



## Article

# Health Monitoring of Conveyor Belt Using UHF RFID and Multi-Class Neural Networks

Fatema Tuz Zohra <sup>1,\*</sup>, Omar Salim <sup>2</sup>, Hossein Masoumi <sup>3</sup> , Nemai C. Karmakar <sup>1</sup> and Shuvashis Dey <sup>4</sup> <sup>1</sup> Department of Electrical & Computer Systems Engineering, Monash University, Melbourne, VIC 3800, Australia<sup>2</sup> School of Engineering & Information Technology, University of New South Wales, Canberra, ACT 2610, Australia<sup>3</sup> Department of Civil Engineering, Monash University, Melbourne, VIC 3800, Australia<sup>4</sup> Department of Electrical and Computer Engineering, North Dakota State University, Fargo, ND 58105, USA

\* Correspondence: fatema-tuz.zohra@monash.edu

**Abstract:** Conveyor belts in mining sites are prone to cracks, which leads to dramatic degradation of overall system performance and the breakdown of operation. Crack detection using radio frequency identification (RFID) sensing technology is recently proposed to provide robust and low-cost health monitoring systems for conveyor belts. The intelligent machine learning (ML) technique is one of the most promising solutions for crack detection and successful implementation within the IoT paradigm. This paper presents a conveyor belt structural health monitoring (SHM) model using ML and Internet of Things (IoT) connectivity. The model is extensively tested, and the classification is conducted based on simulated data obtained from an Ultra High Frequency (UHF) RFID sensor. Here, the sensor is laid on a belt, and the data are obtained at different crack orientations of vertical, horizontal, and diagonal cracks, for varying crack widths of 0.5 to 5 mm at 10 different locations on the sensor. The ML model is tested with different input features and training algorithms, and their performances are compared and analysed to identify the superior input feature and training algorithm. This method produces high accuracy in determining crack width, orientation, and location. The findings show that the proposed detection system based on ML modelling could detect cracks with 100% accuracy. The proposed system can also distinguish between vertical, horizontal, and diagonal cracks with an accuracy of 83.9%, and has a significant identification rate of 84.4% accuracy for detecting crack-width as narrow as 0.5 mm. Moreover, the model can predict the region of the crack with an accuracy of 95.5%. Overall, the results show that the proposed model is very robust and can perform SHM of conveyor belts with high accuracy for a range of parameters and classification scenarios. The method has huge industrial significance in coal mines.



**Citation:** Zohra, F.T.; Salim, O.; Masoumi, H.; Karmakar, N.C.; Dey, S. Health Monitoring of Conveyor Belt Using UHF RFID and Multi-Class Neural Networks. *Electronics* **2022**, *11*, 3737. <https://doi.org/10.3390/electronics11223737>

Academic Editor: Andrea Boni

Received: 16 September 2022

Accepted: 7 November 2022

Published: 15 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** conveyor belt health monitoring; crack detection; UHF RFID sensor; machine learning

## 1. Introduction

Structural health monitoring (SHM) is a vital mechanism to ensure the safe and reliable operation of macroscale essential structures such as bridges, dams, and mining assets such as transporting conveyor belts. Although these structures are expected to operate with high reliability at different loading scenarios within their safety confines, the prolonged exposure to harsh environmental conditions and nonstop operational consignments causes deterioration of the structures and terminates their operations [1]. Such inevitable degradation in operational services is expected due to the introduction of fatigue and cracks. Hence, large-scale structures are always in need of ongoing inspection and maintenance [2]. Smart sensing using passive RFID sensors can monitor and repair structural deteriorations and manage urgent attention requirements [3].

The conveyor belt is the main asset in the mining production process. The economic benefits of mining are mainly dependent on the conveyor belt's healthy operation. Although

steel cord is usually used in belt compositions to increase their tensile strength, the conveyor belt still suffers from tears due to compression, contact with sharp foreign materials, manufacture failure, etc. [2].

Since belt tearing happens frequently, numerous contact and wireless-based damage detection methods have been researched in the literature. Some contact detection methods are linear detection, oscillating roller detection, vibration detection, pressure detection, and bandwidth detection [4]. Contact-based detection is faster, but it can lead to operational interference and inaccurate detection. Non-contact detection methods studied in the literature are electromagnetic induction detection, ultrasonic detection, X-ray detection, and image processing technology (IPT) [5]. However, the extreme environmental conditions of coal mines can limit these methods' reliability. Another method of induction coil sensors based on Eddy current-based testing (ECT) or magnetic loop is also widely employed for the detection of structural cracks, it is not appropriate in the case of non-conductive material like conveyor belts [6,7]. In cases where it is adapted to monitor conveyor belt health, it is fragile and can produce false positives [8]. RFID technology-based sensors are recently adopted in realizing SHM together with the IoT-based smart environment vision [9]. The term Internet of Things (IoT) was first devised for supply chain management [10]. However, with the rise of crossover technology throughout the past decade, the IoT term has evolved to be an umbrella that includes a variety of applications related to industries, health, logistics and utilities, and large-scale infrastructure health monitoring [11]. The objective of IoT is to create a network of interconnected entities that can retrieve (sense) information from the environment, communicate with the World, and utilize the current Internet-based services to deliver application, analysis, and control, all within the realm of the virtual infrastructure provided by cloud computing. Kevin Ashton, the first person to coin the term IoT, believed RFID was a prerequisite for the Internet of Things [10,12]. Much existing content in the Internet of Things has RFID as the foundation and networking core of the construction. For our proposed conveyor belt monitoring, it requires the integration of sensors into the belt and typically a reading distance of 20 cm is sufficient to extract the sensing information. The extracted sensing information can easily be collected, transferred, and processed in a central processing hub or operators' dashboard by using a simple wireless sensor network made of LoRaWan or mesh network topology. Thus, it can be said that RFID is the core technology of IoT [13]. RFID is economical, can communicate wirelessly, and can be designed to have sensing abilities. Extensive research has been conducted to make RFID tags immensely robust [14]. They are frequently used in the supply chain, access/security, and retail ID industry for their flexibility [15] and can have a reliable operation in harsh environments. Moreover, the RFID sensors can be made from flexible materials and embedded inside the belt layers to provide protection to the sensors and avoid any damage during the installation process and the redirection process of the conveyor rollers [16,17]. Apart from its usual implementation in identification and tracking, the tag response can represent physical properties such as strain, fatigue, and humidity. Moreover, passive RFID sensors are wirelessly powered to provide sensing results without needing further electronic devices or batteries [9]. The integration of RFID technology with a Wireless Sensor Network (WSN) can create a collective sensing system that can interpret real-world physical parameters and play a core role in the shaping of smart sensing within the IoT framework.

In recent years, ML has emerged as a powerful technique to predict the health status of large structures. ML has various advantages compared to the traditional prediction methods such as flexibility with large data sets, enhancing detection by eliminating irrelevant input features, adaptability, and prediction ability of nonlinear behavior [18]. Hence, ML is used to detect the structural health of civil constructions such as building columns [19], asphalt road pavement [20], steel pipes [21], and gas turbines [2]. However, most of the studies conducted the machine learning method on images acquired [22], and other studies are based on vibration data. However, the consolidated process in [22] requires large preprocessing and installation steps. Moreover, the system in [11] sometimes mistakenly

identifies dents or bends as cracks. In the case of mining conveyor belts that run for several miles and can suffer from many environmental factors such as dust and external light, crack detection based on image approach is infeasible. Unlike the switching principle, RFID sensors produce a response (RSSI—Received Signal Strength Indicator) at all times, irrespective of the health status of the belt [23]. There is never a blank moment. The introduction of cracks alters the RSSI response and each crack type demonstrates individual patterns, which can then be mined/learned by machine learning [24]. It has to be noted that although the change happening in RSSI response due to the cracks are distinct, there might be overlapping in values between the different types of crack types, locations, and widths. Such decision boundaries and minute variations cannot be characterized and detected without machine learning-based classification.

In this paper, we propose a novel concept, design, and training approach for an artificial neural network (ANN) based on data obtained from the UHF RFID sensor for conveyor belt crack detection. This includes the exclusive collaboration of techniques to ensure high detection accuracy. Simultaneously, the proposed approach can be integrated into the IoT paradigm of smart sensing. However, to be able to fuse ANN classification with UHF RFID-based sensors into the high-profile IOT notion, a robust classifier needs to be designed. The robustness and validity of a classifier are reliable on several intrinsic machine learning parameters. Thus, our second goal is to diversify our designed network's performance by analyzing and comparing it over a range of training algorithms, and different input data parameters. The unique model uses an array of different data obtained by simulation of the UHF RFID sensor and belt. The model is robust enough to produce a very high accuracy rate without requiring complex input parameters. The ANN is capable of predicting three different types of crack orientations. Moreover, it can predict the region of the crack and classify crack widths even when they are as narrow as 0.5 mm. So, our proposed system can classify among several cases. The experimental results show that our proposed method produces crack detection with 100% of accuracy.

The paper is organized as follows: the related literature and studies have been reviewed to provide a succinct summary of existing research. It is proceeded by the detailed theoretical background of UHF RFID sensing principle, ML principles, and the training algorithms implemented in this article. Then the methodology adopted in this paper is described, followed by a projected layout of the integration of the designed model and IoT concept. After that, the experimental details are reported. Later, the results and analysis are discussed in detail and the study is concluded by summarizing the findings and inferences.

## 2. Related Works

IoT has great potential and capacity to simplify a complex system by integrating several sophisticated applications. At a preliminary level, IoT is a system of connected objects with unique identifiers such as RFID tags which are also used as sensors. Here, many objects are connected to the Internet and share data among themselves, such as mobile phones, home appliances, city infrastructures, medical instruments, and plants equipped with sensors [25]. Recent rapid development in the WSN and sensing technologies such as RFID-based sensors is a great boost to IoT-based environmental and structural health monitoring systems. The SHM is integrated into IoT by deploying multiple sensors across the structure, and the sensing data are shared with the processing centres to extract the physical parameters such as the crack's width, and determine the current status of the structure in real time. Then, the operators can remotely access, monitor, and control the operation with the Internet [11]. Such an integrated approach of SHM and IoT enhances reconfiguration flexibility, connecting many platform devices and providing remote access to the operations.

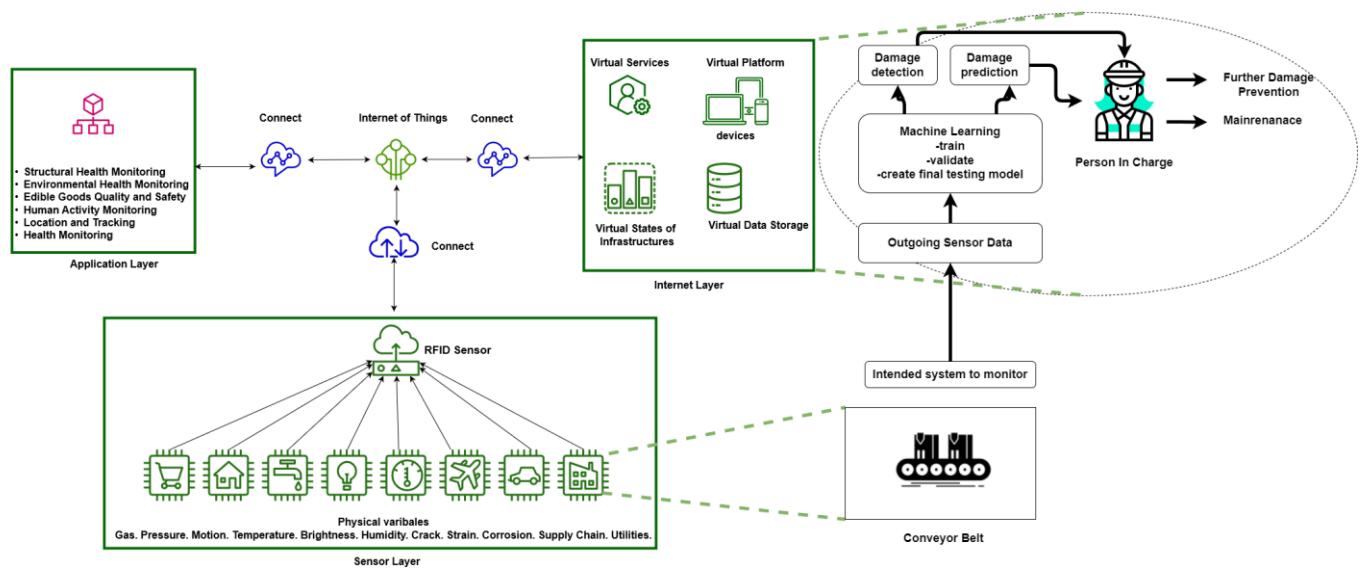
In the mining industry, maintenance requirements of the assets can save costs. For instance, for Goldcorp, a mining company, utilizing gigantic machines to transport materials can incur a loss of two million dollars per day if these machines have a breakdown. Now, the company can save huge costs by applying ML technology to predict its equipment's health

with more than 90% accuracy [26]. SHM can draw immense benefits from the predictive capabilities of machine learning. More specifically, machine learning algorithms can learn a system's typical behaviour from sensors' data and identify when something out of the ordinary begins to happen [27]. The potential of integrating multiple domain applications into IoT increases the amount of data collected from smart structures, which is hard to handle with traditional monitoring systems. Here, machine learning can work together with big data technologies to store and process the monitoring data. ML-based SHM is more suitable for IoT because it can convert simple sensor readings to large-scale industrial actions such as scheduling maintenance. The most useful feature of ML for IoT is that it can develop its prediction efficiency by repeated learning and feedback algorithms [28]. It is worth mentioning that IoT system based on human-generated input is prone to errors, which can also be easily eliminated by ML technology.

RFID sensors can be utilized with ML to create an embedded wireless system with many significant features such as low-cost, and reduced human or manual intervention to maximize the possibility of detecting structural deformation. RFID technology is mainly built up of RFID tags and readers. The UHF tag has a small microchip and an antenna to transfer data via radio waves to the reader, where the data in the microchip (IC) represents the tag's ID and the reader collects and processes the RSSI from the sensors. Here, the RSSI is varied according to the sensing results. The reader also provides the energy required to power the passive RFID tags [29].

This flexible SHM system would be an ideal fit for the services-oriented IoT architecture visioned in Figure 1. For the mining industry, the crack sensing method using RFID sensors and ML can overcome the computational and energy constraints of the IoT domain because ML has the ability to perform such complex computations at reduced energy requirements [26]. The SHM system incorporating IoT consists of data acquired from the sensors that are connected to an external platform. Here, the monitoring centre is automatically alerted when a crack is identified by ML. This allows the person in charge to access this information from anywhere across a range of devices. Tables 1 and 2 summarize crack detection methodologies based on the use of ML and other Non-ML methods presented in the literature, respectively. As shown in Table 1, most of the ML-based detection techniques use images as their inputs and involve civil structures such as roads, buildings, and concrete. Performing identification using images requires complex feature extraction methods leading to high memory usage and lengthy parameters. There is always the requirement to finely detect the underlying damage from impeding factors such as shadow and resolution. Table 2 also shows that the other sensing methodologies using RFID techniques rely on new tag design and the computation of new parameters, and are not implemented on materials such as conveyor belts, despite the SHM of conveyor belts being an immense industrial requirement.

The novelty of detecting the crack pattern of conveyor belt cracks with respect to the cracks' orientations, widths, and locations cannot be achieved without using machine learning. Here, our paper is the first publication in the literature that discusses the various sensor features that can be used for the machine learning training algorithms and also analyses the algorithm which is suitable to provide highly accurate results in determining the crack pattern. Thus, according to our best knowledge, a crack detection for the conveyor belt using RFID based sensor with machine learning is still not addressed in the literature.



**Figure 1.** Envisioned IoT centric Machine Learning based UHF RFID SHM system.

**Table 1.** Summary of Existing SHM Systems Based on ML-Based Techniques.

Study/Year	Features	Disadvantages
Detection of Conveyor Belt longitudinal tears [30] 2019	<ul style="list-style-type: none"> <li>- Visual detection using Visible and Mid-Infrared images.</li> <li>- Accuracy rate 92.04%.</li> </ul>	<ul style="list-style-type: none"> <li>- Works with steel cord conveyor belt.</li> <li>- Cannot detect upper and lower surface tears in a single run.</li> <li>- Actual installation requires continuous dust and water cleaning for the system to operate successfully.</li> </ul>
Surface crack detection of gas turbines [2] 2020	<ul style="list-style-type: none"> <li>- IPT as the pre-processing step followed by convolutional neural network (CNN).</li> <li>- Accuracy rate 96.26%.</li> </ul>	<ul style="list-style-type: none"> <li>- Input data sent to CNN needs to be shuffled so that it consists of more non-crack data.</li> <li>- Does not cover all crack types and angles.</li> <li>- Smoothing might lower the chances of detecting small cracks.</li> <li>- Requires data annotation.</li> </ul>
Multi-surface damage detection of conveyor belt [31] 2019	<ul style="list-style-type: none"> <li>- Requires a focused light source to illuminate while capturing images.</li> <li>- Image Features are extracted using the visual saliency method.</li> <li>- Tested in ideal and wet conditions.</li> </ul>	<ul style="list-style-type: none"> <li>- Undamaged images are not included in the modelling.</li> <li>- Training and testing conducted with only damaged data can cause overfitting.</li> <li>- Cracks misclassified as tears.</li> </ul>
Asphalt pavement crack detection [20] 2017	<ul style="list-style-type: none"> <li>- Images are obtained by Mobile mapping system, then transformed to pavement surface image.</li> <li>- Uses Graphical User Interface (GUI) tool to train and test the SVM model.</li> </ul>	<ul style="list-style-type: none"> <li>- Hard to detect cracks located on the far side of the camera.</li> <li>- Low transverse crack detection rate.</li> <li>- Requires manual evaluation of images for quantitative analysis,</li> </ul>
Asphalt pavement crack detection [22] 2017	<ul style="list-style-type: none"> <li>- Precision 90.13%.</li> <li>- Input Image is captured using 3D laser imaging technology.</li> </ul>	<ul style="list-style-type: none"> <li>- Time-consuming.</li> <li>- Only training the model took 9 days.</li> <li>- Has difficulty detecting hair-line cracks.</li> <li>- Confuses pavement edges with cracks.</li> <li>- Works with 3D data.</li> <li>- Requires modification for other data types.</li> </ul>



Table 1. Cont.

Study/Year	Features	Disadvantages
Building column crack detection [19] 2020	<ul style="list-style-type: none"> <li>- Can detect Horizontal and Vertical cracks.</li> <li>- Acoustic features are captured using Microwave sensor.</li> <li>- Crack Error Rate: - 0.1% (No Crack) - 0.2% (Horizontal and Vertical Crack).</li> </ul>	<ul style="list-style-type: none"> <li>- Requires denoising.</li> <li>- Acoustic features are unsuitable for application in another structural crack detection.</li> </ul>
Conveyor Belt damage detection [32] 2021	<ul style="list-style-type: none"> <li>- Can detect surface wear, surface damage, breakdown, and tear.</li> <li>- Experimentation was conducted with 3000 images.</li> <li>- verify the generalization ability of the proposed model.</li> <li>- Accuracy of 89.12% with the proposed Improved-Yolov3 algorithm.</li> </ul>	<ul style="list-style-type: none"> <li>- Graphics memory overflow occurs.</li> <li>- Light and surface debris might degrade images.</li> </ul>

Table 2. Summary of Existing SHM Systems Based on Non-ML Based Techniques.

Study/Year	Feature	Disadvantages
UHF RFID sensor-based crack analysis of Aluminium beam [33] 2018	<ul style="list-style-type: none"> <li>- Impedance and Reflection coefficient</li> <li>- Circular patch antenna with opening in the middle as sensor.</li> <li>- Sensing is conducted by calibrating the change in resonant frequency with variations in crack depth.</li> </ul>	<ul style="list-style-type: none"> <li>- Suitable for metal structures only.</li> <li>- Requires calibration after every installation.</li> </ul>
Enhancing Low Frequency (LF) RFID sensing system to detect defective metal [34] 2018	<ul style="list-style-type: none"> <li>- Time-Frequency representation of tag response.</li> <li>- Sweep frequency reader is developed to capture the altered resonant frequency occurring due to cracks, wireless environment, and coupling between tag, reader, and defect.</li> <li>- Effect of distance is suppressed.</li> </ul>	<ul style="list-style-type: none"> <li>- The resonant frequency shift occurring due to the metal has to be known beforehand.</li> <li>- During sweeping, cracks cannot be determined when extracted features overlap.</li> </ul>
UHF RFID tag split into two sections to sense cracks in bridges [35] 2013	<ul style="list-style-type: none"> <li>- Power.</li> <li>- The tag is split into two components</li> <li>- tag antenna and moveable chip loop.</li> <li>- Calibration curve is generated based on varying distances and corresponding power.</li> </ul>	<ul style="list-style-type: none"> <li>- Can only work with displacements occurring in a straight line in one direction.</li> <li>- Does not work in orthonormal planes. Function is dependent on the power delivered by the reader and tag antenna.</li> <li>- Chip loop will operate if displacement is too large.</li> </ul>
UHF RFID sensor detect cracks in Ultra High Performance Concrete (UHPC) [36] 2019	<ul style="list-style-type: none"> <li>- Backscattered power.</li> <li>- Tag is directly attached to the surface by removing the substrate between them.</li> <li>- Tested using both fixed and handheld readers to test practical applications.</li> <li>- Tested on UHPC specimens with and without metal fibres.</li> </ul>	<ul style="list-style-type: none"> <li>- Each experiment conducted in identical controlled settings, but RSSI will be affected by environment in real life.</li> <li>- Small detection area.</li> <li>- UHPC structure only.</li> </ul>

### 3. Theory

#### 3.1. Principle of RFID Sensing

A UHF RFID antenna-based crack sensor is founded on the theory that a crack will alter the radiation and impedance pattern of the antenna. The sensing results of cracks are obtained from variations in the received power compared to the power received from the

sensor in the case of a healthy state with no crack. More specifically, the received power at the reader antenna in the presence of a crack is expressed as  $P_R[\Psi]$  [33]:

$$P_R[\Psi] = \left( \frac{4\pi d}{\lambda_0} \right)^2 \frac{P_{th}}{G_R G_T[\Psi] \tau[\Psi] \eta_\rho} \quad (1)$$

where  $d$  is the distance between the reader and tag antennas,  $\lambda_0$  is the free-space wavelength,  $P_{th}$  is the minimum threshold power to activate the tag chip,  $G_R$  is the gain of the reader antenna,  $G_T[\Psi]$  is the gain of the tag antenna and  $\eta_\rho$  is the polarization mismatch between the reader and tag antennas. The power transmission coefficient,  $\tau[\Psi]$ , is the impedance mismatch between the tag chip impedance ( $Z_c = R_c + jX_c$ ) and the tag antenna impedance ( $Z_a[\psi] = R_a[\psi] + jX_a[\psi]$ ), and is given by [33]:

$$\tau[\psi] = 1 - \left| \frac{Z_c - Z_a^*[\psi]}{Z_c + Z_a^*[\psi]} \right|^2 = \frac{4R_c R_a[\psi]}{|Z_c + Z_a[\psi]|^2} \quad (2)$$

where  $*$  indicates the conjugate value,  $[\psi]$  indicates the presence of a crack. As shown in (1), in the presence of a crack, the received power is changed due to the variations of  $G_T[\Psi]$  and  $\tau[\Psi]$  caused by the variation of  $Z_a[\psi]$  in (2) [33]. Considering practical implementations where the output metric is *RSSI*, the relationship between gain, transmission coefficient, and *RSSI* is directly verified [17].

$$RSSI_{r,g}^i(\delta, d, \varepsilon_{eff}) = \frac{\lambda^4}{4\pi d_i^4} \cdot \eta_p \cdot G_r^2 \cdot G_g^2(\delta, \varepsilon_{eff}) \cdot \tau_g(\delta, \varepsilon_{eff}) \quad (3)$$

Hence, we can infer that the theoretically derived parameters are equivalent to *RSSI*. Additionally, machine learning detection accuracy using *RSSI* would produce similar comparable results as to those parameters used in this paper.

### 3.2. Artificial Neural Network Overview

The ANN is a machine learning method that builds on the notion of imitating the biological decision process of the human brain. In contrast to the traditional regression method, the ANN can model complex nonlinear correlations. The ANN is also extremely resistant to errors and faster as it can be implemented in parallel processing. Similar to the neurons in a biological system, the ANN is composed of interconnected nodes and layers as follows.

#### (1) Topology

Figure 2 shows a general architecture of an ANN. The nodes are linearly arranged in phases called layers. Generally, there are three sets of layers, called input layers, output layers, and hidden layers. The nodes can be interconnected in two methods. The first method is a connection with one direction and one has a two-direction connection, where the output can be an input. So, there are two types of neural networks called feedforward and feedback network.

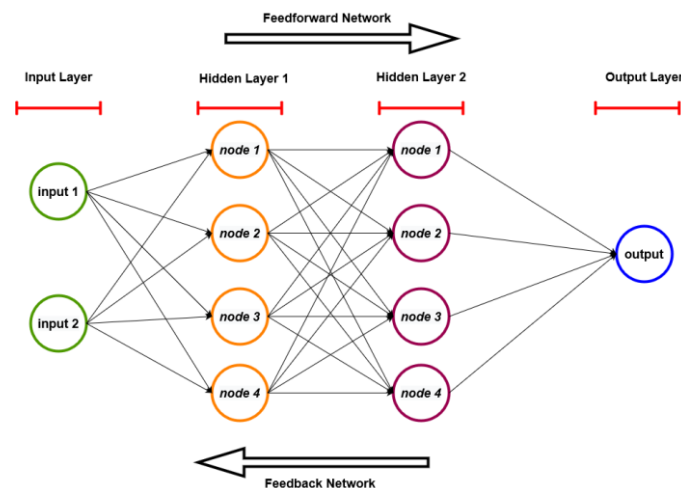
#### (2) Mechanism

Neural network structure is based on three modules: properties of the nodes, network model, and learning methods. Each node relates to other nodes through connections with particular strength factors called weights. It denotes the steepness of a function. On the other hand, bias is similar to the intercept added in a linear equation. Hence, bias is a constant, which helps the model to fit best for the given data. Each node has a threshold value that needs to be achieved by the weighted input sums for it to be activated and allow the signal passage. Once a node is activated, the signal is transformed through a transfer

function  $f$  and permitted to proceed to the next nodes. Mathematically this model can be represented by [37]

$$y = f \sum_{i=0}^n w_i x_i - T \quad (4)$$

where  $y$  is the output of the node,  $w_i$  is the weight of the input  $x_i$ , and  $T$  is the threshold value. There are many types of transfer functions such as Sigmoid and step. Owing to the nonlinear nature of most problems in nature, nonlinear transfer function is more useful than linear ones [38].



**Figure 2.** ANN Topology.

### (3) Learning

During network training, the ANN adopts a learning process to modify or tune the weights to the optimal values. Within the learning process, there are two categories: supervised learning and unsupervised learning. In supervised learning, a sample of inputs and corresponding outputs are provided as a training set. The training set needs to be created such that it exhibits the model's features well, to prevent unreliability in the generated model. When the generated model produces the expected outcomes for a wide range of inputs, the supervised model is finalized by fixing the necessary weights.

Unsupervised learning, on the other hand, studies only the input data to find whether it contains a characteristic pattern within itself or not. It does not examine the target output values. Regression is an example of supervised learning whereas k-means clustering is how an unsupervised method works [37].

### (4) Training Algorithms

The main difficulty in applying ANN-based crack detection is discovering the most suitable training function for the classification task. Another objective of the study is to discover the most suitable training function for the classification task. This is achieved by utilizing 4 types of training algorithms, which are:

1. Gradient descent-based resilience back-propagation algorithm
2. Scaled Conjugate back-propagation
3. Bayesian Regularization (2 and 3 are both under conjugate-based algorithms)
4. Levenberg-Marquardt algorithm is selected from Quasi Newton-based algorithm branch.

The algorithm outcomes are compared for each of the 4 intended detection tasks:

1. Between Healthy and Cracked
2. Between three types of crack orientations
3. Detecting a crack width of 0.5 mm



#### 4. Detecting the location of the crack

In this paper, the performances of four types of training algorithms in feedforward neural networks are compared for each of the four forms of classification. The proposed work compares training algorithms based on their convergence rate and the correctness of the classification.

##### a. Gradient Descent Algorithm

Gradient descent is an optimization procedure for finding a local minimum of a differentiable function. Gradient descent has the goal to diminish the cost function by finding the values of a function's parameters (coefficients). It is one of the most popular training algorithms, which updates weights and biases in the direction of the negative gradient of the performance function [39].

##### b. Resilience Backpropagation

This is a gradient descent-based technique that eliminates the effects of the magnitudes of the partial derivatives [40]. Here the sign of the derivative is used to determine the direction of the weight update and the magnitude of the derivative has no effect on the weight update. The size of the weight change is determined by a separate update value of the derivative of the performance function and the number of iterations [41].

##### c. Conjugate Gradient Algorithm

In the case of the basic gradient descent algorithm, the weights are adjusted towards the negative of the gradient, where the performance function decreases fastest. However, this method does not always give rise to the quickest convergence. In the conjugate gradient algorithm, a search is conducted along conjugate directions, instead of the diminishing direction. It produces faster convergence. The conjugate gradient algorithms can work better for bigger batches of weights with their requirement for low storage spaces [42].

##### d. Scaled Conjugate Gradient

It uses a size scaling mechanism and avoids a time-consuming line search per learning iteration. This mechanism produces a faster convergence by reducing the computation in each iteration, compared to many other algorithms [43].

##### e. Quasi-Newton Algorithm

Newton's method is capable of producing better and faster optimization than conjugate gradient methods. Newton's method utilizes the Hessian matrix (second derivative) of the performance index at the current values of the weights and biases. Newton's method has a faster convergence rate but is complex due to the Hessian matrix calculation. Hence Quasi-Newton method was created that skipped the Hessian matrix calculation with an approximation at each iteration stage.

The Levenberg–Marquardt backpropagation algorithm is an iterative Quasi-Newton method that ensures the decrease of performance function at each iteration, making it the fastest training algorithm. It expresses the minimum of a multi-variable function using the sum of squares of non-linear real-valued functions. This function requires larger memory and computation time due to the calculation of both gradients and the approximated Hessian matrix [44].

##### f. Bayesian Regularization

Bayesian regularized artificial neural networks incorporate the Bayes theorem into regularization to reduce cross-validation requirements while being more robust at the same time. Regularization is the step of adding a weight penalty in the form of a diagonal matrix, to control the problem of inflation of the weights that can occur when the model is complex. Bayes' theorem is the process of making inferences regarding past events using probability. Bayesian Regularization based functions can overcome overtraining because there is a measure to stop training if a certain procedure is reached. It is also hard to overfit because it does not train irrelevant parameters or weights [45].

#### 4. Methodology

The purpose of this study is:

1. To propose a novel machine learning-based coal mining conveyor belt crack detection incorporating the UHF RFID tag sensor that can:
  - detect cracks,
  - detect the type of cracks among vertical, horizontal, and diagonal,
  - detect as low as 0.5 mm of crack width,
  - predict with high precision the location of the crack on the belt.
2. To provide a further detailed comparative analysis of crack detection using a wide range of machine learning algorithms, and inputs.

All these experiments are carried out on a Windows 10 (64-bit) operating system with i5 processor and 8 GB RAM and Matrix Laboratory (MATLAB) software. In order to best replicate the actual surrounding environmental effects such as noise and interferences, the sensor design using the Computer Simulation Technology Microwave Suite (CST) software takes into account the electromagnetic boundary conditions and the electromagnetic effects of the belt and sensor materials involved in the simulation. In our proposed approach, we propose that the reader antenna is fitted and fixed underneath the conveyor belt to collect the sensor response from the sensors which are embedded inside the belt. So there is no effect of the minerals on the sensor response and the scenario for the experiment in the paper is similar to that in real-time. Experiments are conducted by placing the designed UHF sensor on top of the belt structure and subjecting it to different orientations of vertical, horizontal, and diagonal cracks, at 10 different locations, for varying crack widths of 0.5 mm to 5 mm. The measured data to be considered are the sensor complex impedance ( $Z_a[\psi]$ ), transmission coefficient ( $\tau[\Psi]$ ), and the antenna gain ( $G_T[\Psi]$ ). For each crack width, there are a set of impedances, gains, and transmission coefficients corresponding to 10 different locations across the belt. Thus, each orientation of crack has data for six crack widths, which at ten different locations produces a total sample size of 360 features for ( $Z_a[\psi]$ ), and 180 features for ( $\tau[\Psi]$ ) and also ( $G_T[\Psi]$ ). The uncracked healthy state of the tag on the belt has one feature for each calculated parameter. A combination of features for each crack is tested as inputs to the ANN training stage. Following the previous studies in the literature such as [21,46], the dataset data is divided into sets for training (70%), validation (15%), and for testing (15%). Here, as will be demonstrated in Section 6, the simulation results show that the proposed model achieves a classification accuracy rate of 100% for detecting cracked states and this illustrates that the training dataset is sufficient to train the machine learning algorithms.

In order to eliminate any bias in the presentation order of the sample patterns to the ANN, these sample sets are randomized and normalized to fall within the range of 1 to  $-1$ . In the hidden layer, the reported experimental data are predicted with 10 neurons and hyperbolic tangent sigmoid transfer function. The number of neurons in the input layer varies between 1 and 2 depending on the input parameter. The number of neurons in the output layer ranges from 1 to 5 depending on the number of classes to identify. Since our problem is a multiclass classification problem, the discrete probability distribution-based transfer activation function is used in the output layer of the neural network instead of the sigmoid or linear step function.

Figure 3 shows the flowchart of the proposed model with  $Z_a[\psi]$  as an input. Figure 4 shows the ANN model forming our system for detecting cracked states with an input of  $Z_a[\psi]$ . Since the paper intends to classify four different cases as mentioned earlier, when the target classes to determine change, the number of neurons in the output layer will also change. Figure 5 shows the general neural network structure created to test the algorithms where  $w$  and  $b$  are the respective weights and biases of the layers. The output of ANN represents the detected class.

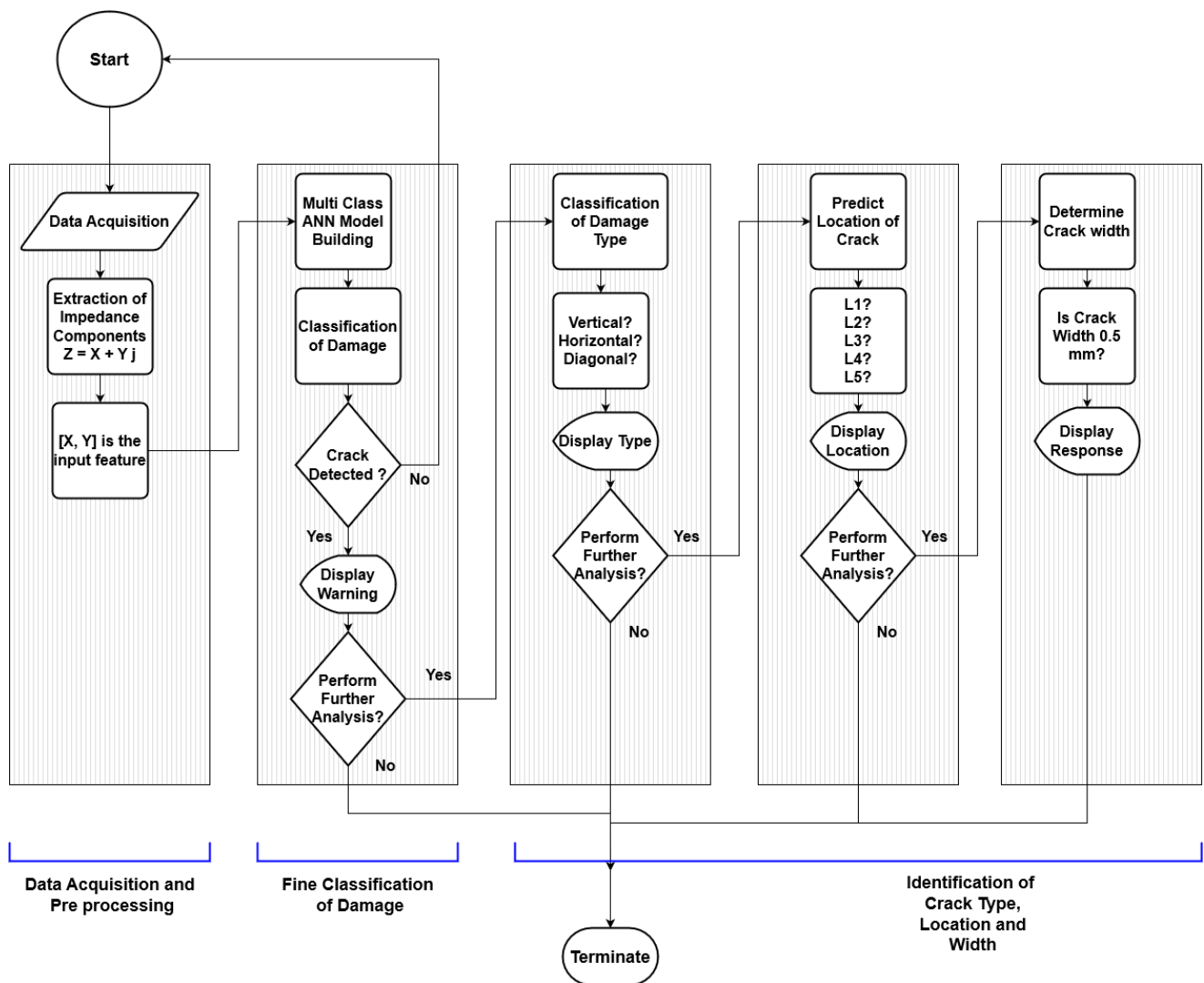


Figure 3. Flowchart of the Machine Learning classification methodology.

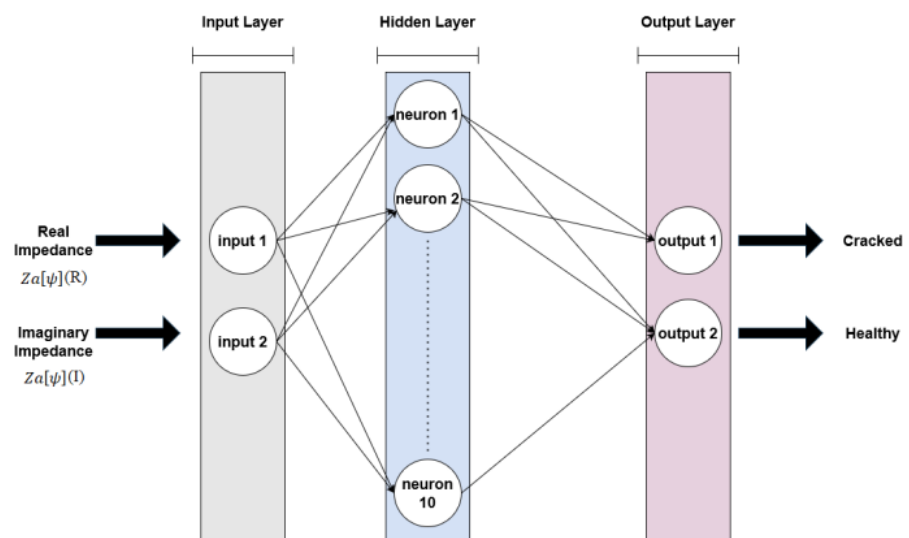
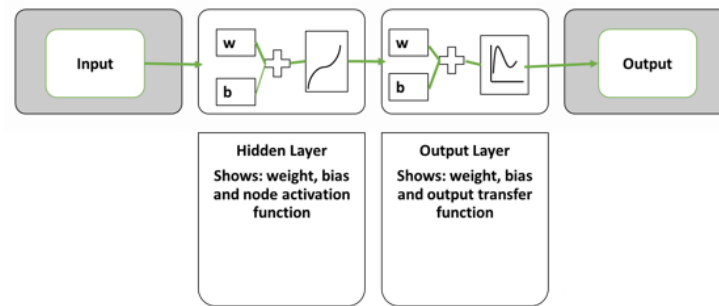


Figure 4. ANN structure used in the paper.



**Figure 5.** ANN structure showing.

The input to the hidden layer of the network is varied and consists of either of the following,

1. real and imaginary impedance values,  $Z_a[\psi]$ ,
2. transmission coefficient,  $\tau[\Psi]$ ,
3. Gain,  $G_T[\Psi]$ ,
4. Transmission coefficient, Gain  $[\tau[\Psi], G_T[\Psi]]$ .

The design goal is to obtain high-accuracy results with low computational complexity that can be implemented without the need for additional operations to improve the accuracy.

## 5. Integration of RFID-Based Sensors into IoT

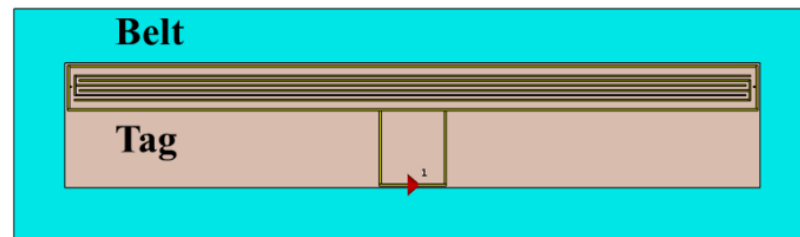
In this section, a scenario where the proposed approach can be integrated into the IoT paradigm of smart sensing has been discussed. UHF RFID sensors are embedded in the conveyor belt and their output parameters are uploaded into the cloud regularly. The ML model is installed into the cloud platform and is fed by the incoming sensor data. The model splits the data into training, validation, and testing sets. The model utilizes the predetermined network algorithm to train and create a hypothesis. The hypothesis is validated and tested to produce the final network model. Once the final model is ready, it is sent to the decision-making endpoint. This trained model continuously performs classification on the incoming live-streamed input data. It produces verdicts on the current state of health of the conveyor belt based on the parameters that were ( $Z_a[\psi]$ ,  $\tau[\Psi]$ , or  $G_T[\Psi]$ ) impacted by the crack, as it was trained to identify.

## 6. Experimentation and Results

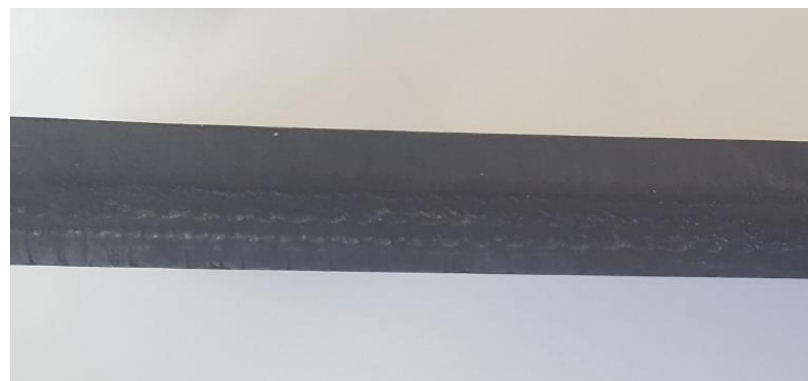
### 6.1. UHF RFID Tag and Belt

Figure 6 shows the designed UHF RFID tag sensor used during experimentation placed on top of the simulated conveyor belt. The sensor is composed of an inductive loop connected to the IDC resonator via a gap-coupled outer box. The resonating fingers are closely distributed to increase the total capacitance, mutual coupling, and eventual sensitivity of the structure. Physical distress such as cracks will be represented through the variation of the electromagnetic properties. The inductive loop is designed to match. The commercial UHF RFID chip has an internal impedance of  $18-j164$  at 915 MHz. The resonator has a response of 898 MHz. The smith chart and reflection coefficient of the sensor in the UHF band are depicted in ([16], Figures 4 and 5). The original sample of the conveyor belt used in the experiment is shown in Figure 7. The total height of the belt is 11 mm. The belt is composed of several different material layers with different dielectric constants. The dielectric properties of the layers will cumulatively affect the transmission of signal into the tag. Thus, the belt's effect on the RFID tag sensor performance needs to be included during our crack detection analysis. The belt is composed of a material called Aramid and Reinforced Fiber layers (Nylon, Polyester, and Cotton) [47,48]. However, since in the practical visioning of a mining plant, the belt affects the RFID system as a block, microwave

characterization using the transmission measurement method was used to determine the dielectric constant of the whole belt for the UHF band [49]. This enabled us to consider the belt as a single block with an individual dielectric constant, instead of mathematically calculating the effective dielectric constant which would require the individual dielectric constant and loss tangent of different layers. Theoretically calculated parameters can sometimes conflict with physical experimental values, introducing uncertainties and errors in the desired outcomes. By considering the practical value of our experiment, we aim to increase our proposed system's practical feasibility.



**Figure 6.** Tag placed on top of the belt.



**Figure 7.** Conveyor belt sample with interleaved layer.

It is worth mentioning that this current work is unique to [16,17] in the sense that it performs Machine Learning based health status detection on the UHF RFID sensor response. Whereas, the previous works [16,17] do not contain the ML application. So, the proposed systems in [16,17] are integrated with ML in this study to provide accurate health status of the conveyor belt.

### 6.2. Characterization of Cracks at Different Locations

The conveyor belt and RFID sensor tag are subjected to three types of cracks of horizontal, vertical, and diagonal orientations as shown in ([16], Figure 12). For each type of crack, further investigation is carried out by varying the crack widths. Figure 8 shows the gradual increase of crack widths introduced during the vertical crack as an example.

### 6.3. Results

Table 3 shows a portion of the set of data that we have computed from our simulation experiments corresponding to the vertical crack scenario. The sensor is subdivided into 10 regions; thus, the crack occurs at ten different locations. The complex input impedance, transmission coefficient, and gain for each position and crack width are calculated. The table lists the measurements for the first two positions.

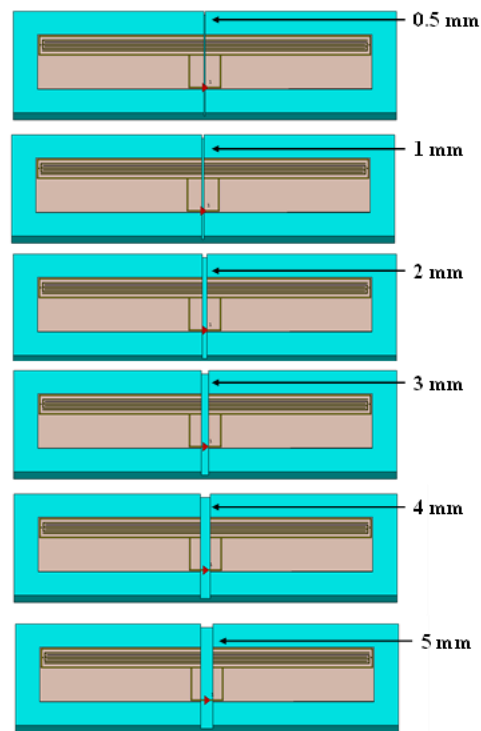


Figure 8. The size of the crack widths experimented.

Table 3. Simulated input parameters for Vertical Crack Orientation (first two positions).

Position Number	1	2	3	4	5	6	7	8	9	10
Location (mm) (w.r.t leftmost corner of sensor considered as 0)	0	14.44	28.88	43.32	57.76	72.2	86.64	101.08	115.52	129.96
Crack width	0.5 mm				1 mm					
position	Transmission coefficient (dB)		Impedance (ohm)		Gain (dB)	Transmission coefficient (dB)		Impedance (ohm)		Gain (dB)
1	−0.27		15.96 + j220.15		2.218	−0.278		16.38 + j220.09		2.197
2	−0.394		33.30 + j263.98		1.579	−0.406		33.90 + j262.40		1.566
Crack width	2 mm				3 mm					
position	Transmission coefficient (dB)		Impedance (ohm)		Gain (dB)	Transmission coefficient (dB)		Impedance (ohm)		Gain (dB)
1	−0.324		18.78 + j217.89		1.973	−0.334		19.38 + j217.84		1.970
2	−0.384		32.22 + j263.34		1.744	−0.369		30.73 + j262.43		1.819
Crack width	4 mm				5 mm					
position	Transmission coefficient (dB)		Impedance (ohm)		Gain (dB)	Transmission coefficient (dB)		Impedance (ohm)		Gain (dB)
1	−0.34		19.7 + j217.73		2.027	−0.346		19.90 + j216.79		2.021
2	−0.365		30.31 + j261.87		1.727	−0.345		28.63 + j262.26		1.848

In order to obtain the optimal model that will produce the best detection accuracies, we varied the inputs fed to the ANN by using different combinations of parameters, for example only impedance, gain or transmission coefficient, or a combination of gain and transmission coefficient.

Based on a range of trials, it was found that for fine classification between cracked and healthy states, all of the inputs worked well. However, for the identification of other details,



proposed ANN models worked best when impedance was used as the input parameter, i.e., impedance manifested the change induced by the crack physiology best.

There are numerous benefits to testing with several algorithms. Experimentations in multiple settings allow for the enhanced study of comprehensive data preprocessing and modelling techniques. In practical terms, it will give the researcher more room to navigate the options at hand and choose the best suitable option depending on the particular project requirements.

### 1. Between Cracked and Healthy

For each input element, at ten locations, there are two recognized classes: Cracked and Uncracked. Table 4 demonstrates the Classification performance comparison between detecting Cracked and Healthy states. The following observations can be deduced from the outputs:

- All of the input parameters are highly efficient in detecting cracked and uncracked states of the belt.
- A maximum accuracy rate of 100% is achieved with all inputs.
- Among the training algorithms, Bayesian Regularization is found to be the most efficient. It produces a 100% identification rate in all instances. However, it also takes the longest time to converge in some cases. Having said that, the execution time is still very low.
- Except the gradient descent technique, the remaining methods produce considerably higher accuracy rates of at least 84.4%, which is more than acceptable.
- Gain, Transmission Coefficient, and Impedance appear to be the optimal input parameter, whereas the combination of the Transmission coefficient and gain fails to produce good rates for all cases.

**Table 4.** Classification Performance between cracked and healthy states.

Input Parameter	Algorithm	Epoch (Iterations to Converge)	Time (s)	Classification Accuracy Rate (%)
Impedance	Scaled Conjugate	47	1	96.7
	Bayesian Regularization	104	4	100
	Quasi Newton	56	1	94.6
	Gradient Descent	49	0.3	75
Transmission Coefficient	Scaled Conjugate	15	0.4	95.8
	Bayesian Regularization	89	2	100
	Quasi Newton	10	0.2	98.3
	Gradient Descent	80	3	96.3
Gain	Scaled Conjugate	19	0.3	100
	Bayesian Regularization	10	0.9	100
	Quasi Newton	8	0.23	100
	Gradient Descent	8	0.33	100
Transmission Coefficient, Gain	Scaled Conjugate	24	0.15	100
	Bayesian Regularization	12	6	100
	Quasi Newton	56	1 s	84.4
	Gradient Descent	47	0.3	61.1

Overall, the proposed network is very robust in detecting underlying cracks and healthy patterns from the inputs provided the appropriate training algorithm is chosen.

Owing to the quality of the input parameters, the designer can expect satisfactory results from a pool of several training functions.

## 2. Between Three Different Types of Cracks

For each input element, at ten locations there are three recognized classes: Vertically Cracked, Horizontally Cracked, and Diagonally Cracked. Table 5 depicts the classification performance comparison between detecting Vertical, Horizontal, and Diagonally Cracked states. The following observations can be deduced from the outputs:

- Each of the three different crack orientations has distinct individual patterns within themselves, which helps in their identification.
- Impedance as an input and Bayesian Regularization as the training algorithm produces the highest accuracy rates of 83.9%. Although not as high as cracked vs healthy detection rates, it is still a good acceptable rate.
- The Transmission Coefficient Gain data set also has the best result with Bayesian Regularization.
- This comparative analysis identifies Impedance and Bayesian Regularization to be the best combination of input and training functions that will provide the best detection between Vertical, Horizontal, and Diagonal Cracks.

**Table 5.** Classification Performance between Vertical, Horizontal, and Diagonally Cracked states.

Input Parameter	Algorithm	Epoch (Iterations to Converge)	Time (s)	Classification Accuracy Rate (%)
Impedance	Scaled Conjugate	19	0.1	60
	Bayesian Regularization	1000	40	83.9
	Quasi Newton	38	4	77.8
	Gradient Descent	74	0.3	58.3
Transmission Coefficient, Gain	Scaled Conjugate	9	0.12	48.9
	Bayesian Regularization	1000	18	70.6
	Quasi Newton	18	0.4	68.3
	Gradient Descent	17	0.5	53.3

Examination of the results shows that the classification of different crack orientations is a challenging task. In the future, suitable feature processing methods can be studied to obtain better results.

## 3. Detecting Crack Width of 0.5 mm

For each input element, at ten locations there are two recognized classes, 0.5 mm of crack width and others. We have used 0.5 mm to test our proposed system as it is the lowest crack width that can be obtained and is also of very little width. We believe the early stage crack starts with a very little typical width such as 0.5 mm [50]. Table 6 shows the classification performance comparison between detecting crack widths. The following observations can be deduced from the outputs:

1. The designed model is quite successful in detecting crack widths as low as 0.5 mm.
2. Both input parameters consistently produce good acceptable accuracy rates of at least 83.3% with all of the training algorithms.
3. Unlike the other cases, the [Transmission Coefficient, Gain] data set is found to be the better input parameter than Impedance, producing a slightly higher identification rate of 84.4%.
4. Because the results are consistent throughout, the decision maker has the freedom to choose from a variety of parametric combinations.

**Table 6.** Classification Performance between crack widths.

Input Parameter	Algorithm	Epoch (Iterations to Converge)	Time (s)	Classification Accuracy Rate (%)
Impedance	Scaled Conjugate	14	0.1	83.3
	Bayesian Regularization	1000	40	83.3
	Quasi Newton	19	5	83.3
	Gradient Descent	13	0.2	83.3
Transmission Coefficient, Gain	Scaled Conjugate	15	1	83.3
	Bayesian Regularization	262	6	84.4
	Quasi Newton	7	0.11	83.3
	Gradient Descent	16	1	83.3

The detection of narrower crack widths is tricky and requires fine scrutinization. Although the identification rates here are considered acceptable, further research should be undertaken to improve the outcome.

#### 4. Detecting Crack Location

For each input of two elements, at ten locations there are five recognized classes, region 1 to region 5. Table 7 shows the classification performance comparison between detecting crack locations. The following observations can be deduced from the outputs:

- Bayesian Regularization and Impedance again outperform the others, producing as high as 95.5% of identification rates. It also produces good results with the Quasi-Newton method.
- Impedance is the better input parameter here, producing higher accuracies than the [Transmission Coefficient, gain] data set. Even the lowest outcome of 61.1% is higher than the maximum rate of 49.4% achieved by the other input.
- [Transmission Coefficient, Gain] appears to struggle in the location of crack detection, failing to reach 50% of accuracy.

**Table 7.** Classification Performance between crack locations.

Input Parameter	Algorithm	Epoch (Iterations to Converge)	Time (s)	Classification Accuracy Rate (%)
Impedance	Scaled Conjugate	79	0.22	63.9
	Bayesian Regularization	1000	20	95.5
	Quasi Newton	56	1	84.4
	Gradient Descent	47	0.3	61.1
Transmission Coefficient, Gain	Scaled Conjugate	21	0.1	38.9
	Bayesian Regularization	100	17	49.4
	Quasi Newton	10	0.18	42.8
	Gradient Descent	11	0.2	40.6

This analysis depicts the benefits of performance comparison using parameter variation. It assists in the determination of the optimal parameters capable of producing consistent good results. For example, it shows that although a combination of Transmission Coefficient and Gain as data produces good results in some classification scenarios, it cannot be chosen as the optimal parameter for our proposed model because of its inconsistencies.

Table 8 shows the summary of the best classification performance parameters for all the discussed cases.

**Table 8.** Summary of the best classification performance parameters for all of the cases.

Input Parameter	Classification			
	<i>Cracked vs. Uncracked</i>	<i>Vertical vs. Diagonal vs. Horizontal</i>	<i>Detecting 0.5 mm Crack Width</i>	<i>Detecting Location</i>
Impedance	Bayesian Regularization—100%	Bayesian Regularization—83.9%	ALL—83.3%	Bayesian Regularization—95.5%
Transmission Coefficient	Bayesian Regularization—100%	Bayesian Regularization—63.3%	Bayesian Regularization—83.3%	Bayesian Regularization—71.1%
Gain	ALL—100%	Quasi Newton—66.1%	ALL—83.3%	Bayesian Regularization—57.8%
Transmission Coefficient, Gain	Bayesian Regularization, Scaled Conjugate—100%	Bayesian Regularization—70.6%	Bayesian Regularization—84.4%	Bayesian Regularization—49.4%

#### 6.4. Analysis

The purpose of this study is to propose a novel machine learning-based coal mining conveyor belt crack detection method incorporating the UHF RFID tag sensor that can identify between cracked and healthy states, detect the type of crack orientation from vertical, horizontal, and diagonal, detect as low as 0.5 mm of crack width, and predict the crack location with high-accuracy. The study also conducted an elaborate comparative analysis of crack detection using a wide range of machine-learning algorithms and inputs.

In order to make the comparison valid, all experiments were conducted in identical simulated environments. The results can be analyzed from different perspectives: the quality of the data set, the parameters of the neural network, and finally the effect of data augmentation. We are not able to perform experiments based on physical experimental data sets because then we would require immense quantities of fabricated tags and access to the actual conveyor belt of the mining plant. Thus, we have resorted to the simulation results to verify our proposed model.

It can be observed that the type of input parameter had a large impact on the model's accuracy. Different data sets are tested, and, finally, it is concluded that during the building and testing of the model to classify between cracked and uncracked states, all the Gain, Transmission coefficient and Impedance values embody the distinction between cracked and healthy states the best. Accuracy as high as 100% is achieved in most cases with all three inputs. This occurs due to the effect of cracks on the tag operation; the crack's presence is revealed through the changes in these parameters. It is noticed that an evenly distributed data set with sufficient non-cracked data provided much better results, preventing any unwanted skewness.

However, for all the other cases: discriminating between the types of cracks from Vertical, Diagonal, and Horizontal, determination of the crack width, and the prediction of crack location, input impedance appeared to be the more robust feature with a better detection rate across all training algorithms, except for classification of 0.5 mm crack width where a combination of transmission coefficient and gain produced a slightly higher accuracy rate of 84.4%.

It is observed that although the use of impedance, gain and transmission coefficient alone resulted in 100% of accuracy rate, the neural network learning algorithm also caused the detection rates to change. Gain is found to be the only input parameter that produced 100% of the identification rate for all of the training functions applied.

Regarding the number of layers, neurons, and other hyperparameters, models with Bayesian Regularization as the learning function is observed to be more responsive in learning the underlying pattern of the input data set. We also found out that ten neurons and a two-layered network suited our problem the best. Increasing the number of neurons will make the model more adaptive to smaller details, and increasing the number of layers will allow the model to work with more complex features. However, this does not guarantee

better classification. In fact, increasing more neurons and layers than is necessary will cause the network to overfit to the training set; that is, it will learn the training data, but it will not be able to generalize to new unseen data. It increases the risk of lower performance. Moreover, higher performance is achieved by utilizing a large number of epochs and regularization implemented in the Bayesian Regularization algorithm. The Bayesian Regularization method is more efficient because it eliminates the dependency on irrelevant and highly correlated inputs and determines the importance of necessary weights.

As can be seen from the tables, for most of the algorithms time elapsed by the end of execution is only a few seconds. Gradient descent is usually the fastest to converge, while it maintains higher learning rates. Scaled Conjugate and Quasi-Newton are almost as fast as Gradient Descent. Bayesian Regularization requires the longest time among all, although, it outperforms others in performance. Due to the fact that it requires the calculation of a very computationally demanding Hessian matrix [51], and would thus account for a long time. Gradient descent and the Scaled Conjugate method converge in a reduced number of iterations versus all the other training functions. It always takes a higher number of epochs (iterations) for Bayesian Regularization to achieve the performance target, while all the other functions require a far lesser number.

Table 8 shows the best classification percentage for each classification network with respect to the input. We can say that depending on the situational requirement, Bayesian Regularization is a very good fit for all the discussed classification scenarios; however, if the execution time is to be considered, there are other functions with acceptable detection rates that converge in lesser times.

It is hard to cover all types of cracks, with different shapes and angles. Data augmentation, by using a combination of different input parameters, in addition to increasing the number of samples, can increase the variance in the data set by arbitrarily generating some of the absent crack shapes and angles. When both the transmission coefficient and gain are used as inputs, accuracy is boosted in some cases. Therefore, in the future, data augmentation can also be considered an important factor in improving classification accuracy. The objective of the proposed method is to be able to detect tears as narrow as 0.5 mm at the earliest opportunity without the need for physical investigation. This method provides the investigator flexibility to assess the location and determine the suitable mode of maintenance without shutting down the plant operation. Stopping the operation is not a desirable choice because it can lead to a capital loss of AUD 1800 per minute [52]. However, our proposed method uses RFID sensors in conjunction with machine learning identification which is ideal in the case of real-time monitoring of conveyor belts with lengths of many kilometers. The monitoring system is implemented by overlaying or embedding RFID sensors into the belt. Each sensor monitors cracks in the area of the belt on top of it. Furthermore, to accurately monitor the whole belt, as future work, the algorithm can be tuned to detect cracks occurring at places not covered by sensors. Machine Learning can help to improve detection during those complicated scenarios by virtue of its superior predictive ability.

As demonstrated in [16,17], the RFID sensor can be embedded inside the conveyor belt to provide protection from damaging or peeling off due to the frictional force of the idlers or rollers during the operation of the conveyor belt. Following the approach of [16,17] in this paper, the proposed system based on ML can obtain the measurements from the sensors while they are embedded inside the conveyor belt. We expect that the RSSI measurements could be impacted by the materials around the embedded sensors. However, as shown in [16,17], the RSSI measurements are varied with crack widths and orientations even though the sensors are embedded inside the belt. For our experiments in the paper, there is no impact on the performance of the proposed system based on ML for either when the sensors are embedded as followed in [16,17] or when the sensors are attached to the surface of the conveyor belt as followed in this paper.

Comparing the proposed method to classic IPTs and other machine learning models also shows the advantage of implementing ANN learning methodology for detecting cracks.

Table 9 demonstrates the performance of the proposed system in comparison to other studies. Image processing techniques are capable of generating the exact shape of the cracks; however, most of the time, the examined surfaces will have more complex characteristics similar to cracks, such as dents, bends, ridges etc., and such damage might have the same color and shape as that of a crack and IPT becomes inefficient. Machine learning methods such as SVM respond well to proper feature selection and feature extraction/transformation techniques, which can be time-consuming to generate and mostly require one-against-all types of classifiers [53]. ANN-based methods can do multi-class identification, use back-propagation to adjust the error weights, can modify the layers, neurons, and transfer functions to intelligently learn the input data pattern. The pre-trained network can also be reused instead of building another model from scratch. Therefore, it can be deduced that deep learning methods using ANN yield better crack detection results.

**Table 9.** Performance Comparison.

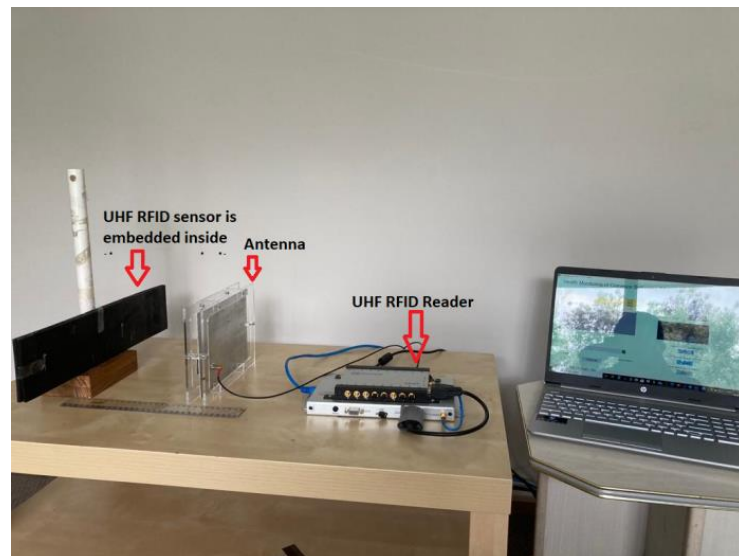
Method	Study	Results	Features	Year
Impedance and Machine Learning	Detection of Conveyor Belt Cracks, orientation, crack widths, and crack location	- Accuracy rate 100%	- Works with coal mining conveyor belts composed of fibre and polyester.	2021
Image Processing Technique	Detection of Conveyor Belt longitudinal tears	- Accuracy rate 92.1%	- Works with steel cord conveyor belt only.	[30] 2019
Image Processing Technique Machine Learning	Asphalt pavement crack detection	- Precision 90.13%	- Has difficulty to detecting hair line cracks. - Confuses pavement edges with cracks.	[22] 2017

## 7. Future Works

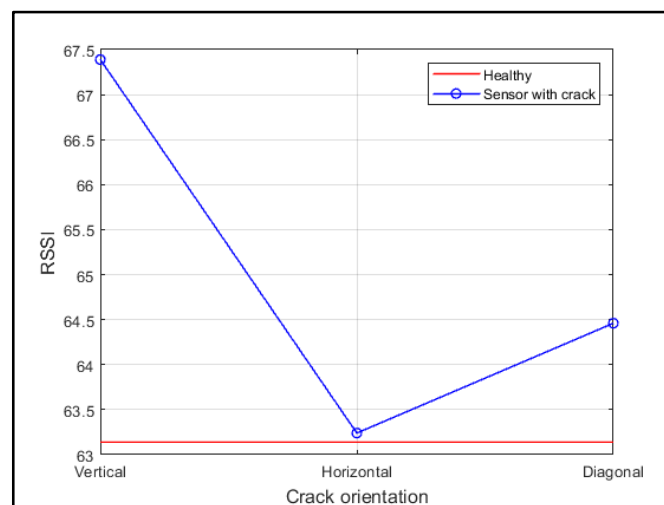
An experiment is conducted for the fabricated UHF RFID sensor at different crack orientations, as shown in Figure 9. Figure 9 shows the experimental setup in Monash Microwave, Antenna, RFID, and Sensor (MMARS) Laboratory within the Department of Electrical and Computer System Engineering, Monash University. Here the CSL468 Class 1 Gen 2 UHF RFID reader is used to measure the RSSI from the sensor. A linear polarised antenna is used with a gain of 7.88 dBi and the total cable loss is 5.7 dBm at 923 MHz [54]. As shown in Figure 10, the response of the sensor in terms of the received signal strength indicator (RSSI) is measured when the sensor has a single vertical, horizontal, or diagonal crack compared to the healthy response of the sensor without a crack. The results in Figure 10 show that the RSSI values show variation with the change of crack orientation. Distinct differences between the three types of readings representing the types of cracks are observed. Thus, the results in Figure 10 show promising potential for further research investigation by further development of the proposed ML algorithms demonstrated in this paper that adopt the RSSI values from the sensors. However, such development of ML algorithms based on RSSI values requires obtaining the RSSI values from hundreds of sensors and can be considered as another research problem to address in the future together with the effects of motion. It is worth mentioning that we have successfully demonstrated the efficiency of the proposed ML algorithms to achieve high accuracy using MATLAB software, so the coding concept of the proposed algorithms can easily be developed more in the future to apply software such as Python and TensorFlow which are better suited for real-time applications [16,55]. Additionally, the issue of debris and its effect on the sensor response can be addressed by fitting a sweeping mechanism at the junction where the belt rolls over to eliminate any debris left. We intend to look into a series of steps to measure the sensor response and its sensitivity to the presence of coal. It is worth mentioning that the proposed system based on ML in this paper can be modified for applications using



multiple sets of tags and judging the distance between the tag signals (distance sensor) to achieve fault diagnosis and can be the subject of further future work.



**Figure 9.** Experimental Setup with fabricated cracked tags.



**Figure 10.** Cracked Tag response.

It is beyond the scope of this paper to validate experimentally the IoT based implementation of our proposed system. The proposed work in the paper is a logical framework to integrate Health monitoring of mining conveyor belts using UHF RFID and Machine Learning to the IoT domain. We are in the same line with the existing publications in the literature that work on the integration of the IoT and RFID-based sensors [50]. Moreover, in order to be able to test the proposed system in real life, we would require a continuous input of real-time data from a mining plant over a significant period. Such implementation of IoT techniques in the real world is hard to achieve as the implementation requires a substantial period, such as a period of two years as addressed in [56]. Furthermore, implanting thousands of sensors in the conveyor belt in the mine site to collect sufficient data for the IoT is also a time-consuming task. Thus, such implementation of IoT techniques can be considered for future work.

## 8. Conclusions

A detailed comparative performance analysis of ML-based UHF RFID crack detection technique for conveyor belts is conducted. This novel concept has been illustrated keeping hypothetical IoT implementation in mind. This is intended to be a pioneering study that will help to usher the way towards further comprehensive exploration in industrial applications, specifically targeting the coal mines. The depicted research analyzed several simulated cases to verify the designed system. A comprehensive investigation with simulation makes the system robust. Here, an UHF RFID tag sensor is placed on a conveyor belt sample and subjected to different crack orientations and crack widths at several locations across the belt. This provided a huge amount of crack scenarios that represent a real-world situation. An ANN is designed to identify the health state of the belt, the orientation of the crack on the belt, the crack width, and the location of the crack. The model is also tested by varying the training algorithms and input attributes, to find the best combination that will produce the maximum classification accuracy. A crack detection rate of as high as 100% is obtained, with good acceptable rates of more than 84.4% for the other cases. Such an elaborate crack detection system based on ML will immensely help to realize remote SHM of conveyor belts, making it a viable addition to a prospective IoT paradigm. The authors are currently carrying out further research to observe this proposed model's response with chipless RFID tag sensors.

**Author Contributions:** Conceptualization, F.T.Z.; formal analysis, F.T.Z.; methodology, F.T.Z.; project administration, N.C.K.; supervision, N.C.K. and S.D.; validation, F.T.Z., O.S. and S.D.; writing—original draft, F.T.Z.; writing—review and editing, O.S., H.M., N.C.K. and S.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research is supported by Monash University, Australian Coal Association Research Program (ACARP) Grant C28036, and the Australian Government Research Training Program (RTP) Scholarship.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Deivasigamani, A.; Daliri, A.; Wang, C.-H.; John, S. A Review of Passive Wireless Sensors for Structural Health Monitoring. *Mod. Appl. Sci.* **2013**, *7*, 57. [CrossRef]
- Khani, M.M.; Vahidnia, S.; Ghasemzadeh, L.; Ozturk, Y.E.; Yuvalaklioglu, M.; Akin, S.; Ure, N.K. Deep-learning-based crack detection with applications for the structural health monitoring of gas turbines. *Struct. Health Monit.* **2019**, *19*, 1440–1452. [CrossRef]
- Zhang, J.; Tian, G.Y.; Zhao, A.B. Passive RFID sensor systems for crack detection & characterization. *NDT E Int.* **2017**, *86*, 89–99. [CrossRef]
- Zhang, A.; Sun, Y.; Yin, Z. Research status and tendency of longitude tearing protection for belt conveyor. *Coal Sci. Technol.* **2007**, *12*, 77–79.
- Zeng, Q.-L.; Wang, J.-G.; Wang, L.; Wang, C.-L. The Research of Coal Mine Conveyor Belt Tearing Based on Digital Image Processing. In *Advances in Intelligent Systems and Computing, Proceedings of the 2012 International Conference on Communication, Electronics and Automation Engineering, Xi'an, China, 23–25 August 2012*; Yang, G., Ed.; Springer: Berlin/Heidelberg, Germany, 2013; Volume 181, pp. 187–191. [CrossRef]
- Ghoni, R.; Dollah, M.; Sulaiman, A.; Ibrahim, F.M. Defect Characterization Based on Eddy Current Technique: Technical Review. *Adv. Mech. Eng.* **2014**, *6*. [CrossRef]
- Hayashi, M.; Saito, T.; Nakamura, Y.; Sakai, K.; Kiwa, T.; Tanikura, I.; Tsukada, K. Extraction Method of Crack Signal for Inspection of Complicated Steel Structures Using a Dual-Channel Magnetic Sensor. *Sensors* **2019**, *19*, 3001. [CrossRef]
- Goodyear Rubber Products. *Heavy Weight Conveyor Belt Catalogue*. 2012. Available online: [https://goodyearrubberproducts.com/2018pdfs/EP\\_Conveyor\\_Belt\\_Catalog/pdf/EP\\_Conveyor\\_Belt\\_Catalog.pdf](https://goodyearrubberproducts.com/2018pdfs/EP_Conveyor_Belt_Catalog/pdf/EP_Conveyor_Belt_Catalog.pdf) (accessed on 1 March 2022).
- Zhang, J.; Tian, G.Y.; Marindra, A.M.J.; Sunny, A.I.; Zhao, A.B. A Review of Passive RFID Tag Antenna-Based Sensors and Systems for Structural Health Monitoring Applications. *Sensors* **2017**, *17*, 265. [CrossRef]
- Ashton, K. That 'internet of things' thing. *RFID J.* **2009**, *22*, 97–114.
- Gubbi, J.; Buyya, R.; Marusic, S.; Palaniswami, M. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Gener. Comput. Syst.* **2013**, *29*, 1645–1660.
- Foote, K.D. A Brief History of the Internet of Things. Available online: <https://www.dataversity.net/brief-history-internet-things/> (accessed on 6 November 2022).

13. Madakam, S.; Lake, V.; Lake, V.; Lake, V. Internet of Things (IoT): A literature review. *J. Comput. Commun.* **2015**, *3*, 16.
14. Abdelmalek, O. Design and Prototyping of Robust Architectures for UHF RFID Tags. Ph.D Thesis, Université Grenoble Alpes, Grenoble, France, 2016.
15. Duroc, Y.; Kaddour, D. RFID Potential Impacts and Future Evolution for Green Projects. *Energy Procedia* **2012**, *18*, 91–98. [[CrossRef](#)]
16. Dey, S.; Salim, O.; Masoumi, H.; Karmakar, N.C. A Novel UHF RFID Sensor Based Crack Detection Technique for Coal Mining Conveyor Belt. *IEEE J. Radio Freq. Identif.* **2021**, *6*, 19–30. [[CrossRef](#)]
17. Salim, O.; Dey, S.; Masoumi, H.; Karmakar, N.C. Crack Monitoring System for Soft Rock Mining Conveyor Belt Using UHF RFID Sensors. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–12. [[CrossRef](#)]
18. Bishop, C.M. *Pattern Recognition and Machine Learning*, 1st ed.; Springer: New York, NY, USA, 2006.
19. Kannadaguli, P.; Bhat, V. Microwave Imaging based Automatic Crack Detection System using Machine Learning for Columns. In Proceedings of the 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT), Gwalior, India, 10–12 April 2020; pp. 5–8.
20. Fujita, Y.; Shimada, K.; Ichihara, M.; Hamamoto, Y. A method based on machine learning using hand-crafted features for crack detection from asphalt pavement surface images. In Proceedings of the Thirteenth International Conference on Quality Control by Artificial Vision 2017, Tokyo, Japan, 14–16 May 2017; Volume 10338, p. 103380. [[CrossRef](#)]
21. Mustapha, S.; Braytee, A.; Ye, L. Detection of surface cracking in steel pipes based on vibration data using a multi-class support vector machine classifier. In Proceedings of the SPIE Smart Structures and Materials + Nondestructive Evaluation and Health Monitoring, Portland, OR, USA, 12 April 2017; p. 101682. [[CrossRef](#)]
22. Zhang, A.; Wang, K.C.P.; Li, B.; Yang, E.; Dai, X.; Peng, Y.; Fei, Y.; Liu, Y.; Li, J.Q.; Chen, C. Automated Pixel-Level Pavement Crack Detection on 3D Asphalt Surfaces Using a Deep-Learning Network. *Comput.-Aided Civ. Infrastruct. Eng.* **2017**, *32*, 805–819. [[CrossRef](#)]
23. Caizzzone, S.; DiGiampaolo, E. Wireless Passive RFID Crack Width Sensor for Structural Health Monitoring. *IEEE Sens. J.* **2015**, *15*, 6767–6774. [[CrossRef](#)]
24. Zohra, F.-T.; Salim, O.; Dey, S.; Masoumi, H.; Karmakar, N. A Novel Machine Learning Based Conveyor Belt Health Monitoring Incorporating UHF RFID Backscattered Power. In Proceedings of the 2021 IEEE 5th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Purwokerto, Indonesia, 24–25 November 2021; pp. 230–234. [[CrossRef](#)]
25. Mahdavinjad, M.S.; Rezvan, M.; Barekatin, M.; Adibi, P.; Barnaghi, P.; Sheth, A.P. Machine learning for internet of things data analysis: A survey. *Digit. Commun. Netw.* **2018**, *4*, 161–175. [[CrossRef](#)]
26. McClelland, C. Applying Machine Learning to the Internet of Things. Available online: <https://medium.com/iotforall/applying-machine-learning-to-the-internet-of-things-5bd0216d4cc3> (accessed on 6 November 2022).
27. Google. Google Nest Thermostat. Available online: [https://store.google.com/au/product/nest\\_learning\\_thermostat\\_3rd\\_gen](https://store.google.com/au/product/nest_learning_thermostat_3rd_gen) (accessed on 6 November 2022).
28. Wuest, T.; Weimer, D.; Irgens, C.; Thoben, K.-D. Machine learning in manufacturing: Advantages, challenges, and applications. *Prod. Manuf. Res.* **2016**, *4*, 23–45. [[CrossRef](#)]
29. Kalansuriya, P.; Bhattacharyya, R.; Sarma, S. RFID Tag Antenna-Based Sensing for Pervasive Surface Crack Detection. *IEEE Sens. J.* **2013**, *13*, 1564–1570. [[CrossRef](#)]
30. Hou, C.; Qiao, T.; Zhang, H.; Pang, Y.; Xiong, X. Multispectral visual detection method for conveyor belt longitudinal tear. *Measurement* **2019**, *143*, 246–257. [[CrossRef](#)]
31. Hao, X.-L.; Liang, H. A multi-class support vector machine real-time detection system for surface damage of conveyor belts based on visual saliency. *Measurement* **2019**, *146*, 125–132. [[CrossRef](#)]
32. Zhang, M.; Shi, H.; Zhang, Y.; Yu, Y.; Zhou, M. Deep learning-based damage detection of mining conveyor belt. *Measurement* **2021**, *175*, 109130K. [[CrossRef](#)]
33. Zhang, J.; Huang, B.; Zhang, G.; Tian, G.Y. Wireless Passive Ultra High Frequency RFID Antenna Sensor for Surface Crack Monitoring and Quantitative Analysis. *Sensors* **2018**, *18*, 2130. [[CrossRef](#)]
34. Zhang, J.; Sunny, A.I.; Zhang, G.; Tian, G. Feature Extraction for Robust Crack Monitoring Using Passive Wireless RFID Antenna Sensors. *IEEE Sens. J.* **2018**, *18*, 6273–6280. [[CrossRef](#)]
35. Cazeca, M.J.; Mead, J.; Chen, J.; Nagarajan, R. Passive wireless displacement sensor based on RFID technology. *Sens. Actuators A Phys.* **2013**, *190*, 197–202. [[CrossRef](#)]
36. Bruciati, B.; Jang, S.; Fils, P. RFID-Based Crack Detection of Ultra High-Performance Concrete Retrofitted Beams. *Sensors* **2019**, *19*, 1573. [[CrossRef](#)]
37. Zou, J.; Han, Y.; So, S.-S. Overview of artificial neural networks. In *Artificial Neural Networks*; Livingstone, D.J., Ed.; Humana Press: Totowa, NJ, USA, 2008; pp. 14–22.
38. Livingstone, D.J. *Artificial Neural Networks: Methods and Applications*; Humana Press: Totowa, NJ, USA, 2008.
39. Donges, N. Gradient Descent: An Introduction to One of Machine Learning’s Most Popular Algorithms. Available online: <https://builtin.com/data-science/gradient-descent>. (accessed on 6 November 2022).
40. Anastasiadis, A.D.; Magoulas, G.D.; Vrahatis, M.N. New globally convergent training scheme based on the resilient propagation algorithm. *Neurocomputing* **2005**, *64*, 253–270. [[CrossRef](#)]

41. Sharma, B.; Venugopalan, K. Comparison of Neural Network Training Functions for Hematoma Classification in Brain CT Images. *IOSR J. Comput. Eng.* **2014**, *16*, 31–35. [[CrossRef](#)]
42. Hager, W.W.; Zhang, H. A survey of nonlinear conjugate gradient methods. *Pac. J. Optim.* **2006**, *2*, 35–58.
43. Møller, M.F. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Netw.* **1993**, *6*, 525–533. [[CrossRef](#)]
44. Pham, D.; Sagioglu, S. Training multilayered perceptrons for pattern recognition: A comparative study of four training algorithms. *Int. J. Mach. Tools Manuf.* **2001**, *41*, 419–430. [[CrossRef](#)]
45. Burden, F.; Winkler, D. Bayesian regularization of neural networks. In *Artificial Neural Networks*; Livingstone, D.J., Ed.; Humana Press: Totowa, NJ, USA, 2008; pp. 23–42.
46. Chen, F.-C.; Jahanshahi, M.R. NB-CNN: Deep Learning-Based Crack Detection Using Convolutional Neural Network and Naïve Bayes Data Fusion. *IEEE Trans. Ind. Electron.* **2018**, *65*, 4392–4400. [[CrossRef](#)]
47. Shah, K. Construction and Maintenance of Belt Conveyors for Coal and Bulk Material Handling Plants. April 2018, pp. 1–269. Available online: [www.practicalmaintenance.net](http://www.practicalmaintenance.net) (accessed on 4 July 2021).
48. Dunlop, F. Conveyor Belt Manual, Dunlop. Available online: [https://www.fennerdunlopamericas.com/sites/default/files/u562/conveyor\\_belt\\_manual.pdf](https://www.fennerdunlopamericas.com/sites/default/files/u562/conveyor_belt_manual.pdf). (accessed on 6 November 2022).
49. Karmakar, N.C.; Amin, E.; Saha, J.K. Characterization of Smart Materials. In *Chipless RFID Sensors*; John Wiley & Sons: Hoboken, NJ, USA, 2016; pp. 99–124. [[CrossRef](#)]
50. Dey, S.; Bhattacharyya, R.; Sarma, S.E.; Karmakar, N.C. A Novel “Smart Skin” Sensor for Chipless RFID Based Structural Health Monitoring Applications. *IEEE Internet Things J.* **2020**, *8*, 3955–3971. [[CrossRef](#)]
51. Sariev, E.; Germano, G. Bayesian regularized artificial neural networks for the estimation of the probability of default. *Quant. Finance* **2019**, *20*, 311–328. [[CrossRef](#)]
52. Owen, P. Condition monitoring for conveyors. In Proceedings of the 9th International Materials Handling Conference, Fellbach, Germany, 7–9 August 1997.
53. Hsu, C.-W.; Lin, C.-J. A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Netw.* **2002**, *13*, 415–425.
54. C.S. Limited (Version 2.0.1) CSL CS468 16-Ports EPC Class 1 Gen 2 RFID Reader User’s Manual. Available online: <https://www.convergence.com.hk/downloads/cs468/> (accessed on 6 November 2022).
55. Loo, N.; Chiew, Y.; Tan, C.; Arunachalam, G.; Ralib, A.; Nor, M.B.M. A Machine Learning Model for Real-Time Asynchronous Breathing Monitoring. *IFAC-PapersOnLine* **2018**, *51*, 378–383. [[CrossRef](#)]
56. Wang, J.; Fu, Y.; Yang, X. An integrated system for building structural health monitoring and early warning based on an Internet of things approach. *Int. J. Distrib. Sens. Netw.* **2017**, *13*, 12. [[CrossRef](#)]