. Wei, L. Jia, and Y. Li, "Railway track fastener defect detection based on image processing and deep learning techniques: A comparative study," *Eng. Appl. Artif. Intell.*, vol. 80, pp. 66–81, Apr. 2019, doi: 10.1016/j.engappai.2019.01.008.

[169] F. Ghofrani, Q. He, R. M. P. Goverde, and X. Liu, "Recent applications of big data analytics in railway transportation systems: A survey," *Transp. Res. Part C Emerg. Technol.*, vol. 90, pp. 226–246, May 2018, doi: 10.1016/j.trc.2018.03.010.

[170] A. Berry, B. Nejikovsky, X. Gilbert, and A. Tajaddini, "High speed video inspection of joint bars using advanced image collection and processing techniques," in *Proc. 2008 World Congr. Railway Res.*, Seoul, Korea, Oct. 28–Nov. 1, 2008.

[171] J. Liu, L. Qian, Y. Zhang, J. Han, and J. Sun, "Towards safety-risk prediction of CBTC systems with deep learning and formal methods," *IEEE Access*, vol. 8, pp. 16618–16626, Jan. 2020, doi: 10.1109/access.2020.2967634.

[172] X. Xu, Y. Lei, and F. Yang, "Railway subgrade defect automatic recognition method based on improved faster R-CNN," *Scientific Program.*, vol. 2018, pp. 1–12, Jun. 2018, doi: 10.1155/2018/4832972.

[173] Y. Wang, S. Hu, G. Wang, C. Chen, and Z. Pan, "Multi-scale dilated convolution of convolutional neural network for crowd counting," *Multimedia Tools Appl.*, vol. 79, no. 1-2, pp. 1057–1073, Oct. 2019, doi: 10.1007/s11042-019-08208-6.

[174] B. Guo, G. Geng, L. Zhu, H. Shi, and Z. Yu, "High-speed railway intruding object image generating with generative adversarial networks," *Sensors*, vol. 19, no. 14, p. 3075, Jul. 2019, doi: 10.3390/s19143075.

[175] N. Cai, H. Chen, Y. Li, and Y. Peng, "Intrusion detection and tracking at railway crossing," in *Proc. 2019 Int. Conf. Artif. Intell. Adv. Manuf. - AIAM 2019*, Dublin, Ireland, Oct., 2019, p. 57.

[176] S. A. Sumon, M. T. Shahria, M. R. Goni, N. Hasan, A. M. Almarufuzzaman, and R. M. Rahman, "Violent Crowd Flow Detection Using Deep Learning," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2019, vol. 11431 LNAI, pp. 613–625, doi: 10.1007/978-3-030-14799-0_53.

[177] N. Bakalos *et al.*, "Protecting Water Infrastructure from Cyber and Physical Threats: Using Multimodal Data Fusion and Adaptive Deep Learning to Monitor Critical Systems," *IEEE Signal Process. Mag.*, vol. 36, no. 2, pp. 36–48, 2019, doi: 10.1109/MSP.2018.2885359.

[178] S. Akcay and T. Breckon, "Towards Automatic Threat Detection: A Survey of Advances of Deep Learning within X-ray Security Imaging," arXiv preprint arXiv:2001.01293 ,Jan. 2020.

[179] S. Liu, Q. Wang, and Y. Luo, "A review of applications of visual inspection technology based on image processing in the railway industry," *Transp. Saf. Environ.*, vol. 1, no. 3, pp. 185–204, Dec. 2019, doi: 10.1093/tse/tdz007.

[180] K. Ljubymenko and M. Adamek, "Security personnel of new generation," in *2018 Int. Carnahan Conf. Secur. Technol. (ICCST)*, Montreal, QC, Oct. 22–25, 2018, pp. 1–4.

[181] J. Seo and I.-J. Chung, "Face Liveness Detection Using Thermal Face-CNN with External Knowledge," mdpi.com, 2019, doi: 10.3390/sym11030360..

[182] A. M. Rakoczy, S. T. Wilk, and M. Jones, "Security and safety of rail transit tunnels," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2673, no. 1, pp. 92–101, Jan. 2019, doi: 10.1177/0361198118822819.

[183] Y. Aytar, C. Vondrick, and A. Torralba, "Soundnet: Learning sound representations from unlabeled video," in *Advances Neural Inf. Process. Syst.*, Barcelona, Spain, Dec. 5–Dec. 10, 2016, pp. 892–900.

[184] A. van den Oord *et al.*, "Wavenet: A generative model for raw audio," 2016, *arXiv preprint arXiv:1609.03499*.

[185] Mohana and H. V. R. Aradhya, "Simulation of object detection algorithms for video survillance applications," in *2018 2nd Int. Conf. I-SMAC (IoT Soc. Mobile Analytics Cloud) (I-SMAC)I-SMAC (IoT in Soc. Mobile Analytics Cloud) (I-SMAC), 2018 2nd Int. Conf.*, Palladam, India, Aug. 30–31, 2018, 2018, pp. 651–655.

[186] X. Zhang *et al.*, "Deep fusion of multiple semantic cues for complex event recognition," *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1033–1046, Mar. 2016, doi: 10.1109/tip.2015.2511585.

[187] W. Wang, Y. Yan, S. Winkler, and N. Sebe, "Category specific dictionary learning for attribute specific feature selection," *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1465–1478, Mar. 2016, doi: 10.1109/tip.2016.2523340.

[188] S. Ronaghan. "Deep learning: Overview of neurons and activation functions." https://medium.com/@srnghn/deep-learning-overview-of-neurons-and-activation-functions-1d98286cf1e4 (accessed Jul. 1, 2019).

[189] J. Thrane, D. Zibar, and H. L. Christiansen, "Model-aided deep learning method for path loss prediction in mobile communication systems at 2.6 GHz," *IEEE Access*, vol. 8, pp. 7925–7936, Jan. 2020, doi: 10.1109/access.2020.2964103.

[190] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 221–231, Jan. 2013, doi: 10.1109/tpami.2012.59.

[191] G. Chandan, A. Jain, H. Jain, and Mohana, "Real time object detection and tracking using deep learning and openCV," in *2018 Int. Conf. Inventive Res. Comput. Appl. (ICIRCA)*, Coimbatore, India, Jul. 11–12, 2018, pp. 1305–1308.

[192] J. Krüger and R. Westermann, "Linear algebra operators for GPU implementation of numerical algorithms," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 908–916, Jul. 2003, doi: 10.1145/882262.882363.

[193] M. H. Jarrahi, "Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making," *Bus. Horiz.*, vol. 61, no. 4, pp. 577–586, Jul. 2018, doi: 10.1016/j.bushor.2018.03.007.

[194] M. Walport and M. Sedwill. "Artificial intelligence: opportunities and implications for the future of decision making. Assets.publishing.service.gov.uk." https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/566075/gs-16-19-artificial-intelligence-ai-report.pdf (accessed Jul. 1, 2019).

[195] Y. Tian, G.-H. Lee, H. He, C.-Y. Hsu, and D. Katabi, "RF-based fall monitoring using convolutional neural networks," *Proc. ACM Interactive Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 3, pp. 1–24, Sep. 2018, doi: 10.1145/3264947.

[196] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances Neural Inf. Process. Syst.*, Lake Tahoe, Nevada, Dec. 3–6, 2012, pp. 1097–1105.

[197] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in *2014 IEEE Conf. Comput. Vision Pattern Recognit.*, Columbus, OH, Jun. 23–28, 2014, pp. 1725–1732.

[198] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *2017 IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, Honolulu, HI, Jul. 21–26, 2017, pp. 1800–1807.

[199] S. Kaewunruen, JM Sussman and HH Einstein, Strategic framework to achieve carbon-efficient construction and maintenance of railway infrastructure systems. Front. Environ. Sci. 3:6, 2015. doi: 10.3389/fenvs.2015.00006

[200] I. Charfi, J. Miteran, J. Dubois, M. Atri, and R. Tourki, "Optimized spatio-temporal descriptors for real-time fall detection: Comparison of support vector machine and Adaboost-based classification," *J. Electron. Imag.*, vol. 22, no. 4, p. 041106, Jul. 2013, doi: 10.1117/1.jei.22.4.041106.

[201] S. Kaewunruen, J. M. Sussman, and A. Matsumoto, "Grand challenges in transportation and transit systems," *Frontiers Built Environ.*, vol. 2, p. 4, Feb. 2016, doi: 10.3389/fbuil.2016.00004.