

Yunus Emre KARABACAK
Nurhan GÜRSEL ÖZMEN
Levent GÜMÜŞSEL

WORM GEAR CONDITION MONITORING AND FAULT DETECTION FROM THERMAL IMAGES VIA DEEP LEARNING METHOD

MONITOROWANIE STANU I WYKRYWANIE BŁĘDÓW PRZEKŁADNI ŚLIMAKOWEJ NA PODSTAWIE TERMOGRAMÓW Z WYKORZYSTANIEM METODY GŁĘBOKIEGO UCZENIA

Worm gearboxes (WG) are often preferred, because of their high torque, quickly reducing speed capacity and good meshing effectiveness, in many industrial applications. However, WGs may face with some serious problems like high temperature at the speed reducer, gear wearing, pitting, scoring, fractures and damages. In order to prevent any damage, loss of time and money, it is an important issue to detect and classify the faults of WGs and develop the maintenance plans accordingly. The present study addresses the application of the deep learning method, convolutional neural network (CNN), in the field of thermal imaging that were gathered from a test rig operating on different loads and speeds. Deep learning approaches, have proven their powerful capability to exploit faulty information from big data and make intelligently diagnostic decisions. Studies concerning the condition monitoring of WGs in the literature are limited. This is the first study on WGs with infrared thermography rather than vibration and sound measurements which have some deficiencies about hardware requirements, restricted measurement abilities and noisy signals. For comparison, CNN was also trained, with vibration and sound data which were collected from the healthy and faulty WGs. The results of fault diagnosis show that thermal image based CNN model on WG has achieved 100% success rate whereas the vibration performance was 83.3 % and sound performance was 81.7%. As a result, thermal image based CNN model showed a better diagnosing performance than the others for WGs. Moreover, condition monitoring of WGs, can be performed correctly with less measurement costs via thermal imaging methods.

Keywords: fault diagnosis, worm gears, thermal imaging, convolutional neural networks, GoogLeNet, condition monitoring.

W wielu zastosowaniach przemysłowych preferuje się przekładnie ślimakowe, ze względu na ich wysoki moment obrotowy, możliwość szybkiej redukcji prędkości i dobrą sprawność zazębnienia. Jednakże przekładnie tego typu narażone są często na poważne problemy, takie jak wysoka temperatura przy reduktorze prędkości czy też zużycie, pitting (wżery), zatarcie, pęknięcie lub uszkodzenie kół zębatych. Zapobiec takim uszkodzeniom, i związanym z nimi stratom finansowym i czasowym, można poprzez wykrywanie i klasyfikowanie błędów przekładni i odpowiednie opracowanie planów konserwacji. Niniejsze badanie dotyczy zastosowania metody głębokiego uczenia oraz splotowych sieci neuronowych (SSN) do monitoringu stanu przekładni na podstawie termogramów zarejestrowanych na stanowisku testowym pracującym przy różnych obciążeniach i prędkościach. Podejście oparte na uczeniu głębokim umożliwia efektywne wykorzystanie informacji o błędach pochodzących z dużych zbiorów danych i podejmowanie trafnych decyzji diagnostycznych. Niewiele z dostępnych publikacji poświęconych jest monitorowaniu stanu przekładni ślimakowych. Niniejsza praca jako pierwsza przedstawia badania przekładni ślimakowej z zastosowaniem termografii zamiast zwyczajowo prowadzonych pomiarów drgań i dźwięku, które mają pewne wady dotyczące wymagań sprzętowych, ograniczonych możliwości pomiarowych i głośności sygnałów. SSN opartą na danych termicznych porównano z siecią, którą uczono na zbiorach danych wibracyjnych i akustycznych pochodzących z prawidłowo działających i uszkodzonych przekładni ślimakowych. Wyniki diagnostyki uszkodzeń pokazują, że model SSN przekładni ślimakowej oparty na obrazie termicznym osiągnął stuprocentową (100%) skuteczność, podczas gdy skuteczność modeli opartych na danych wibracyjnych i akustycznych wyniosła, odpowiednio, 83,3% i 81,7%. Tym samym, model SSN oparty na obrazie termicznym pozwalał na trafniejsze diagnozowanie przekładni ślimakowej niż pozostałe modele. Ponadto zastosowanie metod opartych na termografii pozwala na poprawne monitorowanie stanu przy niższych kosztach pomiaru.

Słowa kluczowe: Diagnostyka błędów, przekładnie ślimakowe, termografia, splotowe sieci neuronowe, GoogLeNet, monitorowanie stanu.

1. Introduction

Industrial condition monitoring applications have improved a lot depending on the novel monitoring technologies and artificial intelligence decision making methods. In places where production continues uninterrupted such as factories, power generation facilities, lifting appliances, elevators, it is extremely important for systems to operate without any faults in terms of cost and work safety [31]. Due to the

increased complexity and precision of these systems and engineering applications, condition monitoring and reliability problems become more prominent [41]. During condition monitoring, the system is real time observed and real time measurements are done. By interpreting the measurements, whether there is a fault in the machine or not, is determined and the maintenance plan is applied accordingly. Establishing a reliable condition monitoring system, especially for gear faults,

ensures a smooth operation during the working of many industrial equipment [33].

Worm gear condition monitoring studies also plays a critical role in the maintenance plan since they are commonly used for power and motion transmission in many industrial applications and machines. They contain a mechanism consisting of worm screw and worm wheel that work together and differ from other types of reducers due to their lightness, simplicity and high gear ratio [11]. The main problem with a WG is how it transfers power. The spiral motion allows huge amounts of reduction in a comparatively small amount of space however this motion can also cause a problematic condition that is called sliding wear. During the working of a WG set, as the worm slides across the apex of the wheel, it slowly rubs off the lubricant film, and as a result, the worm rubs at the metal of the wheel in a boundary lubrication regime. When the worm surface leaves the wheel surface, it picks up more lubricant, and starts the process over again on the next revolution. This contact between the worm and the wheel in less lubricant conditions can cause wear and high temperature which need to be diagnosed. The most common problems faced with WGs are the high temperature at the speed reducer causing oil leakage, gear wearing, pitting, scoring and bearing fractures and damages. In order to prevent any damage, loss of time and money, serious precautions should be taken or predictive maintenance methods should be applied.

There are commonly used techniques such as vibration monitoring, acoustic monitoring, wear debris analysis, motor current analysis, in the literature for monitoring the conditions of gearboxes. Ghodake et al. [8] reviewed fault detection studies for worm gearboxes. Goyal et al. [9] performed a review study for condition monitoring and fault diagnosis for fixed axis gearboxes. Carden and Fanning [5] reviewed condition monitoring studies based on vibration analysis. Liu and Zhang [26] put ahead failure modes, condition monitoring and fault diagnosis methods for wind turbine bearings. Sait and Sharaf-Eldeen [35] conducted a review study on fault diagnostic and prognostic with the vibration analysis technique. Sharma et al. [37] reviewed different statistical condition indicators in time and frequency domains for gear faults. Lei et al. [18] reviewed studies of empirical mode decomposition method to diagnose faults of rotating machinery. Osman and Wang [32] proposed an improved Hilbert–Huang transform technique to diagnose faults of bearings using vibration signal analysis. They verified the effectiveness of the suggested technique in feature extraction and analysis by experimental tests. Wang et al. [45] proposed a hybrid technique by using complex wavelet transform for health diagnosis of rotating machines. They used numerical simulation and experimental studies under varying operating conditions to show the effectiveness of the hybrid technique. Loutas et al. [28] used on-line signal measurements and studied condition monitoring of a single-stage gearbox having artificial cracks. They used acoustic emission and vibration measurements to make use of this purpose. Zhang et al. [48] improved a new feature extraction method named Narrow-band Interference Cancellation to diagnose gear faults easier. Zhang and Zhao [49] proposed a compound fault detection approach based on time synchronous resampling and adaptive variational mode decomposition for gearboxes. With the experimental data analysis, they showed that the method is valid and can be used for fault detection of gearbox. Zhang et al. [50] proposed a fault diagnosis method based on singular value decomposition and radial basis function neural network to detect weak gear fault signals. Zhao et al. [51] proposed a local feature-based recurrent unit networks for monitoring health of machines. The method is a hybrid approach that integrates feature design with automatic feature learning.

Vibration monitoring has been considered as the most prevalent technique because it is easier to gather vibration data with vibration sensors [36]. Vibration sensors are versatile tools that measure acceleration for various applications [43]. However, the signals obtained from vibration sensors are very prone to the position of placements,

harsh working conditions and high temperature [10]. Vibration and acoustic measurements were also conducted together to improve the performance [27]. Thermal Imaging Technique (TIT) is introduced as a contactless, continuous and easy to implement technique [3] that can sense the radiation in a long-infrared range (9-14 μ m) of the electromagnetic spectrum and produce thermal images known as thermograms [29]. A single thermogram recorded in a very short time interval can contain multiple temperature points and provides information about the system being studied [34]. In addition, infrared thermography technique is used as a condition monitoring tool in which contactless and real-time abnormal temperature distribution can be detected [3]. Singh et al. [39] presented a comparative fault detection approach on asynchronous motors using the infrared thermography. They also detected in-turn and cooling system failures of induction motor with thermal imaging techniques [40]. Wakar and Demetgöl [46] collected thermal images for normal and faulty conditions under different speeds through an experimental system containing worm gearbox, and a multilayer perceptron was developed based on vibration and acoustic emission signals. Janssens et al. [13] detected eight different faults of bearings using thermal image based fault detection system and proposed two new features for thermal imaging of rotating machines. Al-Musawi et al. [2] developed a new coloring model and classified different bearing faults of a three-phase induction motor based on thermal image segmentation. Younus and Yang [47] developed a new intelligent detection system to classify different operating conditions based on the support vector machine and linear discriminant analysis using data from infrared thermography. Zhang et al. [50] used singular value decomposition and Radial Basis Function neural network for the detection of gear faults. In different condition monitoring studies, artificial neural network (ANN) based methods and computational modelling methods were used [22].

Despite the use of many fault detection and condition monitoring studies with the use of numerous diagnosing algorithms, such as support vector machines, empirical mode decomposition, Wigner-Ville distribution, short-time Fourier Transform, etc [22], as the evaluation data become larger, an efficient deep learning method, CNN, was introduced, in which extracting special features from the entire data is not needed. Nowadays, condition monitoring of mechanical faults in rotating components, are being prescribed by intelligent diagnosing methods such as artificial intelligence techniques and deep learning [9]. Deep learning is one of the newest machine learning techniques and has recently been used in condition monitoring studies and has been proved to be an effective methodology to improve the safety and reliability of gearboxes [17]. Zhao et al. [51] was the first who redefined the representation learning of raw data with the deep learning method. They proposed a local feature-based gated recurrent unit network for fault diagnosis. Li et al. [25] presented a modified deep learning method that can be used in case of limited data access and compared the method with traditional deep learning methods in terms of diagnostic success using datasets of two different gearboxes. Li et al. [22] proposed a new method named augmented convolution sparse auto-encoder to diagnose pitting faults of gears by using acoustic emission signals. Comparative analysis shows that the proposed method gives better results than fully-connected layer neural network and convolutional neural network. Few researches were carried out fault detection and classification with thermal image based deep learning methods. Li et al. [24] have used thermal imaging as a cost-effective and practical predictive maintenance method. In their study, they processed thermal images with a convolutional neural network (CNN) and allowed online remote condition monitoring of a gearbox. Their proposed method predicted faults on the gearbox with higher accuracy rate than their vibration-based counterparts. Janssens et al. [14] analyzed infrared thermal video images with deep learning algorithm to determine the condition of the machine automatically. Then, with deep learning, they detected the machine's faults and oil

level. Li et al. [23] proposed a CNN method based on infrared thermal images to detect faults of rotors and bearings on rotating machines and compared the results with conventional vibration or sound based approaches and other CNN types. Condition monitoring studies based on thermal images with CNN are limited to these publications. However, many studies have been carried out for fault detection with vibration and sound based deep learning. Jing et al. [15] proposed a CNN based feature learning and fault diagnosis method for gearboxes and showed that the CNN is more appropriate to learn features from vibration signal. Li et al. [19] proposed a deep random forest fusion method by using acoustic and vibration signals. They showed that the deep learning fusion can develop the ability of fault detection and diagnosis for gearbox. In another study, Li et al. [20] studied deep statistical feature learning to diagnose rotating machinery faults. The experiments show that the proposed method based on deep learning has the best fault classification rate and potential for diagnosing of rotating machinery faults.

This study presents a condition monitoring application with CNN, based on thermal images of a WG test rig which simulates different operating loads and speeds. The novelty of this study lies in the application of CNN for detecting faults on WGs from thermal imaging data. Vibration and sound signal measurement and analysis devices are relatively expensive than thermal cameras. Moreover, compared to signal analysing devices, a camera can examine the entire surface of WG. The operating speed and loading rate have a major impact on the gearbox temperature, sound and vibration behaviour. In order to distinguish the fault and the effect of working conditions, measurements should be done for all of the operating conditions. For this reason, in this study a test rig that can work at different speeds and loading rates was used to collect thermal images, vibration and sound signals from healthy and faulty WGs. Since an industrial WG is loaded at different rates and operated at different speeds in the real working environment, measurements at constant speed and load are only valid for a limited operating range. For comparison reasons, vibration and acoustic data were also gathered for all working conditions. The effects of varying resolution on performance of CNN are also depicted. As far as the authors' search, the results of the present study are of vital importance to the development of WG knowledge and industry.

The rest of this paper is formed as follows. Section 2 describes the experimental system, test rig and measurement elements. Section 3 describes the methodology and theoretical background of CNN-based fault diagnosis and proposed models for WG. Section 4 describes

experimental validation and performance comparison with proposed models. Finally, section 5 gives conclusions.

2. Experimental System

2.1. Test Rig

The test rig shown in Fig. 1 was built to verify the CNN-based fault diagnosis method proposed for WG. The test rig consists of three main units: control panel, loading device and single stage WG connected to an induction motor. The loading device, control panel and WG are mounted on a steel platform. The induction motor is driven by an inverter and coupled to the WG. An elastic coupling is used between the output shaft of the WG and an electromagnetic powder brake is the loading device. A radial fan protects the loading device from excessive heat.

The inverter is adjusted by a digital panel placed on the control panel. The DC loading device is powered by a transformer which has 5 different outputs so that the system can be loaded at different rates as 0%, 25%, 50%, 75% and 100%. Real working conditions are simulated with the inverter and the loading device powered by a multiple output transformer. The output speed of the gearbox is measured with a tachometer and the input voltage of the loading device with an AC/DC voltmeter.

2.2. Measurement Elements

Three different measuring devices were used in the experiments: thermal camera, accelerometer and microphone. The TESTO 880 thermal camera with 9 Hz image refreshing frequency (Fig. 2 (a)). Thermal sensitivity of the camera is less than 0.1 °C and it can focus manually up to a distance of 10 cm. It has 160x120 pixel detector and 32°x24° standard lens. The second one is a PCB ICP® type piezoelectric accelerometer for vibration data. (Fig. 2 (b)). The accelerometer has a sensitivity of 1.02 mV/(m/s²) and allows measurements in the frequency range of 0.5-10000 Hz. Phantom powered Behringer ECM8000 condenser microphone is used for sound measurements (Fig. 2 (c)). The microphone has 600-ohm resistance and -60 dB sensitivity within the frequency range of 15 Hz - 20 kHz. In addition, m + p VibPilot dynamic signal analyser with 4 analog outputs was used to collect and monitor data with accelerometer and microphone.

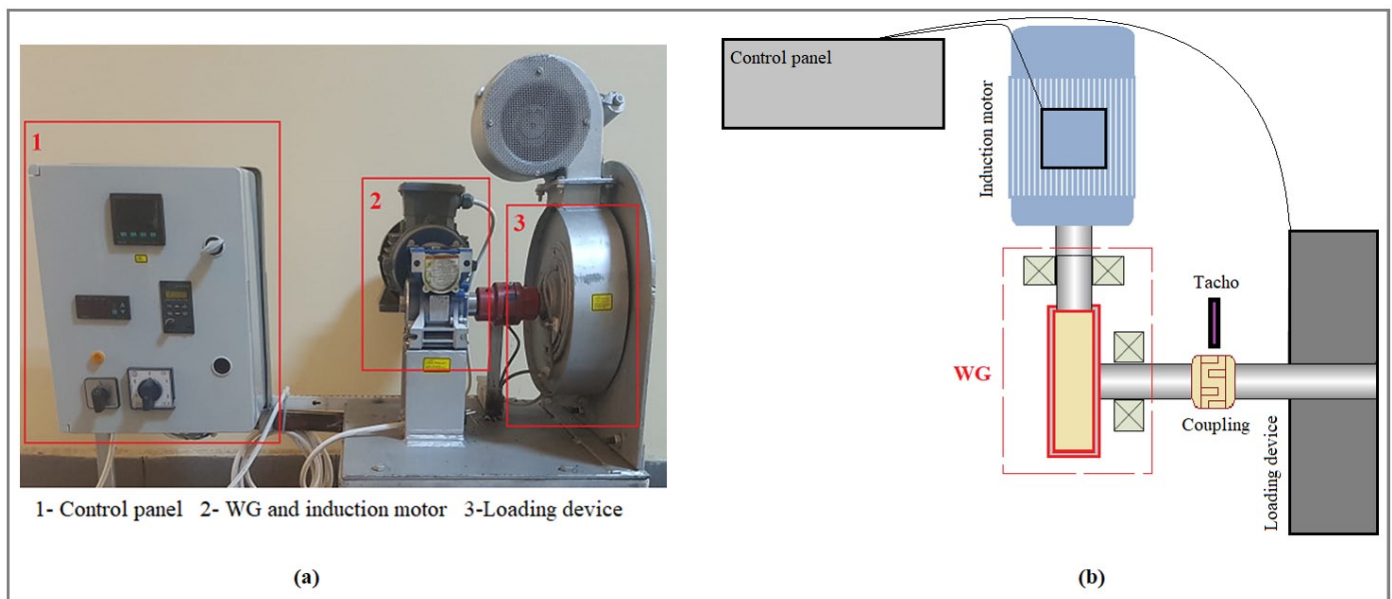


Fig. 1. Experimental setup: (a) test rig (b) the schematic diagram

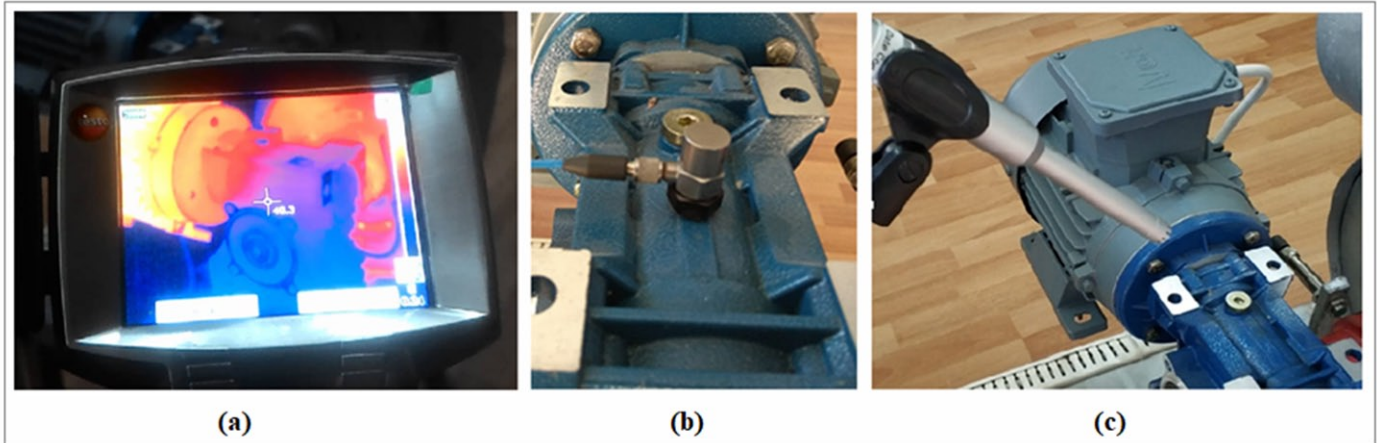


Fig. 2. (a) thermal camera, (b) accelerometer, (c) microphone

3. Methodology

This section presents the methodology used to classify WG faults that are healthy, wear, pitting and tooth breakage. Deep learning, one of the recent machine learning methods, is used as the main tool. As stated in previous works [22], deep learning do not require manual extraction of fault features and also achieve better fault detection results. It has the ability to reduce the number of network parameters through local weight sharing and to avoid the over-fitting of the network when the number of samples is insufficient.

3.1. Theoretical Background of CNN

Convolutional neural network (CNN) is one of the most used types of deep learning. CNN consists of convolutional and subsampling layers. Each of these layers has a certain topographic structure and the layers contain a different set of neurons. Each neuron is also linked to neurons in the previous layers. Fig. 3 shows a typical CNN architecture consisting of input, convolutional and subsampling layers, feature maps, fully connected layer and softmax regression (final stage). Convolutional and sub-sampling layers are arranged to reduce

the computation time and to create spatial and configurable invariance gradually [21].

Convolutional layers consist of a number of filters. These filters convolute input from the previous layer with a set of weights and create an output called a feature map. The neurons in the filters are connected to the input data points and these points are multiplied with the weights. Because all neurons in the same filter share their weight, the optimization time and complexity of CNN is reduced [12].

$$C_{cn} = f(X * W_{cn} + b_{cn}) \quad (1)$$

Assuming that the convolutional layer input is $X \in R^{M \times N}$, the layer output can be computed as in Eqs. 1. Here, M and N are dimensions of input data; $*$ is convolution operator; C_{cn} is the cn -th feature map of the convolutional layer; X is the input data matrix; W_{cn} is the weight matrix of cn -th filter of the actual layer; b_{cn} is the cn -th bias; and f is nonlinear activation function that applied to the result [15].

To reduce the size of the features and parameters of the network by subsampling, the pooling layers come after the convolution lay-

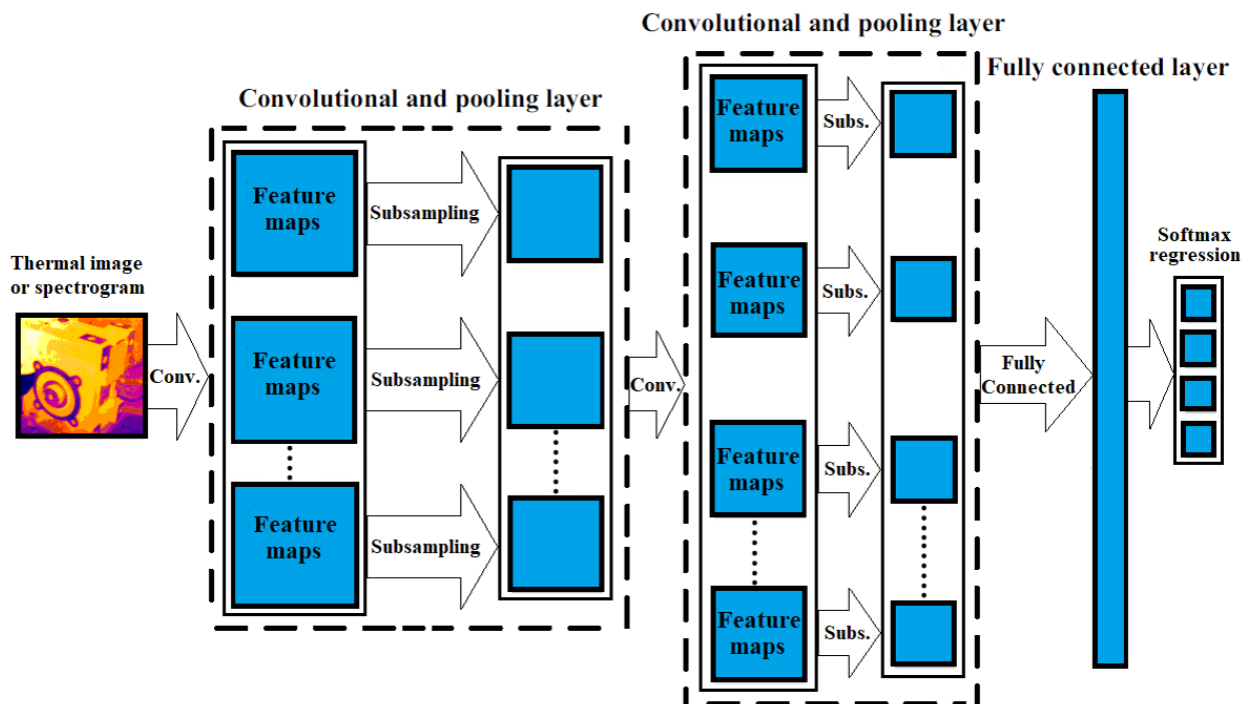


Fig. 3. Typical architecture of CNN

ers. There are three different pooling functions that calculate activation statistics: max pooling, mean pooling, and weighted pooling. Among these, max pooling is the most preferred function in CNN architecture [21]. S is pooling block size and $C_{cn} \in S$; P_{cn} is the output of the pooling layer; then max pooling activation can be written as in Eqs. 2 [15].

$$P_{cn} = \max C_{cn} \quad (2)$$

Finally, fully connected layer follows the combinations of convolutional and pooling layers. The fully connected layer is similar to a traditional neural network. So it can be applied to different classification problems. To achieve fast and accurate results, one hidden layer and softmax regression were chosen as the last layer. In this paper, as different gearboxes, which are healthy and faulty, are classified, the output of the softmax regression can be expressed as in Eqs. 3. Here, K is the label number, W_j is the weight matrix; b_j is bias; and O is the final result of the CNN [15].

$$O = \begin{bmatrix} P(y=1|x;W_1,b_1) \\ P(y=2|x;W_2,b_2) \\ \dots \\ P(y=K|x;W_K,b_K) \end{bmatrix} = \frac{1}{\sum_{j=1}^K \exp(W_j x + b_j)} \begin{bmatrix} \exp(W_1 x + b_1) \\ \exp(W_2 x + b_2) \\ \dots \\ \exp(W_K x + b_K) \end{bmatrix} \quad (3)$$

3.2. Fault Diagnosis Protocol

When gears are in operation they are subject to dynamic operating loads that affect the temperature, vibration and sound magnitudes directly [1]. Therefore, the high temperature values or the increase in vibration or sound amplitudes of WGs cannot always be a cause of a fault in condition monitoring applications. Considering all working loads and speeds of the test rig [46], it should be demonstrated that the increase in temperature is due to fault or normal operating load (Table 1).

Table 1. Gearbox output speeds, loading rates and measurements

	WG Output Speeds	Loading Rates	Measurements
Healthy and faulty WG	30 rpm	0%, 25%, 50%, 75% and 100%	Thermal imaging, vibration measurement, and sound measurement
	50 rpm		
	70 rpm		

One healthy and three faulty WGs are selected. Faulty WGs were made by machining techniques. Healthy (H) and faulty wheels of gearboxes are given in Fig. 4. Faulty gearboxes are labelled as F1 (wear), F2 (pitting) and F3 (tooth breakage).

3.2.1. Infrared Thermography

Infrared thermography (IRT) has become one of the most effective condition monitoring tools [3], as they can enable reliable fault diagnosing results. Real-time and non-contact measurement of the temperature of machine equipment and elements can be carried out with IRT. In this way, it is possible to eliminate the failures occurring in the machines without causing catastrophic damage and production losses [30]. Under actual operating conditions, WGs are heated depending on the loading rate, operating speed and environment temperature. Measuring the radiant thermal energy distribution emitted from the surface of the WG and converting this energy distribution into a surface thermogram constitutes the basis of the condition monitoring with infrared thermography and gives information about the current operating status and possible gearbox faults such as wear, pitting, tooth breakage [23]. Therefore, thermal images of healthy and faulty WGs with 160x120 pixel were collected for the working conditions given in Table 1 and then, they were fed as visual data to train and validate CNN.

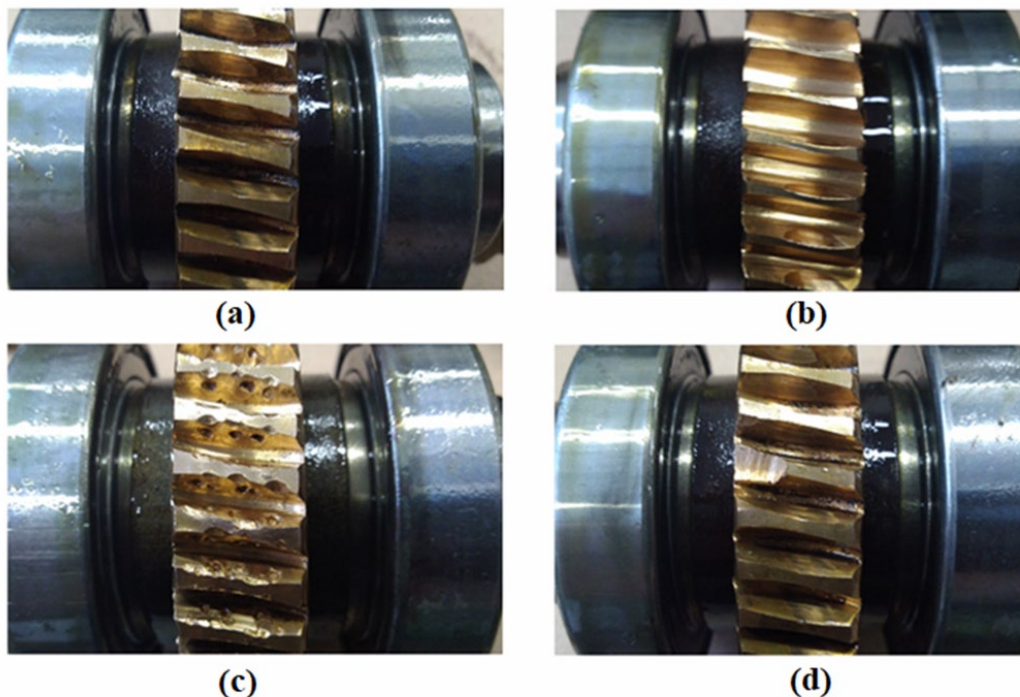


Fig. 4. Healthy (a), F1-wear (b), F2-pitting (c), F3-tooth breakage (d)

3.2.2. Short Time Fourier Transform and Spectrogram Images of Signals

Time-frequency analysis, one of the most common signal processing approach, can be used to understand the changes of sound and vibration signal components over time [6]. Fourier transform, is used to visualize the change of frequency components of the signal over time. This visual representation is called the spectrogram of the signal [38]. Vibration and sound spectrograms were fed as visual data for CNN. In this study, short time Fourier transform (STFT) was preferred in order to obtain spectrograms.

In STFT, the signal, the function of time, is divided into short segments and the Fourier transform is applied for each segment. In the case of continuous time, the signal is multiplied by a window function when applying STFT operation as seen in Eqs. 4. Here, $x(t)$ is the time domain signal and $w(t)$ is the window function [44]:

$$STFT \{x(t)\}(\tau, \omega) = \int_{-\infty}^{+\infty} x(t)w(t-\tau)e^{-j\omega t} dt \quad (4)$$

STFT for discrete time is given as in Eqs. 5. Here, $x[n]$ is the signal and $w[n]$ is the window function. Based on this, the spectrogram of the STFT function (Fig. 5) is calculated as in Eqs. 6 [44]:

$$STFT \{x[t]\}(m, \omega) = X(m, \omega) = \sum_{-\infty}^{+\infty} x[n]w[n-m]e^{-j\omega n} \quad (5)$$

$$Spectrogram \{x[t]\}(m, \omega) = |X(m, \omega)|^2 \quad (6)$$

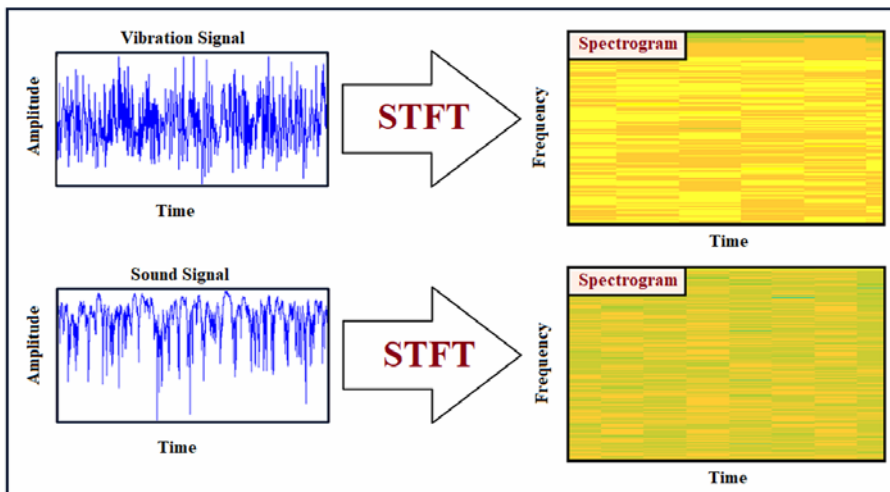


Fig. 5 Vibration and sound signals and spectrogram images

3.2.3. Fault Detection and Classification with Modified GoogLeNet Models

GoogLeNet architecture is used to construct CNN models in this study for the diagnosis of WG failures. GoogLeNet, offered by Szegedy et al. [42], is a deep convolutional neural network architecture based on new software technologies for classification and detection [42]. GoogLeNet is a pretrained CNN that has 22 layers deep, and the networks have an image input size of 224-by-224. The modified

GoogLeNet properties and detailed structure of modified architecture are given in Table 2 and Table A1.

It is commonly agreed that the CNN mostly performs well with an enormous amount of data. However, it is possible to leverage deep learning even with limited data [4]. In some instances, where you

Table 2. The modified GoogLeNet properties

Depth of layers	22
Number of layers	144
Number of connections	170
Input type	Image
Input size	224x224x3
Output type	Classification
Output size	4
Weight learn rate factor	10
Bias learn rate factor	10
Loss Function	Cross-entropy

can't gather more data due to cost or inconvenience, then there are some ways to follow during the training process. Fine tuning, data augmentation, using cosine loss function, or using autoencoders are some of them. This study has the data that can be considered as limited amount. So that, we did finetuning and data augmentation, respectively. In finetuning, we start with a pretrained model and updated all of the model's parameters for our new task, in essence retraining the whole model. All model parameters which are the number of classes in the dataset, the batch size used for training and the number of training

epochs we want to run are updated. The train model function trains for the specified number of epochs and after each epoch runs a full validation step. After each epoch, the training and validation accuracies are reported. For data augmentation, we used a standard GoogleNet architecture that change the inputs in such a way that provides new data while not changing the label value. The 160x120 size raw input images are converted to 224x224x3 image input size with data augmentation.

Four different deep learning models were selected in this study (Table 3). They are thermal image based CNN(IRT-CNN), vibration signal based CNN(V-CNN), sound signal based CNN(S-CNN) and vibration-sound signals based CNN(VS-CNN). IRT-CNN is trained with raw thermal images; V-CNN is trained with vibration spectrogram images; S-CNN is trained with sound spectrogram images; and VS-CNN is trained with both of vibration and sound spectrogram images. The technical parameters; validation frequency is selected as 6 Hz and the learning rate is 0.001 for all models in the table.

A simple flowchart of training process and the images used for the classification of faults with different CNN models are given in Fig. 6.

Accordingly, thermal images and signal spectrogram images of H, F1, F2 and F3 gearboxes are collected separately for all operating conditions at first. The data of each gearbox is divided into three groups as train, validation and test. GoogLeNet outputs are modified to classify healthy and faulty gearboxes. Using train and validation data, CNN

Table 3. Experimental parameters for different deep learning models

IRT-CNN	V-CNN	S-CNN	VS-CNN	Fault Diagnosis	
Input Parameters				Output Parameters	
Raw thermal images	Vibration spectrogram images	Sound spectrogram images	Vibration-sound spectrogram images	Healthy (H)	
				Fault 1 (F1)	
Train, Validation and Test Samples				Fault Detection	Fault 2 (F2)
50%; 25%; 25%					Fault 3 (F3)
Loading Rates (LR)		0%; 25%; 50%; 75%; 100%			
Gearbox output speed (GOS)		30 rpm; 50 rpm; 70 rpm			

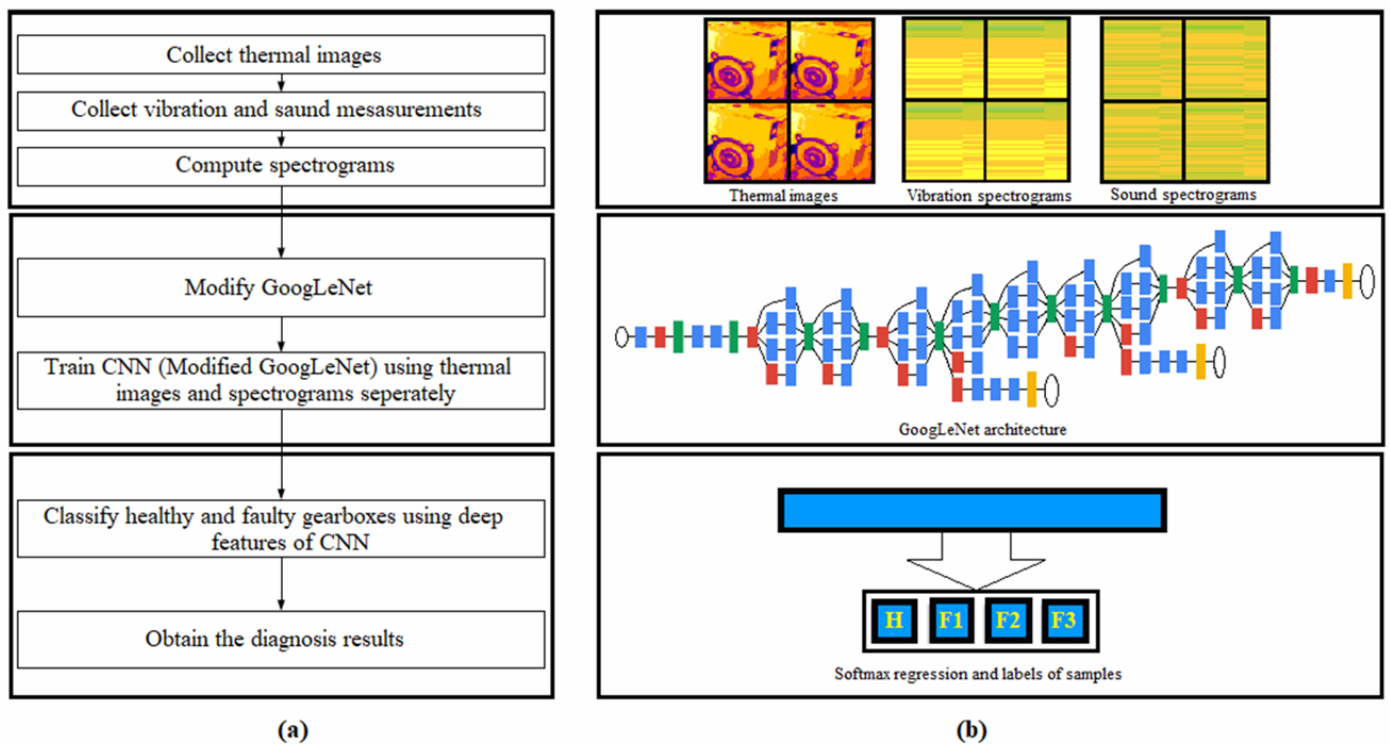


Fig. 6. CNN-based fault diagnosis of WG (a) flowchart of training process, and (b) Images

is trained and classification is performed based on deep features. Finally, the trained network is tested and fault diagnosis is carried out according to data labels.

4. Result and Discussions

This section contains the findings of the experiments and a brief discussion on the analysis based on past and current studies.

4.1. Findings of Thermal Imaging and Time-Frequency Signal Analysis

A total of 120 thermal images were collected in different loading rates (LR) and gearbox output speeds (GOS) for each gearbox. A total of 480 thermal images were obtained from all gearboxes for training, validation and testing of the IRT-CNN model.

The difference between the temperature distributions on the surfaces of healthy and faulty gearboxes increases gradually with increas-

ing LR and GOS. For all LR and GOS, the temperature distributions of F1, F2 and F3 gearboxes are higher than H. The thermal images of H, F1, F2 and F3 for GOS = 50 rpm and LR = 50% are given in Fig. 7. The maximum temperature for F1 is 105°C and the average surface temperature is 75°C. The hottest gearbox is observed as F1 during the experiments. The maximum temperature value for F2 is 93°C and the average surface temperature is 66°C. F2 is hotter than F3 for the same conditions. The maximum temperature value for F3 is 88°C and the average surface temperature is 64°C. The less heated gearbox is observed as H. The maximum temperature value for H is 83°C and the average surface temperature is 60 °C. For F1, F2, F3 and H gearboxes, the minimum surface temperature values measured under these conditions are 18.4°C; 16.7° C; 16.1°C; 15.7 °C, respectively. Accordingly, the worn gearbox is the hottest gearbox. Pitting failure causes more heat than the broken tooth. The healthy gearbox is the coolest one during experiments.

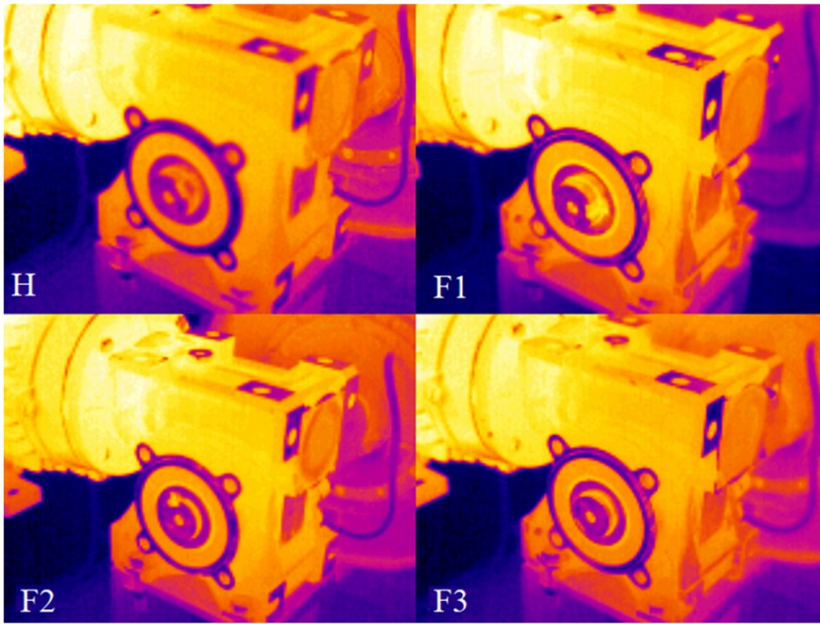


Fig. 7. The thermal images of H, F1, F2 and F3 gearboxes

The number of vibration spectrogram images were 120 for each gearbox in different LR and GOS. A total of 480 vibration spectrograms were obtained from all gearboxes for training, validation and testing of the V-CNN model. In Fig. 8, time waveforms and spectrogram images for H, F1, F2 and F3 gearboxes are given in the interval

of 0-1 s for LR = 50%, GOS = 50 rpm. The sampling rate for time-frequency analysis was chosen as 1600 Hz and frequencies higher than 1000 Hz were eliminated by Butterworth low-pass filter for vibration and sound analysis. Fig. 9, shows the time waveforms and spectrogram images of sound measurements for LR = 50%, GOS = 50 rpm. The time interval for sound waveform is 0-0.4 s. A total of 480 sound spectrogram images were also collected for training, validation and testing of the S-CNN model. It was seen that the amplitude of sound and vibration signals were increasing with increasing load and speed. Moreover, different faulty gearboxes produced different signal amplitudes as in Fig. 8 and Fig. 9 that the amplitudes of the sound and vibration signal of the faulty gearboxes are higher than the healthy gearbox. From the figures, F1 has the highest vibration amplitudes. These amplitudes vary from -31.7 m/s^2 to 42.6 m/s^2 in the 0-1 s time interval and 0-750 Hz frequency range. The H gearbox vibration signal amplitudes range from -28.4 m/s^2 to -29.4 m/s^2 and they are the lowest in the same conditions. The vibration amplitudes of the F2 generally vary from -32.4 m/s^2 to 30 m/s^2 and larger than amplitudes of the F3. In addition, F3 has the highest sound amplitudes, ranging from -13 dB to 75 dB in the 0-0.4 s time interval and 0-1900 Hz frequency range. H reducer has the lowest sound amplitudes varying from -10.1 dB to 68 dB in the same conditions. The sound amplitudes of the F2 are generally between -9 dB and 72 dB which are larger than the amplitudes of the F3.

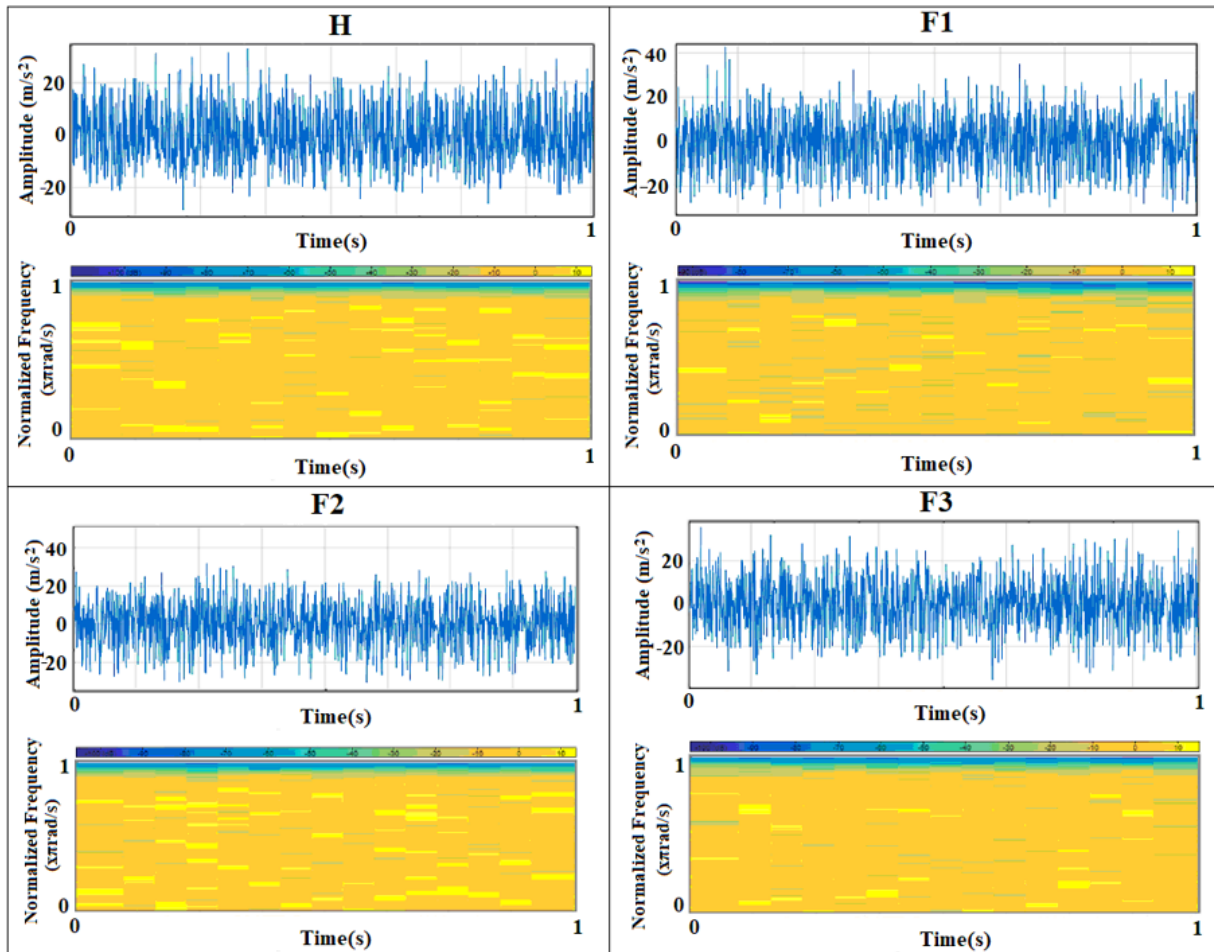


Fig. 8. Time waveforms and spectrogram images of vibration measurements

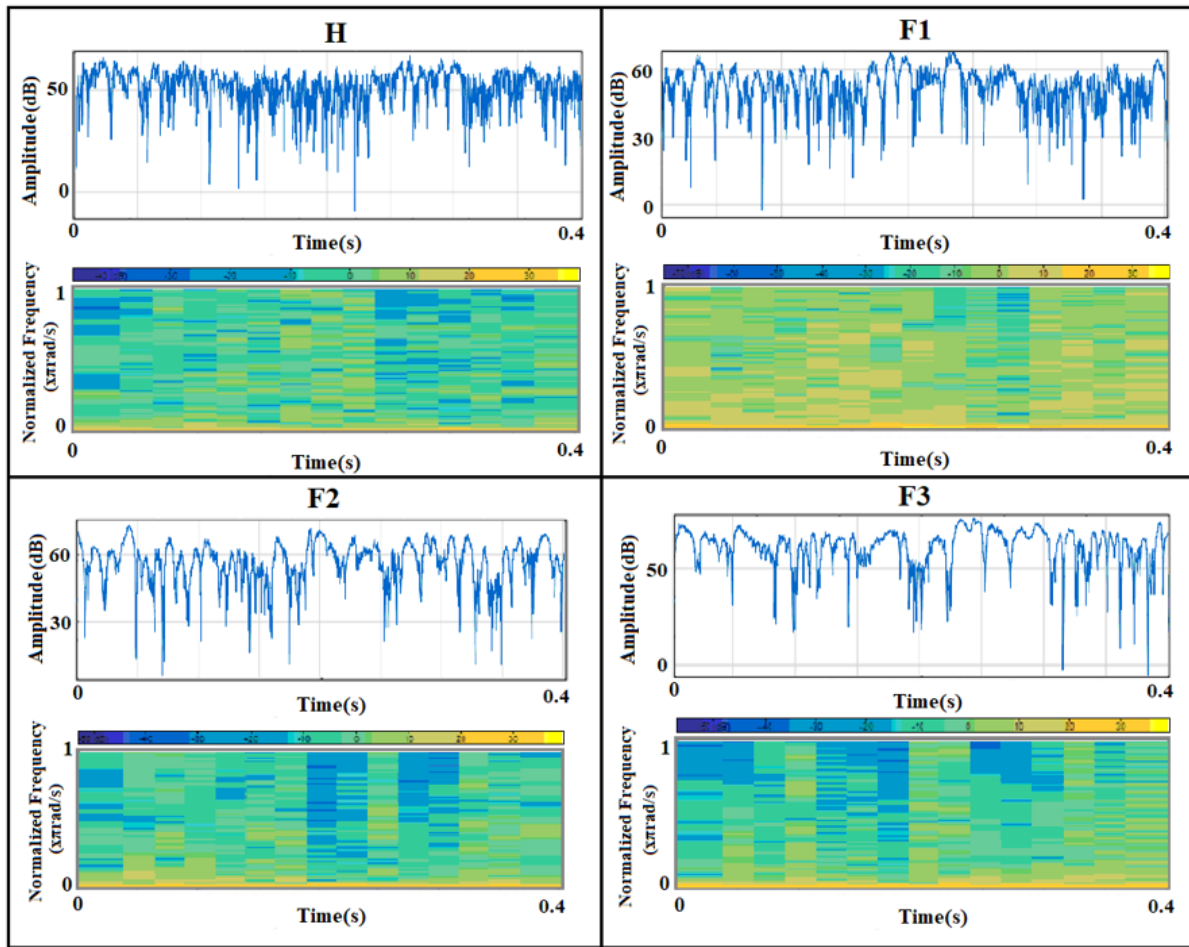


Fig. 9. Time waveforms and spectrogram images of sound measurements

4.2. Results of Fault Diagnosis with Different CNN Models

The spectrogram images of vibration and sound signals were used together in training, validation and test of the VS-CNN model. The targets for CNN models are healthy and faulty gearboxes. Table 4 shows class labels of gearboxes and the number of samples used for training, validation and testing of different CNN models.

Fig. 10 shows the accuracy rates for different CNN models. Accordingly, the highest training and validation accuracy rates were achieved with the IRT-CNN model based on thermal images. It reached 100% in the 30th iteration and 10 epochs. This success rate is 87.5% for V-CNN and 81.67% for S-CNN model. VS-CNN, which uses spectrogram images of vibration and sound samples together, has the lowest training and validation accuracy with 73.33%. Li et al. [24] achieved similar results in their study for the condition monitoring of bevel gearbox. However, they did not use spectrograms

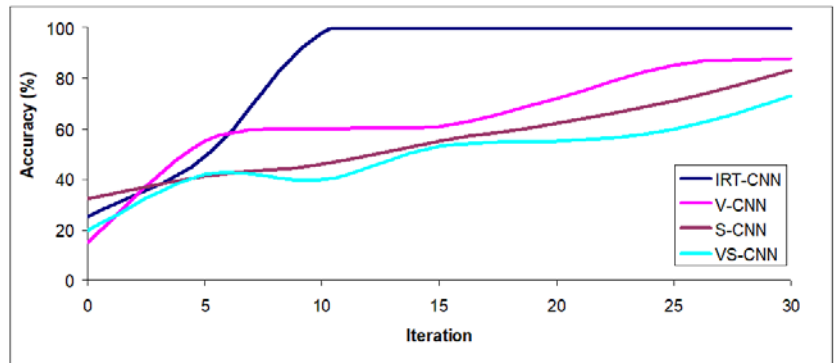


Fig. 10. Accuracy rates for different CNN models

Table 4. Class labels and number of samples for CNN models

Class Label	Number of Training Samples	Number of Validation Samples	Number of Test Samples
H	60	30	30
F1	60	30	30
F2	60	30	30
F3	60	30	30

for training and validation of CNN models. In addition, their experiments were carried out under constant speed and loading rate, and the effect of real working conditions on temperature change was not taken into account sufficiently (Table 6, Table 7).

Fig. 11 shows the losses in training and validation progresses for different CNN models. Losses are a quantitative measure of the difference between the predicted output and the actual output,

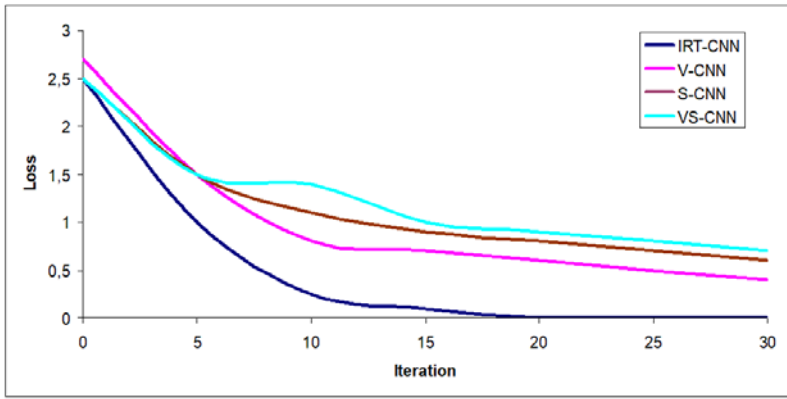


Fig. 11. Losses for different CNN models

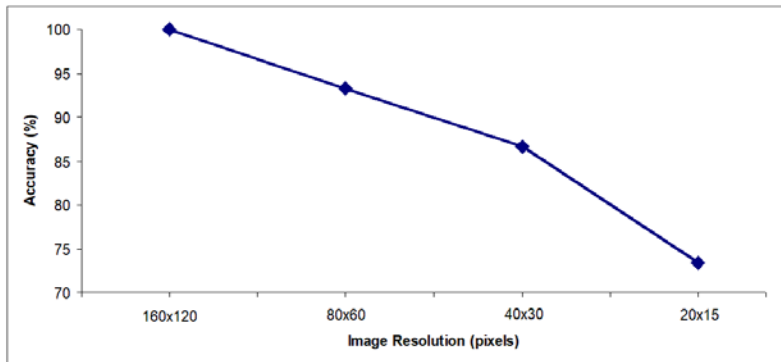


Fig. 12. The change of accuracy rates due to thermal image resolutions

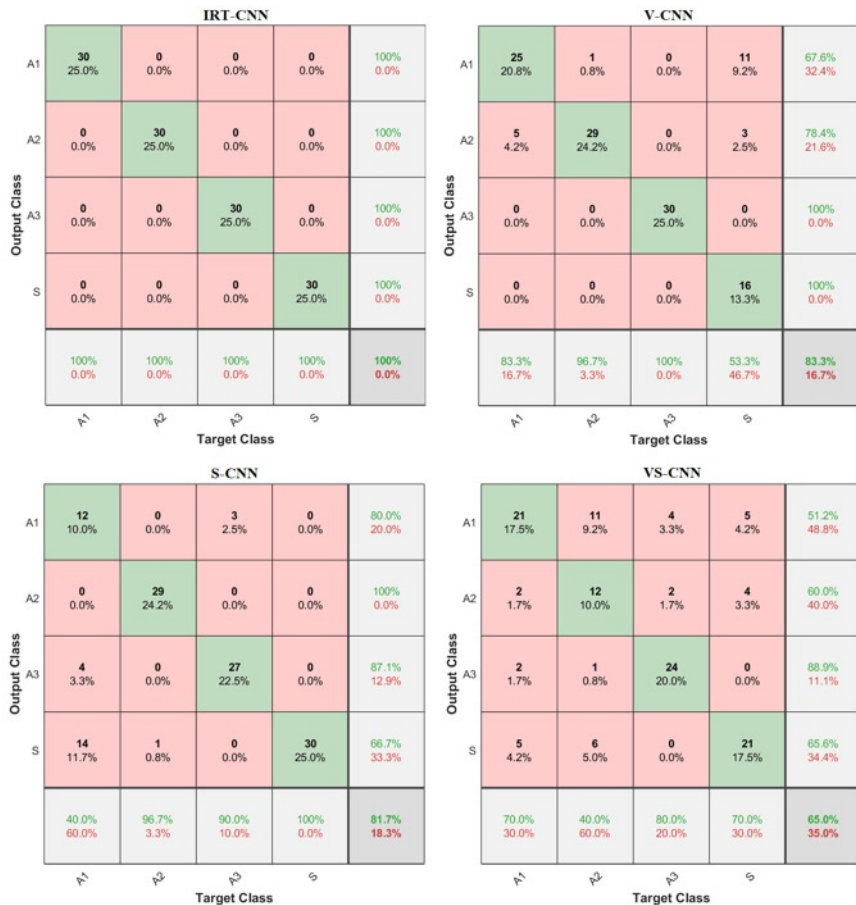


Fig. 13. Confusion matrices of CNN models

and decrease with the number of iterations. Depending on the suitability and adequacy of the data set, losses and the number of iterations decrease. Accordingly, the highest loss occurred during the training progress of the VS-CNN model. Training and verification loss for the IRT-CNN model decreased to 0% in the 30th iteration and 10 epochs.

To test CNN models, 25% of the samples collected from the gearboxes was used. Confusion matrices of different CNN models are given in Fig. 13. The IRT-CNN model classified all test inputs for H, F1, F2 and F3 gearboxes correctly. For this success rate, the test was repeated 10 times and the standard deviation was found to be 0. Whereas the confusion matrix of V-CNN model's success rate is 83.3% and half of the samples of the H gearbox is estimated incorrectly. When the S-CNN model is tested, 81.7% of the samples are estimated correctly. The rate of misclassification of F1 is high. The lowest test success rate belongs to VS-CNN model. In this model, only 65% of all test samples were estimated correctly. It can be inferred from confusion matrix that the test samples of the A2 gearbox are classified incorrect considerably.

Table 5 shows the success rates of CNN models trained and tested with different numbers of samples. Accordingly, as the number of samples increases, validation and test accuracy rates and training time of all CNN models increase. The highest validation and test accuracy rates for all models are achieved when 120 samples are used. Validation and test accuracy rates of IRT-CNN models are higher than all V-CNN, S-CNN and VS-CNN models. It is notable that even if the IRT-CNN training sample numbers are 30 or 60, it still has a very high diagnostic success (89-90%). The accuracy rates with the increasing amount of training and testing data of V-CNN and S-CNN models are more affected compared to IRT-CNN models. Even with the least number of samples, IRT-CNN has 89.3 % accuracy.

One other topic searched in this study is the resolution of images that are to be classified. Resolution of the image varies due to different input sources, different imaging devices and different environment noises [16]. Variation in images resolution alters the visual information of images [7]. Fig. 12 shows the variation of accuracy rates depending on the image resolutions. When the performance comparison of classifier for IRT-CNN dataset analysed from Figure 12, it is noticeable that accuracy rate decreases when image resolution decreases. The results are in accordance with the literature [7].

These results have shown that the IRT-CNN algorithm is effective for using the original temperature values for different fault diagnosis. It is worth noting that this study is conducted as a special case study in a factory where the unwanted working noises are surpassed. However, the system was tried to be made as a real as possible. Further effort can be put and the method can be tried on a real shift.

Table 6 and Table 7 show a general comparison of CNN models proposed by Li et al. [24] and CNN models proposed from this study.

Table 5. Effect of the number of samples on the performance of CNN models

CNN Models	Number of Samples	Validation Accuracy (%)	Test Accuracy (%)	Training Time (s)
IRT-CNN	30	90.63	89.3	122
IRT-CNN	60	95	90	123
IRT-CNN	120	100	100	366
V-CNN	30	59.38	57.1	129
V-CNN	60	65	61.7	145
V-CNN	120	87.50	83.3	456
S-CNN	30	59.38	42.9	127
S-CNN	60	58.33	41.7	142
S-CNN	120	81.67	81.7	395
VS-CNN	30	40.63	39.3	128
VS-CNN	60	46.67	55	147
VS-CNN	120	73.33	65	423

Table 6. The proposed CNN models in the study

CNN Models	Test Accuracy (%)	Training Time (s)
IRT-CNN	100	366
V-CNN	83.3	456
S-CNN	81.7	395
VS-CNN	65	423

Table 7. CNN models proposed by Li et al. [24]

CNN Models	Test Accuracy (%)	Training Time (s)
IRT-CNN	100	470
V-CNN	71.53	542

5. Conclusions

In this paper, a CNN based condition monitoring study based on infrared thermal images, vibration spectrogram images and sound spectrogram images was performed and the model is used to diagnose different types of faults in WGs. In real working conditions with a test rig, which can operate at different speed and loading rates, the CNN model based on thermal imaging yielded the most successful results than other models. Thermal images collected at different loading rates and different speeds increased the success rate of IRT-CNN. The operating conditions directly affect the temperature of the gearbox and the thermal image pattern. Also, different types of faults tend

to show different temperature distribution under different operating conditions. The results showed that healthy and faulty gearbox can be correctly classified 100% with the IRT-CNN model; 83.3% with the V-CNN model; 81.7% with the S-CNN model; and 65% with the VS-CNN model. Based on this comparison, it was found that the use of thermal images with convolutional neural networks (CNN) generates the highest classification rates. With the high resolution detector thermal cameras, successful fault detection can be made with small data sets considering different operating conditions. A new possible advantage of the IRT-CNN model is that it allows remote fault diagnosis. It can be useful to work with the thermal cameras, which are relatively cheaper than other signal processing devices. Faults of WG can be accurately detected with loss of cost, time and money.

References

- Al-Arbi S K. Condition Monitoring of Gear Systems using Vibration Analysis, Doctoral Dissertation. University of Huddersfield, 2012.
- Al-Musawi A K, Anayi F, Packianather M. Three-Phase Induction Motor Fault detection based on Thermal Image Segmentation. *Infrared Physics & Technology* 2019; 104: 103140, <https://doi.org/10.1016/j.infrared.2019.103140>.
- Bagavathiappan S, Lahiri B B, Saravanan T, Philip J, Jayakumar T. Infrared thermography for condition monitoring, A review. *Infrared Physics & Technology* 2013; 60: 35-55, <https://doi.org/10.1016/j.infrared.2013.03.006>.
- Barz B, Denzler J. Deep learning on small datasets without pre-training using cosine loss. *The IEEE Winter Conference on Applications of Computer Vision* 2020; 1371-1380, <https://doi.org/10.1109/WACV45572.2020.9093286>.
- Carden E P, Fanning P. Vibration based condition monitoring: a review. *Structural health monitoring* 2004; 3(4): 355-377, <https://doi.org/10.1177/1475921704047500>.
- Cohen L. *Time-Frequency Analysis*, New Jersey: Prentice Hall PTR, 1995.
- Dodge S, Karam L. Understanding how image quality affects deep neural networks. In 2016 eighth international conference on quality of multimedia experience (QoMEX) IEEE 2016; 1-6, <https://doi.org/10.1109/QoMEX.2016.7498955>.
- Ghodake S, Mishra A K, Deokar A V. A Review Paper on Fault Detection of Worm Gearbox. *International Advanced Research Journal in Science, Engineering and Technology* 2016; 3(1): 161-164.
- Goyal D, Pabla B S, Dhami S S. Condition monitoring parameters for fault diagnosis of fixed axis gearbox: a review. *Archives of Computational Methods in Engineering* 2017; 24(3): 543-556, <https://doi.org/10.1007/s11831-016-9176-1>.
- Goyal D, Pabla B S, Dhami S S. Non-contact sensor placement strategy for condition monitoring of rotating machine-elements. *Engineering Science and Technology* 2019; 22(2): 489-501, <https://doi.org/10.1016/j.jestch.2018.12.006>.
- Haberhauer H, Bodenstern F. *Maschinenelemente*. Berlin: Springer, 2009.

12. Hoang D T, Kang H J. Rolling element bearing fault diagnosis using convolutional neural network and vibration image. *Cognitive Systems Research* 2019; 53: 42-50, <https://doi.org/10.1016/j.cogsys.2018.03.002>.
13. Janssens O, Schulz R, Slavkovikj V, Stockman K, Loccufier M, Van de Walle R, Van Hoecke S. Thermal image based fault diagnosis for rotating machinery. *Infrared Physics & Technology* 2015; 73: 78-87, <https://doi.org/10.1016/j.infrared.2015.09.004>.
14. Janssens O, Van de Walle R, Loccufier M, Van Hoecke S. Deep learning for infrared thermal image based machine health monitoring. *IEEE/ASME Transactions on Mechatronics* 2017; 23(1): 151-159, <https://doi.org/10.1109/TMECH.2017.2722479>.
15. Jing L, Zhao M, Li P, Xu X. A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox. *Measurement* 2017; 111: 1-10, <https://doi.org/10.1016/j.measurement.2017.07.017>.
16. Kannoja S P, Jaiswal G. Effects of Varying Resolution on Performance of CNN based Image Classification: An Experimental Study. *International Journal of Computer Sciences and Engineering* 2018; 6(9): 451-456, <https://doi.org/10.26438/ijcse/v6i9.451456>.
17. Kłosowski G, Rymarczyk T, Kania K, Świć A, Cieplak T. Maintenance of industrial reactors supported by deep learning driven ultrasound tomography. *Eksploracja i Niezawodność - Maintenance and Reliability* 2020; 22(1): 138-147, <https://doi.org/10.17531/ein.2020.1.16>.
18. Lei Y, Lin J, He Z, Zuo M J. A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing* 2013; 35(1-2): 108-126, <https://doi.org/10.1016/j.ymssp.2012.09.015>.
19. Li C, Sanchez R V, Zurita G, Cerrada M, Cabrera D, Vásquez R E. Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals. *Mechanical Systems and Signal Processing* 2016; 76: 283-293, <https://doi.org/10.1016/j.ymssp.2016.02.007>.
20. Li C, Sánchez R V, Zurita G, Cerrada M, Cabrera D. Fault diagnosis for rotating machinery using vibration measurement deep statistical feature learning. *Sensors* 2016; 16(6): 895, <https://doi.org/10.3390/s16060895>.
21. Li C, Sanchez R V. Gearbox fault identification and classification with convolutional neural Networks. *Shock and Vibration* 2015; Article ID: 390134, <https://doi.org/10.1155/2015/390134>.
22. Li X, Li J, He D, Qu Y. Gear pitting fault diagnosis using raw acoustic emission signal based on deep learning. *Eksploracja i Niezawodność - Maintenance and Reliability* 2019; 21(3): 403-410, <https://doi.org/10.17531/ein.2019.3.6>.
23. Li Y, Du X, Wan F, Wang X, Yu H. Rotating machinery fault diagnosis based on convolutional neural network and infrared thermal imaging. *Chinese Journal of Aeronautics* 2020; 33(2): 427-438, <https://doi.org/10.1016/j.cja.2019.08.014>.
24. Li Y, Gu J X, Zhen D, Xu M, Ball A. An Evaluation of Gearbox Condition Monitoring Using Infrared Thermal Images Applied with Convolutional Neural Networks. *Sensors* 2019; 19(9): 2205, <https://doi.org/10.3390/s19092205>.
25. Li Y, Wang K. Modified convolutional neural network with global average pooling for intelligent fault diagnosis of industrial gearbox. *Eksploracja i Niezawodność - Maintenance and Reliability* 2020; 22(1): 63-72, <https://doi.org/10.17531/ein.2020.1.8>.
26. Liu Z, Zhang L. A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings. *Measurement* 2020; 149: 107002, <https://doi.org/10.1016/j.measurement.2019.107002>.
27. Loutas T H, Roulias D, Pauly E, Kostopoulos V. The combined use of vibration, acoustic emission and oil debris on-line monitoring towards a more effective condition monitoring of rotating machinery. *Mechanical Systems and Signal Processing* 2011; 25(4): 1339-1352, <https://doi.org/10.1016/j.ymssp.2010.11.007>.
28. Loutas T H, Sotiriades G, Kalaitzoglou I, Kostopoulos V. Condition monitoring of a single-stage gearbox with artificially induced gear cracks utilizing on-line vibration and acoustic emission measurements. *Applied Acoustics* 2009; 70(9): 1148-1159, <https://doi.org/10.1016/j.apacoust.2009.04.007>.
29. Maldague X. *Theory and Practice of Infrared Technology for Non-Destructive Testing*. New York: John Wiley-Interscience, 2001.
30. Meola C. *Infrared thermography recent advances and future trends*. Sharjah: Bentham Science Publishers, 2012, <https://doi.org/10.2174/97816080514341120101>.
31. Mohanty A R. *Machinery Condition Monitoring: Principles and Practices*. New York: CRC Press, Taylor&Francis Group, 2015, <https://doi.org/10.1201/9781351228626>.
32. Osman S, Wang W. An enhanced Hilbert-Huang transform technique for bearing condition monitoring. *Measurement Science and Technology* 2013; 24(8): 085004, <https://doi.org/10.1088/0957-0233/24/8/085004>.
33. Randall R B. *Vibration-based Condition Monitoring: Industrial, Aerospace and Automotive Applications*. West Sussex: John Wiley & Sons Ltd., 2011, <https://doi.org/10.1002/9780470977668>.
34. Ring E F J, Ammer K. Infrared thermal imaging in medicine. *Physiological Measurement* 2012; 33(3): R33, <https://doi.org/10.1088/0967-3334/33/3/R33>.
35. Sait A S, Sharaf-Eldeen Y I. A review of gearbox condition monitoring based on vibration analysis techniques diagnostics and prognostics. In: Proulx T (Ed.), *Rotating Machinery, Structural Health Monitoring, Shock and Vibration* 2011; 5: 307-324, https://doi.org/10.1007/978-1-4419-9428-8_25.
36. Sasai S, Zhen Y X, Suetake T, Tanita Y, Omata S, Tagami H. Palpation of the skin with a robot finger: an attempt to measure skin stiffness with a probe loaded with a newly developed tactile vibration sensor and displacement sensor. *Skin Research and Technology* 1999; 5(4): 237-246, <https://doi.org/10.1111/j.1600-0846.1999.tb00136.x>.
37. Sharma V, Parey A. A review of gear fault diagnosis using various condition indicators. *Procedia Engineering* 2016; 144: 253-263, <https://doi.org/10.1016/j.proeng.2016.05.131>.
38. Sharma V, Parey A. Frequency domain averaging based experimental evaluation of gear fault without tachometer for fluctuating speed conditions. *Mechanical Systems and Signal Processing* 2017; 85: 278-295, <https://doi.org/10.1016/j.ymssp.2016.08.015>.
39. Singh G, Kumar T C A, Naikan V N A. Induction motor inter turn fault detection using infrared thermographic analysis. *Infrared Physics & Technology* 2016; 77: 277-282, <https://doi.org/10.1016/j.infrared.2016.06.010>.
40. Singh G, Naikan V N A. Infrared thermography based diagnosis of inter-turn fault and cooling system failure in three phase induction motor. *Infrared Physics & Technology* 2017; 87: 134-138, <https://doi.org/10.1016/j.infrared.2017.10.007>.
41. Sun B, Li M M, Liao B P, Yang X, Cao Y T, Cui B F, Feng Q, Ren Y, Yang, D Z. Time-variant reliability modeling based on hybrid non-probability method. *Archive of Applied Mechanics* 2020; 90(2): 209-219, <https://doi.org/10.1007/s00419-019-01605-1>.
42. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: 28th IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston MA, 2015: 1-9, <https://doi.org/10.1109/CVPR.2015.7298594>.

43. Tanaka Y, Horita Y, Sano A. Finger-mounted skin vibration sensor for active touch. In: Isokoski P, Springare J. (Eds.), EuroHaptics: International Conference on Human Haptic Sensing and Touch Enabled Computer Applications. Tampere, Finland, 2012: 169-174, https://doi.org/10.1007/978-3-642-31404-9_29.
44. Vashisht R K, Peng Q. Crack detection in the rotor ball bearing system using switching control strategy and Short Time Fourier Transform. *Journal of Sound and Vibration* 2018; 432: 502-529, <https://doi.org/10.1016/j.jsv.2018.06.061>.
45. Wang J, Gao R X, Yan R. Multi-scale enveloping order spectrogram for rotating machine health diagnosis. *Mechanical Systems and Signal Processing* 2014; 46(1): 28-44, <https://doi.org/10.1016/j.ymssp.2013.06.001>.
46. Waqar T, Demetgul M. Thermal analysis MLP neural network based fault diagnosis on worm gears. *Measurement* 2016; 86: 56-66, <https://doi.org/10.1016/j.measurement.2016.02.024>.
47. Younus A M, Yang B S. Intelligent fault diagnosis of rotating machinery using infrared thermal image. *Expert Systems with Applications* 2012; 39(2): 2082-2091, <https://doi.org/10.1016/j.eswa.2011.08.004>.
48. Zhang X, Kang J, Bechhoefer E, Zhao J. A new feature extraction method for gear fault diagnosis and prognosis. *Eksplatacja i Niezawodnos - Maintenance and Reliability* 2014; 16(2): 295-300.
49. Zhang X, Zhao J. Compound fault detection in gearbox based on time synchronous resample and adaptive variational mode decomposition. *Eksplatacja i Niezawodnos - Maintenance and Reliability* 2020; 22(1): 161-169, <https://doi.org/10.17531/ein.2020.1.19>.
50. Zhang Q, Zhao W, Xiao S G. Fault Diagnosis of Gear Based on Singular Value Decomposition and RBF Neural Network. *2nd International Conference on Frontiers of Sensors Technologies* 2017; 470-474, <https://doi.org/10.1109/ICFST.2017.8210559>.
51. Zhao R, Wang D Z, Yan R Q, Mao K Z, Shen F, Wang J J. Machine Health Monitoring Using Local Feature-Based Gated Recurrent Unit Networks. *IEEE Transactions on Industrial Electronics* 2018; 65: 1539-1548, <https://doi.org/10.1109/TIE.2017.2733438>.

Appendix A

Table A1. Detailed structure of modified GoogLeNet

Layer type	Patch size/ stride	Output size	Depth	#1x1	#3x3 reduce	#3x3	#5x5 reduce	#5x5	Pool proj.
Convolution	7x7/2	112x112x64	1	-	-	-	-	-	-
Max pool	3x3/2	56x56x64	0	-	-	-	-	-	-
Convolution	3x3/1	56x56x192	2	-	64	192	-	-	-
Max pool	3x3/2	28x28x192	0	-	-	-	-	-	-
Inception(3a)	-	28x28x256	2	64	96	128	16	32	32
Inception(3b)	-	28x28x480	2	128	128	192	32	96	64
Max pool	3x3/2	14x14x480	0	-	-	-	-	-	-
Inception(4a)	-	14x14x512	2	192	96	208	16	48	64
Inception(4b)	-	14x14x512	2	160	112	224	24	64	64
Inception(4c)	-	14x14x512	2	128	128	256	24	64	64
Inception(4d)	-	14x14x528	2	112	144	288	32	64	64
Inception(4e)	-	14x14x832	2	256	160	320	32	128	128
Max pool	3x3/2	7x7x832	0	-	-	-	-	-	-
Inception(5a)	-	7x7x832	2	256	160	320	32	128	128
Inception(5b)	-	7x7x1024	2	384	192	384	48	128	128
Avg pool	7x7/1	1x1x1024	0	-	-	-	-	-	-
Dropout (40%)	-	1x1x1024	0	-	-	-	-	-	-
Fully connected	-	1x1x4	1	-	-	-	-	-	-
Softmax	-	1x1x4	0	-	-	-	-	-	-

Yunus Emre KARABACAK

General Directorate of Tea Enterprises
53080 Rize, Turkey

Nurhan GÜRSEL ÖZMEN

Levent GÜMÜŞEL

Karadeniz Technical University, Mechanical Engineering Department
61080 Trabzon, Turkey

E-mails: Y_emre_karabacak@hotmail.com, gnurhan@ktu.edu.tr, gumusel@ktu.edu.tr