Acoustic Emissions:

Diagnosing Tribological Phenomenon in Artificial Joint Materials

by

KHADIJAT ABIOLA OLORUNLAMBE

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Department of Mechanical Engineering

School of Engineering

University of Birmingham

JULY 2022

Abstract

Studies have shown that many reported causes of failure of artificial joints such as hip, knee and spine are wear and friction related. Current modes of diagnosing failed artificial joints involve the use of imaging techniques like X-rays and CT scans, which although effective, are costly, time-consuming and harmful to patient health due to frequent exposure to radiation. There is the added limitation of the delay experienced before signs of failure become visible, causing further discomfort to the patient and, at times, health complications resulting from possible migration of wear debris into blood tissues. These complications have necessitated the need for a simpler and more dynamic system for identifying and diagnosing failed artificial joints, which is where the acoustic emission (AE) testing has shown promise.

AE testing is a non-destructive test method used to detect the onset and progression of mechanical flaws that has proven advantageous in the analysis and understanding of tribological interactions in mechanical systems. In recent times, it has been increasingly used in the study of the tribology of artificial and natural human joints thereby showing potential as a tool for the identification and diagnosis of failed artificial joints. Thus, this research aimed to use AE to monitor the tribological characteristics of artificial joint materials as a first step toward using AE to diagnose artificial and natural joint pathologies.

To gain an initial understanding of how AE features can be related to tribological mechanisms such as friction, in particular, a bio-tribo-acoustic tests system was developed. This enabled the acquisition of AE signals during biotribological testing of artificial joint materials. This proof-of-concept study showed that time-dependent (TDD) AE features can be used to predict the friction profile of a simulated polymer-metal artificial joint articulation. The prediction was carried out using a Non-linear Auto Regression with Exogeneous inputs (NARX) model. During testing of the trained model, predicted data had R^2 values of 94% in tests on PEEK reciprocating at 2 Hz test and 98.6% for UHMWPE at 2 Hz. These regression

results support the hypothesis that AE TDD features can be used to predict the friction profile which can then be related to the wear behaviour of the simulated joint articulation.

Having proved the potential of AE as a biotribological diagnostic tool, the next step is to be able to use the acquired AE signals to identify the perceived damage mode prompting the need for a method by which AE signals can be differentiated according to different wear mechanisms. To this end, AE signals from adhesive and abrasive wear, simulated under controlled joint conditions, were classified using supervised learning. Principal component analysis was used to derive uncorrelated AE features and then classified using three methods – logistic regression, k-nearest neighbours and back propagation (BP) neural network. The BP network emerged as the best performing network with a classification accuracy of 98%.

One of the limitations of traditional artificial neural networks (ANN) such as the BP network is the complex feature engineering required to obtain a model with high accuracy and high sensitivity. To mitigate this, deep transfer learning, with GoogLeNet as the base convolutional neural network (CNN) model, was used to classify AE signals from simulated damage mechanisms observed in retrieved polyethylene inserts of failed knee implants - burnishing and scratching wear. It was found that using CNN to extract features to be trained with an SVM model obtained a higher classification accuracy (99.3%) than just training with CNN model (96.5%).

The work presented in this thesis has shown that AE testing can be used to monitor the tribological properties of simulated articulating joint surfaces. With machine learning and deep transfer learning techniques, models with high accuracy and high sensitivity can be built to classify the acquired AE signals based on simulated real-life artificial joint damage modes. This confirms the initial hypothesis that with AE testing, a more dynamic, highly specific and highly sensitive process of identifying and diagnosing artificial joint pathologies can be developed, thereby reducing patient discomfort and NHS expenditure.

To my parents and siblings

- for their continued support and encouragement

Acknowledgements

First and foremost, my gratitude goes to Allah for making it possible for me to complete my PhD study, Alhamdulillah Robbil Al-Ameen!!!

I would like to express my heartfelt gratitude to my supervisors, Professors Karl D. Dearn and Duncan E.T. Shepherd from the Department of Mechanical Engineering at the University of Birmingham for their supervision and guidance throughout the entirety of my doctoral studies. I am eternally grateful to them for making this PhD journey an exciting and rewarding one and for their constant encouragement, which has made me a better researcher.

I would also like to thank my colleagues, Dr David G. Eckold and Dr Iestyn Stead, for their academic support, without whom this thesis would not be what it is today.

My thanks go to the School of Engineering, the University of Birmingham, for funding my PhD studies with the School Scholarship making my studies less stressful and allowing me to advance my research further.

The support of the Biomedical Engineering Research Group members and the Mechanical, Innovation and Tribology Group members has also been very valuable to completing of my thesis. A big thanks to Diana, Weiqi, Joe, Sanj, Nan, Zainab and everyone else for their companionship, support, advice, and ideas and most importantly, for providing me with necessary distractions throughout my studies.

I would like to thank my friends and extended family members who have provided me with some much-needed distractions throughout my studies.

I would also like to thank my husband for being a huge source of support and encouragement over the past year and a half. Writing this thesis has been a time-consuming task for the past year and a half, and I thank you for your patience and understanding.

Lastly, I would like to thank my parents (Engr & Mrs Olorunlambe) and my brothers (Abubakr & Ali Olorunlambe), who have been a huge source of support (both financial and emotional), inspiration and huge determination throughout the duration of my education. I am grateful for their unwavering support and encouragement, particularly at difficult moments, without which it would have been difficult to complete this thesis.

KAO 07/2022

Contents

1 Introduction **1**

1.2 Thesis Overview ……...4

2 Literature Review **6**

2.1 Introduction ……..7

2.2 Acoustic Emission ……....7

2.3 Bio-Tribology ……15

2.4 Tribological Applications of Acoustic Emission ……...18

2.5 Bio-tribo-acoustic study – Aims and Objectives ……28

3 General Materials and Methods **29**

3.1 Introduction ……30

3.2 Materials ……..30

3.3 Methods ……...30

Chapter Summary ……..39

4 Condition Monitoring of a Simulated Joint Articulation **40**

4.1 Introduction ……41

4.2 Materials and Methods ……41

4.3 Results and Discussion ……45

4.4 Conclusion ……...62

5 Diagnostic Capabilities of AE Testing **64**

5.1 Introduction ……65

5.2 Materials and Methods ……65

5.3 Results and Discussion ……72

5.4 Conclusion ……...88

6 Deep Transfer Learning of AE Signals **89**

6.1 Introduction ……90

6.2 Materials and Methods ……90

6.3 Results and Discussion ……..100

6.4 Conclusion ……118

7 Overall Conclusions **119**

7.1 Introduction ……120

7.2 Overall Conclusions ……120

7.3 Potential Further Work ……122

Appendices **125**

Appendix A1 Test Geometry Determination ……126

Appendix A2 MATLAB Codes ……127

Bibliography **133**

List of Figures

Figure 1.1: An overview of this thesis 5
Figure 2.1: Illustration of the AE test system showing progression of signal from generation to acquisition and storage. A sample signal, with basic parameters labelled, is also shown. 8
Figure 2.2: The three modes of AE waveforms. The red dash line indicates the amplitude threshold. 9
Figure 2.3: Duration against Amplitude Scatter Plot 12
Figure 2.4: Typical frequency and power spectral representations of a raw AE waveform 13
Figure 2.5: Typical CWT and STFFT transforms of a raw AE waveform 14
Figure 2.6: Illustration showing the progression from (a) Normal lumbar spine to (b) Lumbar spine with herniated disc L4-L5(1) and (c) Lumbar spine where the herniated disc has been replaced with an artificial one (2). 16
Figure 2.7: Illustration showing the progression from (a) Normal functioning Knee joint to (b) Osteoarthritic Knee (1) and to (c) where the affected knee joint has now been replaced with an artificial one (2) 17
Figure 2.8: Illustration of the Joint Acoustic Analysis System 22
Figure 2.9: Illustration showing acquisition of AE signal from the hip implant 24
Figure 3.1:The TE77 test configurations used in this thesis. (a) The face contact was used for tests in chapters 4 & 6; and (b) The line contact was used for tests in chapter 5. 31
Figure 3.2: Experimental layout of the TE77 tribometer and the AE acquisition system. (1) AE sensor mounted on (2) polymeric upper specimen, (3) flat lower specimen, (4) lubricant bath, (5) lubricant, (6) heater block. 32
Figure 3.3: Pictures of the 2/4/6 Preamplifier (Left) and the Nano-30 AE Sensor (Right) 33
Figure 3.4: Image of a Lumbar Spinal Implant 36
Figure 4.1: Raw AE signal at (a) 37 dB and (b) 41 dB amplitudes. The 37 dB signal is noisier than the 41 dB signal 43
Figure 4.2: NARX Neural Network. Obtained from MATLAB. 45
Figure 4.3: RMS (blue line) & CoF (black line) plots for (a) PEEK at 2 Hz, (b) UHMWPE at 2 Hz, (c) PEEK at 4 Hz and (d) UHMWPE at 4 Hz. Red dash line indicates transition from running-in (stage I, after lowest CoF is reached) to prolonged sliding. 46
Figure 4.4:Plots of average signal level (ASL) and absolute energy (AbsE) in relation to coefficient of friction (CoF) 47

Figure 4.5: NARX Net Training Performance for PEEK (left) and UHMWPE (right) 48
Figure 4.6: NARX Net Training Response for PEEK (top) and UHMWPE (bottom).....49
Figure 4.7: Test Response for PEEK Tests at 2 Hz (top) and 4 Hz (bottom).....50
Figure 4.8: Test Response for UHMWPE Tests at 2 Hz (top) and 4 Hz (bottom) 51
Figure 4.9: Silhouette plot showing silhouette values of hits in each cluster 54
Figure 4.10: Cluster Assignments.....54
Figure 4.11: Distribution of AE signals across all test stages for (a) PEEK at 2 Hz, (b) PEEK at 4 Hz, (c) UHMWPE at 2 Hz and (d) UHMWPE at 4 Hz. The generation of more continuous emissions during prolonged sliding (stage II) is common to all tests.....58
Figure 4.12: Polymeric specimens wear scar images for (a) PEEK at 2 Hz, (b) UHMWPE at 2 Hz, (c) PEEK at 4 Hz and (d) UHMWPE at 4 Hz. The regions of high z value indicate adhesion of displaced wear particles back onto the surface.....60
Figure 4.13: Peak Frequency Distribution Plots for (a) PEEK at 2 Hz, (b) PEEK at 4 Hz, (c) UHMWPE at 2 Hz and (d) UHMWPE at 4 Hz. Peak frequencies of the AE hits are mostly in the range 150 to 275 kHz. 61
Figure 5.1: Illustration of the KNN Classifier 71
Figure 5.2: Illustration of a three-layer backpropagation neural network. The hidden layer has 10 neurons. 72
Figure 5.3:Cumulative AE hits vs time plot for adhesive (red) and abrasive (black) wear tests 73
Figure 5.4:Coefficient of friction vs time curves for adhesive (red) and abrasive (black) wear tests 74
Figure 5.5: Adhesive test specimens, contour image & corresponding graphical profiles of worn region after testing for (a) steel plate & (b) PEEK rods. Regions of high height shows wear particle transfer from PEEK unto steel plate indicating adhesion. 75
Figure 5.6: Abrasive test specimens, contour image & corresponding graphical profiles of the worn regions after testing for (a) steel plate & (b) PEEK rods. Low topographical height shows material has broken away indicating abrasive wear mechanism. 76
Figure 5.7: Clustering solutions for adhesive wear test showing (a) duration vs amplitude and (b) hits vs amplitude plots.....77
Figure 5.8: Waveforms of hits from adhesive wear test clusters in time domain (top) and frequency domain (bottom). Cluster 1 event is a burst emission while cluster 2 event is a continuous emission.....78

Figure 5.9: Clustering solutions for abrasive wear showing (a) duration vs amplitude and (b) hits vs amplitude plots. 80
Figure 5.10: Waveform of hits from abrasive wear test clusters in time domain (top) and frequency domain (bottom). Event in cluster 1 is a burst emission while that of cluster 2 is a mixture of continuous and burst emissions.....82
Figure 5.11: Principal Component Analysis Output. The principal components that accounted for 95% variance were selected for training. 84
Figure 5.12:Confusion matrix for all three classifiers 87
Figure 6.1: Illustration of classification of AE signals using a pre-trained CNN and CWT images 93
Figure 6.2: GoogLeNet Framework. The expanded view shows the layers that are finetuned before retraining on new dataset (Image recreated from MATLAB). 98
Figure 6.3: Illustration of the Support Vector Machine.....99
Figure 6.4:Image field capture of UHMWPE wear scar for burnishing tests. Region labelled A shows adhesion of displaced wear particles back unto the UHMWPE surface.....102
Figure 6.5:Image field capture of UHMWPE wear scar for scratching tests. There are grooves on the surface caused by the scratching of the grinding grits on the surface during sliding. 103
Figure 6.6: Image field capture of the CoCrMo counterface post-scratching tests. The grooves indicate scratching wear having taken place.....103
Figure 6.7: Mean CoF and Mean AE RMS plots for burnishing test 106
Figure 6.8: Mean CoF and Mean AE RMS for scratching tests 106
Figure 6.9: Mean CoF and Mean AE RMS plots for wear mode progression tests 107
Figure 6.10: Box Plots of Average AE Features showing result of ANOVA test.....109
Figure 6.11: Denoised waveform (top) and the corresponding CWT Image (bottom) of a sample signal for Burnishing Wear tests.....110
Figure 6.12:Denoised waveform (top) and the corresponding CWT Image (bottom) of a sample signal for Scratching Wear tests. 111
Figure 6.13: CNN Transfer Learning (based on GoogLeNet) Training Progress. The top image is the accuracy plot and the bottom image is the corresponding loss plot. 112
Figure 6.14:Minimum classification error output for CNN-SVM using un-trained GoogLeNet features.....113
Figure 6.15: Minimum Classification Error Plot for CNN-SVM using trained GoogLeNet features. 113
Figure 6.16: GoogLeNet Transfer Learning Test Confusion Matrix.....115

Figure 6.17: Test confusion matrix for CNN-SVM using un-trained GoogLeNet features...116
Figure 6.18: Test confusion matrix for CNN-SVM using trained GoogLeNet features 116

List of Tables

Table 1.1:Joint replacement surgeries carried out in 2020. Data obtained from the National Joint Registry 18th Annual Report 2021 2
Table 1.2: Knee and Hip revision procedures (and the reason for revision) in the year 2020. Data collated from the National Joint Registry 18th Annual Report 2021. 3
Table 1.3: Total disc replacement failure modes and their corresponding causes. Data collated from (Reeks and Liang 2015). 3
Table 2.1: Modes of wear detected by acoustic emission testing (Wang, Wood, and Sun 2008; Price, Lees, and Friswell 2005; Mukhopadhyay et al. 2012; Rodgers et al. 2017). 10
Table 2.2: Overview of selected biotribological applications of AE 27
Table 3.1: Materials used in alphabetical order. Properties were either provided by the suppliers or from other public sources (Geetha et al. 2009; Xin, Shepherd, and Dearn 2012). 30
Table 3.2: AE Timing Parameters and their descriptions 34
Table 3.3: Selected AE features and their definitions in alphabetical order. 34
Table 3.4:Maximum contact pressure a ball-and-socket Charite Lumbar Spinal Implant is subjected to. Calculated using Hertzian contact mechanics. 36
Table 4.1: Summary of Test Parameters 43
Table 4.2: NARX Neural Network Training Result Summary 48
Table 4.3: A summary of CoF prediction results with test data 48
Table 4.4: Mean and 95% confidence intervals (C.I.) of AE features in each cluster for PEEK at 2 Hz 55
Table 4.5: Mean and 95% confidence intervals (C.I.) of AE features in each cluster for PEEK at 4 Hz 56
Table 4.6: Mean and 95% confidence intervals (C.I.) of AE features in each cluster for UHMWPE at 2 Hz 56
Table 4.7: Mean and 95% confidence intervals (C.I.) of AE features in each cluster for UHMWPE at 4 Hz 56
Table 5.1: TE77 and AE Acquisition Parameters 66
Table 5.2: Selected AE features after hierarchical clustering. Definitions can be found in Section 3.2.1 67
Table 5.3: Range and mean (including standard deviation) of five features per cluster for adhesive wear tests 77

Table 5.4: Features of the sample AE waveforms (Figure 5.8) from adhesive wear tests clusters. 78
Table 5.5: Range and mean (including standard deviation) values of five features per cluster for the abrasive wear test 81
Table 5.6: Features of the sample AE waveforms (Figure 5.10) from abrasive wear tests clusters. 82
Table 5.7: Extract of the data for supervised learning after PCA. Column seven is the known classification. 0 is for adhesive wear and 1 is for abrasive wear. 84
Table 5.8: Summary of Classifiers' Performance 85
Table 6.1: TE 77 and AE Acquisition Parameters 92
Table 6.2: The signal-to-noise ratio and mean square error of different wavelets. The wavelet with the best denoising effect is shown in the red font. 95
Table 6.3: Transfer Learning Training Options 98
Table 6.4: Average Coefficient of Friction and Frictional Work Done 101
Table 6.5: Summary of ANOVA test results. StDev is standard deviation and CI is confidence interval. 108
Table 6.6: Training and Test accuracy of all three CNN models. 114
Table A1.1: Load and Contact Mechanics for a TE77 sphere-on-plane configuration 126

List of Equations

Equation 3.1: Calculation for equivalent elastic modulus, E^*, where E_1 & E_2 represent the elastic moduli of the socket and ball materials, respectively, and $\nu 1$ & $\nu 2$ are the Poisson's ratios of the materials. 35
Equation 3.2: Calculation for equivalent radius, R, where R_1 & R_2 represent the radii of the socket and balls materials, respectively. 35
Equation 3.3: Calculation for contact area radius, a, where P is the normal applied load. 35
Equation 3.4: Calculation for maximum contact pressure, $p0$ 35
Equation 4.1: Calculation for equivalent TE77 load, P*, in a face contact where a* is the nominal contact area. 42
Equation 4.2: NARX model equation 44
Equation 5.1: TE77 Load calculation for a line contact where l is length of the cylindrical rod. 66
Equation 5.2: Equation for normalising the discrete AE features. μ represents the mean value of the descriptor and σ the standard deviation. 67
Equation 5.3: Matrix, X, of normalised AE features. 68
Equation 5.4: Covariance of Matrix X 68
Equation 5.5: Equation for calculating covariance 68
Equation 5.6: New matrix, Y. 69
Equation 5.7: Formula for precision. 69
Equation 5.8: Formula for recall. 69
Equation 5.9: Formula for f-score. 69
Equation 5.10: Sigmoid function 70
Equation 5.11: z index calculation 70
Equation 5.12: Logistic regression model 70
Equation 6.1: Formula for computing SNR where fn is the original signal, $f(n)$ is the denoised signal and N is the length of the signal. 95
Equation 6.2: Formula for computing MSE where f(n) is the original signal, f ̂(n) is the denoised signal and N is the length of the signal. 95
Equation 6.3: CWT Integral where x(t) is the original signal, $\psi(t)$ is the analysing function (wavelet), a is a scale parameter and b is a position in time. 96
Equation 6.4: Mathematical representation of the Morlet wavelet 96

Acronyms

ABR Abrasive

AbSE Absolute Energy

ADH Adhesive

AE Acoustic Emission

ANN Artificial Neural Network

ANOVA Analysis of Variance

ASL Average Signal Level

BP Back Propagation

BUR Burnishing

CNN Convolutional Neural Network

CoCrMo Cobalt Chromium Molybdenum

CoF Coefficient of Friction

CONV Convolutional

CT Computed Tomography

CWT Continuous Wavelet Transform

FC Fully connected

FFT Fast Fourier Transform

FWd Frictional work done

GPU Graphics Processing Unit

HDT Hit Definition Time

HLT Hit Lockout Time

KNN K-Nearest Neighbours

MSE Mean Squared Error

NARX Nonlinear Autoregressive with external inputs

NDT Non-Destructive Test

PCA Principal Component Analysis

PDT Peak Definition Time

PE Polyethylene

PEEK Poly-ether-ether-ketone

POOL Pooling

R Regression

RA Ratio of risetime to amplitude

ReLU Rectified Linear Unit

RMS Root mean square

SCR Scratching

SNR Signal-to-noise ratio

STFFT Short Time Fast Fourier Transform

SVM Support Vector Machine

TDD Time dependent

UHMWPE Ultra high molecular weight Polyethylene

Chapter 1

Introduction

1.1 Introduction

More people are suffering from joint pathologies such as acute trauma and osteoarthritis, necessitating the need for joint replacement procedures. In 2020 alone, over 110,00 joint replacement procedures were carried out in the UK (Table 1.1) (National Joint Registry 2021). Tribological interactions are fundamental to the operation of these artificial joints, with wear and other wear-related factors such as aseptic loosening, subsidence and periprosthetic fracture emerging as the principal means of failure in these devices (Table 1.2 and Table 1.3) (Abu-Amer, Darwech, and Clohisy 2007; Punt et al. 2008; Hossain, Patel, and Haddad 2010; Dowson 2012; Reeks and Liang 2015). As shown in Table 1.2, the migration of wear debris into the blood system can also cause adverse soft tissue reactions leading to the need for revision surgery (Eckold, Dearn, and Shepherd 2015).

Table 1.1:Joint replacement surgeries carried out in 2020. Data obtained from the National Joint Registry 18th Annual Report 2021

Reference Joint	No of replacement procedures	Selected reasons for replacement
Hips	54,858	Acute trauma – 7 % Osteoarthritis – 88%
Knees	50,904	Osteoarthrtis – 97% Unicondylar knee replacements – 13%
Ankles	465	Osteoarthritis – 92% Rheumatoid arthritis and other inflammatory joint problems – 8 %
Elbows	561	Total elbow replacement – 38% Radial head replacements – 45% Distal humeral hemiarthroplasty – 14%
Shoulders	3,833	Acute trauma – 17% Osteoarthritis – 59% Effective cuff tear arthroplasty – 27%

Table 1.2: Knee and Hip revision procedures (and the reason for revision) in the year 2020. Data collated from the National Joint Registry 18th Annual Report 2021.

Reason for revision surgery	**Hip revision surgeries – Total is 4,910**		**Knee revision surgeries – Total is 3,907**	
	No	Percentage, %	No	Percentage, %
Aseptic Loosening	1794	35	1359	34
Infection	1066	21	1262	31
Periprosthetic fracture	1120	22	285	7
Dislocation/Subluxation	932	18	157	4
Implant Fracture	228	5	56	1
Lysis	660	13	449	11
Wear	495	10	402	10
Adverse soft tissue reaction to particulate debris	390	8	N/A	N/A
Malalignment	205	4	162	4
Unexplained Pain	129	3	213	5
Component dissociation	91	2	76	2
Instability	N/A	N/A	553	14
Stiffness	N/A	N/A	160	4
Progressive arthritis remaining knee	N/A	N/A	619	15

Table 1.3: Total disc replacement failure modes and their corresponding causes. Data collated from (Reeks and Liang 2015).

Failure Mode	**Cause of Failure**
Degradation	Wear and corrosion
Inflammation	Presence of large UHMWPE wear particles
Surface wear and damage	Adhesive and abrasive wear
Damaged UHMWPE core	Plastic deformation and fracture of the rim
Corrosion	Fretting wear
Osteolysis	Micro-motion of implant and presence of wear debris

Artificial joint failures are traditionally diagnosed using X-rays, computed tomography (CT) scans and recently nuclear medicine bone scans, but these are expensive, time-consuming and harmful to health due to frequent radiation exposure (Devin, Myers, and Kang 2008; Karl

2013; David 2007; Lu et al. 2015; Leuridan et al. 2017; Malham and Parker 2017; Lee et al. 2021). There is also a concern with these traditional diagnostic methods that clear signs of failure do not present early enough to prevent pathologies, causing patients to experience pain and allowing the migration of wear debris into the bloodstream leading to further medical complications (Abu-Amer, Darwech, and Clohisy 2007; Karl 2013; Brown et al. 2011).

Previously, researchers have studied the tribological phenomena of artificial joint materials during the implant design process to assess its suitability for use in the body and for long term wear predictions (Moghadas et al. 2012; Baykal et al. 2014; Alnaimat, Shepherd, and Dearn 2016; Jin et al. 2016; Siskey et al. 2016; Saikko 2017). Although these have successfully predicted wear of artificial joints and identified which biomaterials are most appropriate, there are some limitations. One such limitation is that these traditional biotribological studies can only be used to evaluate the cause of failure of artificial joints after retrieval from the body. This still does not solve the problem of monitoring the condition of artificial joints *in vivo.* Hence, there is a need for a simplified, dynamic and faster way of non-invasively monitoring the condition of artificial joints with the added advantage of early diagnosis of failed artificial joints. This is where applying acoustic emission (AE) testing has potential.

1.2 Thesis Overview

The focus of this thesis is to use Acoustic Emission (AE) testing to diagnose tribological phenomena in artificial joint materials using two approaches. First, a bio-tribo acoustic test method was developed to monitor tribological phenomena of artificial joint materials, and second, traditional machine learning and deep learning techniques are used to analyse AE signals for damage mechanism identification. The thesis overview is presented in Figure 1.1.

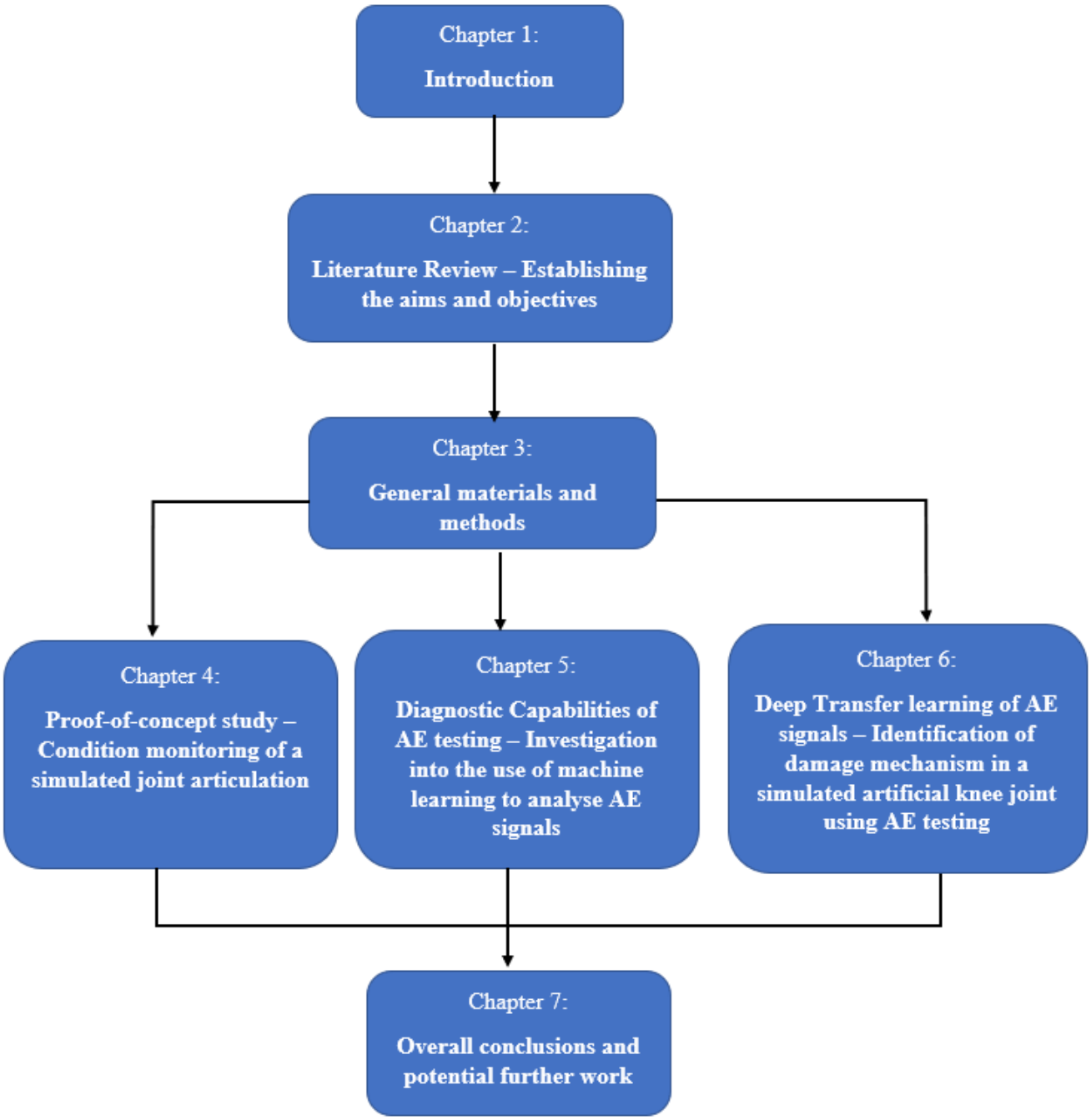


Figure 1.1: An overview of this thesis

Chapter 2

Literature Review

This chapter is based on the following publication:

K. A. Olorunlambe, D. E. T. Shepherd & K. D. Dearn (2019). “A review of acoustic emission as a biotribological diagnostic tool.” *Tribology - Materials, Surfaces & Interfaces* 13 (3): 161–71. https://doi.org/10.1080/17515831.2019.1622914.

2.1 Introduction

This chapter presents a review of previous research relevant to this thesis. Section 2.2 introduces the concept of the acoustic emission (AE) testing system, its characteristics and the techniques used to analyse the waveforms. An overview of the tribology of artificial joints is presented in section 2.3 and section 2.4 details the tribological applications of AE. Last but not least, section 2.5 details the synergy of AE testing and biotribology and presents the research aims and objectives.

2.2 Acoustic Emission

Acoustic emission (AE) testing is a non-destructive test (NDT) method used to detect and locate faults in mechanically loaded structures and components (Hellier 2003). Joseph Kaiser first developed the AE testing technique in the 1950s whilst he was researching the sounds metals emit upon mechanical stress (Tensi 2004). It involves the use of piezoelectric sensors that convert mechanical movement to an electrical voltage signal which is then pre-amplified before being detected and recorded via a data acquisition unit (Figure 2.1). The signal is then analysed and characterised based on the source location, voltage intensity and frequency content (Aggelis et al. 2015). A sample raw waveform along with the basic parameters is shown in Figure 2.1. Unlike other NDT methods, AE testing detects energy release that is initiated within the material due to applied loads. AE has become widely used to provide information on both the origin of a flaw and its progression due to repetitive stress (Hellier 2003).

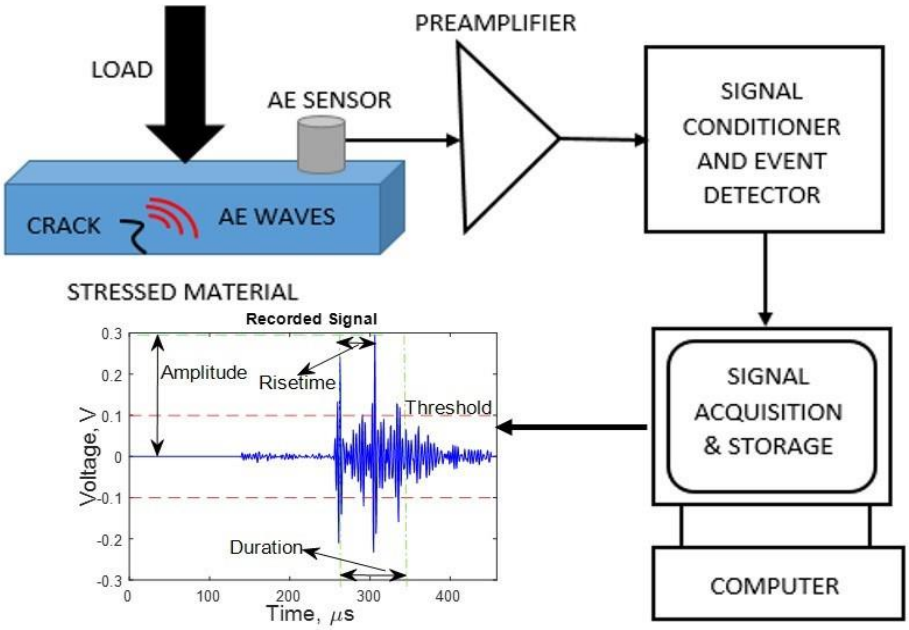


Figure 2.1: Illustration of the AE test system showing the progression of signal from generation to acquisition and storage. A sample signal, with basic parameters labelled, is also shown.

Acoustic emission waveforms are generally categorised into three modes (Figure 2.2) – continuous, burst and mixed emission (Unnporsson 2013). Continuous emissions are those with short amplitudes and long duration. They are most commonly generated by plastic deformation, sliding and rubbing (Rao 1990; Unnporsson 2013). Conversely, burst emissions are known to have high amplitudes and short duration. They are generated due to damage formation such as asperities breakage, crack propagation and delamination (Rao 1990; Unnporsson 2013). Mixed emissions are waveforms consisting of both continuous and burst emissions. This is the most common mode of emission observed as damage to most structures and components tends to consist of two or more damage/failure processes.

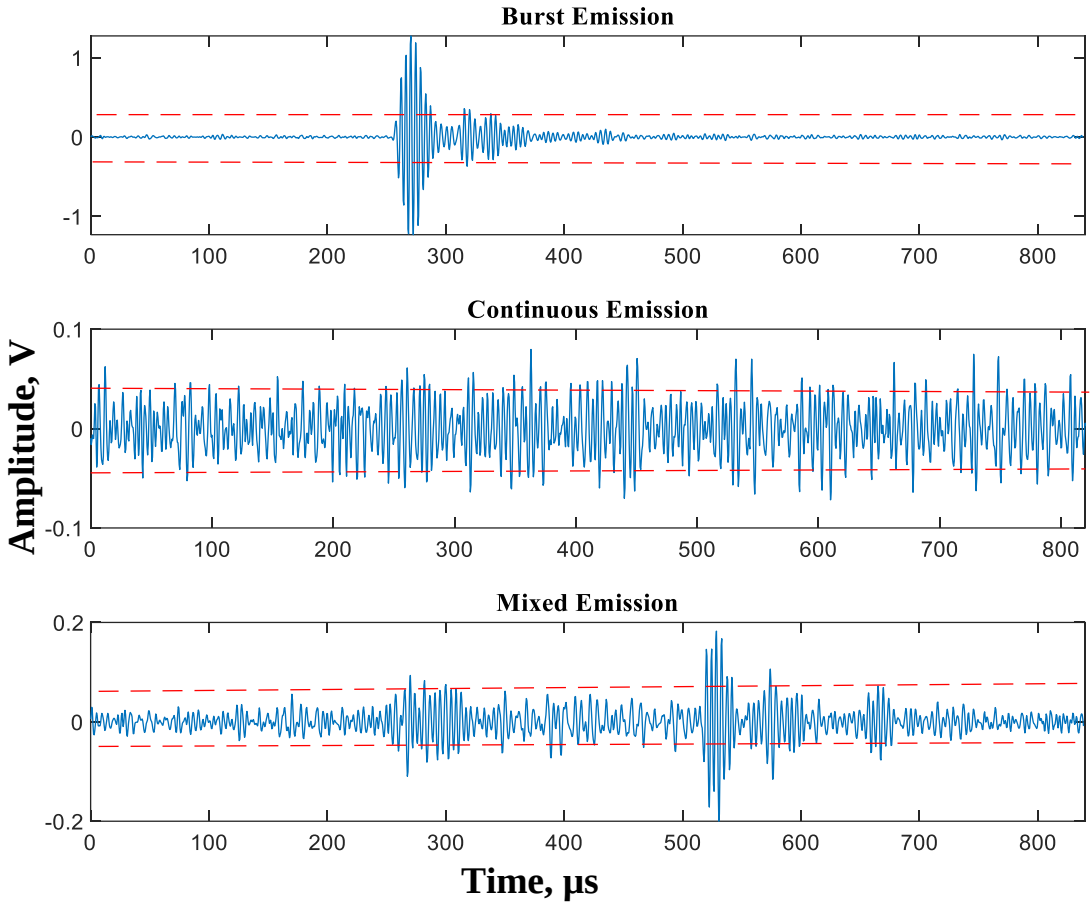


Figure 2.2: The three modes of AE waveforms. The red dash line indicates the amplitude threshold.

AE testing has been utilised in all engineering applications in one form or another. Some of its earliest applications were in the study of plastic deformation, crack propagation in materials and the detection and location of flaws in metallic and composite structures (Rao 1990). Its application has since evolved, and it is now being used as a tool for monitoring the conditions of structures such as cables, condition monitoring of tribological characteristics of mechanical systems and the study of the tribology of artificial and natural human joints (Li et al. 2012; Olorunlambe, Shepherd, and Dearn 2019).

As far back as 1980, there have been various studies on how AE can be used to analyse the tribological characteristics of surfaces in contact. These studies have been successfully used for mechanical condition monitoring like bearings in motors, pumps and transport systems

(Rowland et al. 2004) and tool wear experienced during machining (Wang, Wood, and Sun 2008). Many of these studies have been focused on the use of AE to measure wear, particularly adhesive and abrasive wear. Adhesive wear occurs when the atomic bonds between two contacting surfaces become stronger than the subsurface strength causing material removal/displacement whilst abrasive wear occurs when a hard material slides over a soft material causing material removal and/or when a third body is introduced to the bearing surface causing third body abrasive wear. Table 2.1 shows other modes of wear that are detectable by AE.

Table 2.1: Modes of wear detected by acoustic emission testing (Wang, Wood, and Sun 2008; Price, Lees, and Friswell 2005; Mukhopadhyay et al. 2012; Rodgers et al. 2017).

Application	Modes of wear	AE Features
Transmission Systems	Scuffing, pitting fatigue and bearing wear	High signal-to-noise ratio; scuffing wear had higher RMS amplitude than pitting
Machine tool monitoring	Flank, abrasive and crater wear	There is a linear relationship between tool wear and AE parameters
Orthopaedic applications	Adhesive, abrasive and fatigue wear	Implant AE is generally of the continuous type

2.2.1 Attenuation Characteristic of AE Signals

Attenuation is the phenomenon whereby the amplitude of acoustic waves decreases as they travel through a structure, mostly due to energy absorption, geometrical spreading and structural scattering (Hellier 2003). Geometrical spreading is one of the leading causes of attenuation as it relates to the geometry of the structure under observation. In compact structures like pipes, the narrowness of the space through which the wave propagates will reduce its dispersion, whereas in a solid object like concrete, there will be a dispersion of waves in all three dimensions leading to a rapid reduction of the wave amplitude as it propagates from the signal source to the sensor. Structural scattering happens when the wave encounters

discontinuities, for example, having to change direction, causing the wave to reflect; hence, attenuation of the signal occurs. The last cause is energy absorption which happens when kinetic and elastic wave energies are absorbed and turned to heat by the material. The type of material determines how much energy is absorbed; for example, non-metallic materials absorb more energy than metallic ones. It is important to consider attenuation effects during the monitoring of structures to ensure that all factors that could affect the propagation of the waves are considered when interpreting AE results.

2.2.2 Analysis of AE Waveforms

The analyses of AE waveforms are traditionally performed in one of two ways: parameter-based and frequency-based analysis. The parameter-based analysis involves evaluating the waveforms generated using parameters such as amplitude, count rate, number of hits, peak frequency and AE r.m.s (root mean square) value. Any trends in these parameters are observed and quantified using distribution plots such as scatter plots (Figure 2.3), histograms and line graphs. One of the early applications of AE in tribology by Belyi et al. (Belyi, Kholodilov, and Sviridyonok 1981) used this technique to investigate the possible relationship between AE parameters and mode of wear for polymers. Jiaa and Dornfeld (Jiaa and Dornfeld 1990) also analysed the AE waves generated during sliding tests and found the parametric-based method to be effective in showing the progression of wear from running-in to steady-state and beyond.

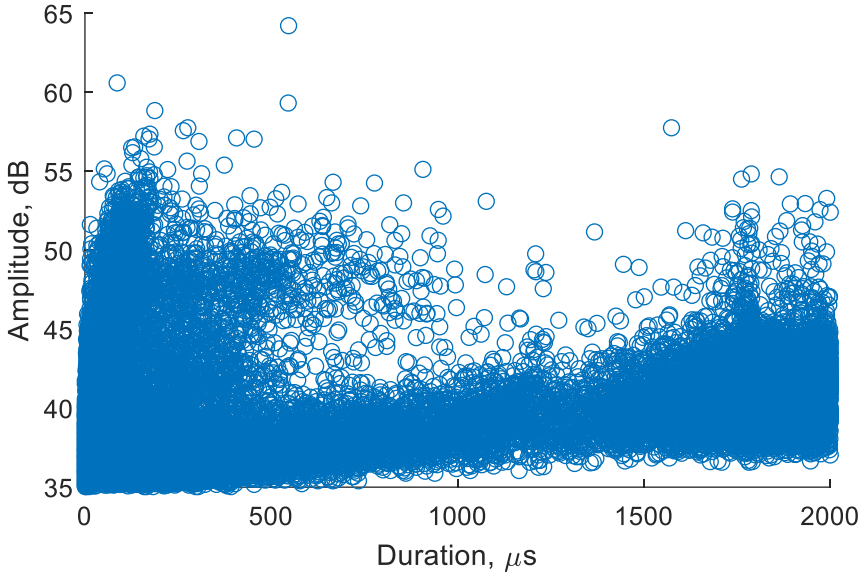
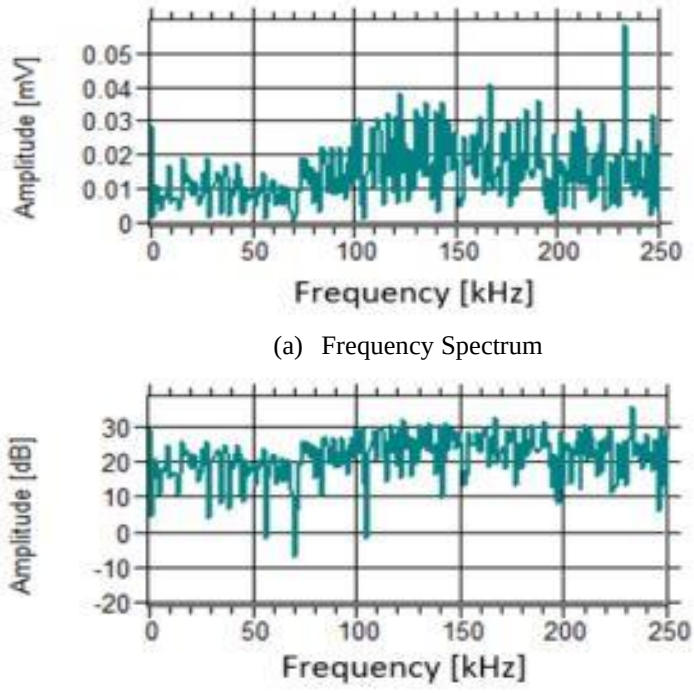


Figure 2.3: Duration against Amplitude Scatter Plot

Despite the effectiveness of the technique, it gives limited information and there is a need for the frequency-based investigations as it allows for the significant frequencies of the signal to be identified. This can aid in determination of tribological properties related to the signals. The importance of frequency-based analysis was illustrated in the work by Lingard et al. (Lingard, Yu, and Yau 1993). By analysing the frequency characteristics of the AE signals, they were able to identify the peak frequencies and found that the wearing materials had different peak frequencies suggesting that AE could be used to distinguish between the different materials. Zykova et al. (Zykova, Mazal, and Pazdera 2006) also used frequency analysis to distinguish between running-in and pitting stages during fatigue tests of grey cast iron and carbon steel samples showing how useful frequency-based investigation can be for tribological investigations that involve the use of acoustic emission test system.

Frequency-based investigations can take different forms: Continuous wavelet transform (CWT), power spectral density, frequency spectrum and short-time fast Fourier transform (STFFT). Their function is to extract hidden features of the waves usually in terms of the wave frequency. Examples of power spectral density and frequency spectral outputs from an AE signal are shown in Figure 2.4. The STFFT and CWT transform produce an image analysis of

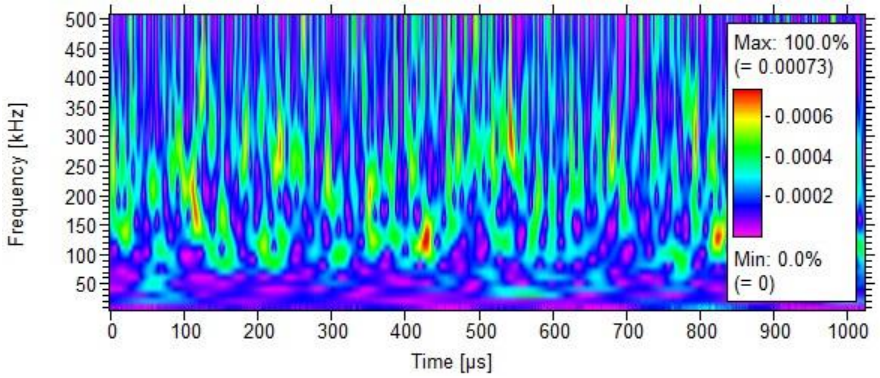
the AE waveform in terms of frequency against time with a colour map showing the intensity of the different frequencies present in the signal. STFFT involves spectral analysis of the waves generated in small segments called wavelets (i.e., short waves) and averaging the results to reveal the hidden features of the generated wave. One of the limitations of FFT is its inability to localise the signal in the time domain, which the STFFT attempts to correct, but, as seen in Figure 2.5, it is difficult to extract the exact time at which each significant frequency component arrived at the sensor. Unlike the STFFT, wavelet transforms such as the CWT are able to show the actual time intervals for the different frequency components with no overlapping (Polat and Siraç 2018). This is useful because, despite the frequency components of the wave being generated at the source simultaneously, they arrive at the sensor at different times (Asamene and Sundaresan 2012).



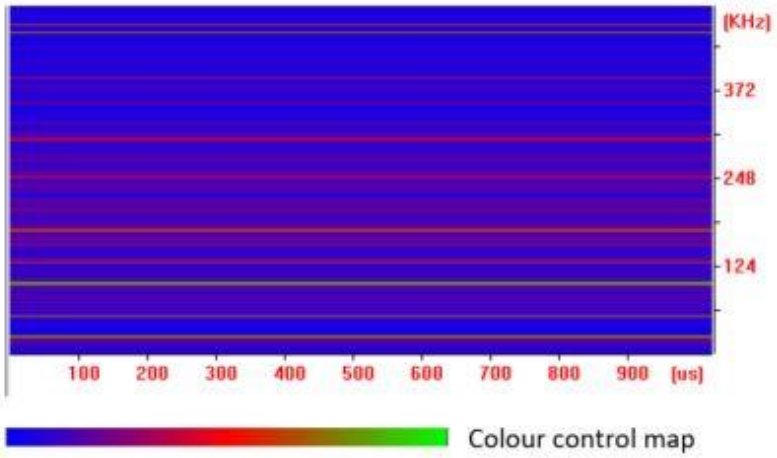
(a) Frequency Spectrum

(b) Power Spectral Density

Figure 2.4: Typical frequency and power spectral representations of a raw AE waveform



(a) Continuous Wavelet Transform (CWT)



(b) Short-time Fast Fourier Transform

Figure 2.5: Typical CWT and STFFT transforms of a raw AE waveform

2.2.3 Application of Machine Learning in AE Signal Analysis

The use of machine learning techniques to analyse and interpret data has become more widespread within the engineering industry due to its ability to find patterns within data that basic statistical analysis cannot provide. Applying some of these techniques to analysing AE data has been advantageous for achieving the diagnostic capabilities of AE testing.

Gutkin et al. (Gutkin et al. 2011) analysed AE signals from carbon fibre reinforced polymers under various test configurations using three pattern recognition algorithms: k-means

clustering, Self Organising Map (SOM-k) combined with k-means and Convolutional Neural Network (CNN). SOM-k means was the most effective at classifying AE responses to failures. Qiao et al. (Qiao, Weng, and Li 2019) also used k-means clustering to classify AE signals from ceramic thermal barrier coatings during indentation testing into three distinct categories associated with different failure modes. Using Back Propagation (BP) neural networks to further identify failure types of the coatings after thermal exposure, the results showed that AE measurements could distinguish between the mechanisms of high-temperature oxidation that accelerated thermal barrier degradation. Yao et al. (Yao, Li, and Yuan 1999) used a combination of the wavelet fuzzy neural network with AE and fuzzy classification combined with a motor current to estimate tool wear successfully. Machine learning approaches, combined with AE, enable a deeper categorisation of signals based on damage recognition and failure modes. Deep transfer learning, in combination with wavelet transforms, have also been used to classify AE signals from fracture processes and for wear monitoring in sliding bearing systems (Xin et al. 2020; Ren and Chen 2021; König et al. 2021). This involves using already trained convolutional neural networks such as GoogLeNet and VGG Net to classify AE signals that have been transformed into their representative scalograms using wavelet analysis. Through the use of deep transfer learning, the complex feature engineering process required in a traditional machine learning approach is eliminated, thereby saving time and reducing the computing power required for the analysis.

2.3 Bio-Tribology

Tribology is the study of friction, wear and lubrication whilst bio-tribology can be defined as the aspect of tribology related to biological systems (Jin et al. 2006). Examples of the application of tribology to biology include, but are not limited to, the tribology of natural and artificial joints, wear of dentures, and the friction of skin and garments. The focus of this thesis is the application of tribology to the bearing surfaces of artificial joints. The following

section provides an understanding of the concept of biotribology about the study of artificial joints.

2.3.1 Tribology of artificial joints

Joint diseases such as herniated discs, osteoarthritis and rheumatoid arthritis affect the functionality of a natural joint which often leads to a need for the affected joint to be replaced with an artificial one (Figure 2.6 and Figure 2.7) (Zhou and Jin 2015). The mechanics of natural joints mimic that of surfaces articulating against each other such as the multi-directional reciprocating motion of the hip and spine. The manufacturing of an artificial joint that can effectively replace a natural one would require an understanding of the tribological characteristics of a natural joint.

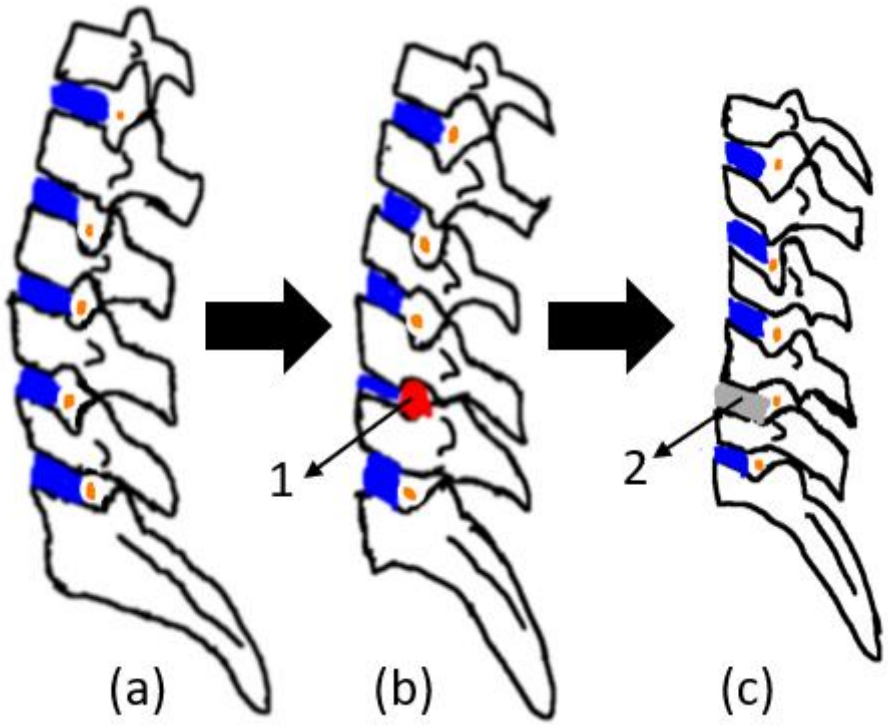


Figure 2.6: Illustration showing the progression from (a) Normal lumbar spine to (b) Lumbar spine with herniated disc L4-L5(1) and (c) Lumbar spine where the herniated disc has been replaced with an artificial one (2).

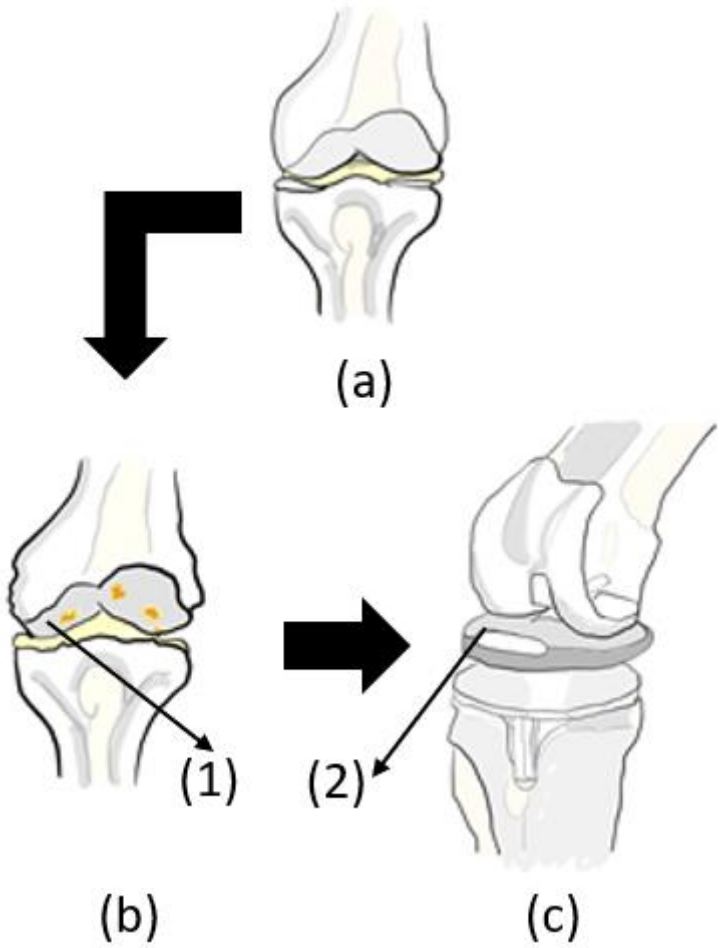


Figure 2.7: Illustration showing the progression from (a) Normal functioning Knee joint to (b) Osteoarthritic Knee (1) and to (c) where the affected knee joint has now been replaced with an artificial one (2).

Understanding the tribological characteristics of natural joints has several benefits such as choosing suitable biomaterials and for determining the required life of the artificial joint. Tribological testing has two aspects: bench testing using tribometers and long-term wear tests using joint simulators. The bench tests allow for *in vitro* wear and friction tests to be carried out at the material level first to build a profile of how the intended joint would behave *in vivo* before the manufactured implant is tested for durability using joint simulators. This would save time and money as unsuitable materials and/or designs would have been eliminated during bench testing, ensuring that only devices that have shown suitability for the intended use are tested on joint simulators.

Joint simulator tests are essential and necessary as they are used to decide if the device would meet the required conditions of use and therefore be allowed to go to market (Jin et al. 2016). These tests are carried out by fixing the device to the corresponding simulator (i.e., an artificial hip joint to a hip simulator and inter-vertebral (IVD) spinal discs to a spine simulator) and applying standardised test parameters to simulate physiological conditions and to determine the life cycle of the device. Wear measurements are taken periodically (~ every1 million cycles) until either the cycle limit is reached (10 million for IVDs and 5 million for hips) or a device failure occurs (BS ISO 18192-1 2011; BS ISO 14242-1 2014).

Material bench testing has been used in the characterisation of the wear behaviour of ultra-high-molecular-weight polyethylene (UHMWPE) using multidirectional pin-on-disk tests for an improved understanding of how different formulations of UHMWPE behave in biotribological applications (Baykal et al. 2014) and to assess the suitability of different lubricants (Alnaimat, Shepherd, and Dearn 2016) for better and more optimised lubricant selection. Joint simulation testing has also been used in the comparison of bearing surfaces for use in IVD devices to determine optimal materials (Moghadas et al. 2012) and to replicate impingement loading conditions of lumbar total disc replacements in *in vitro* bench tests (Siskey et al. 2016) and so to be included in pre-clinical tests. Biotribological tests have also proven useful in identifying the causes of failure of artificial joints as they are used to assess the condition of retrieved failed implants.

2.4 Tribological applications of AE

2.4.1 Mechanical Systems

AE is capable of identifying the different wear and friction mechanisms, such as adhesive and abrasive wear, that are active during tribological processes. During sliding three phases of wear that take place: running in, steady state and rapid wear. Studies by Belyi et al.

(Belyi, Kholodilov, and Sviridyonok 1981) and Jiaa and Dornfeld (Jiaa and Dornfeld 1990) showed that an AE signal can be used to distinguish between these. There was variation in the AE r.m.s and AE intensity values during running-in compared to during steady-state wherein the values were much more stabilised. It was also found that when the major wear mode developed from adhesive to abrasive, there was an increase in AE intensity (Belyi, Kholodilov, and Sviridyonok 1981), suggesting that AE can be used to distinguish between adhesive and abrasive wear modes. In addition to identifying wear mechanisms, the study by Wang et al. (Wang, Wood, and Sun 2008) identified a correlation between AE signals and wear mechanisms. The oil starvation strategy they employed allowed for the observation of this relationship. After turning off the oil supply, four distinct stages of decay were identified and there was a decrease in the duration of burst raw AE signal acquired from stage A (start of starvation) to C (onset of scuffing). At point D where scuffing was visible, the AE signal acquired was of the continuous type with no burst emission superimposed as with the three previous stages.

Studies have also shown that there is a relationship between AE signal energy and wear loss. By integrating the AE r.m.s value of the signal obtained, Boness et al (R. Boness, McBride, and Sobczyk 1990; R. J. Boness and McBride 1991) observed an empirical relationship between the r.m.s and the wear scar volume in the form of a power law expression. This expression was shown to be independent of lubricant used when it was used to estimate wear scar volume for lubricated tests, and there was an excellent agreement with experimental wear scar volumes. Similarly, Hase et al. (Hase, Mishina, and Wada 2013) observed a linear relationship between AE pulse energy(μV.s) and the amount of wear elements generated.

AE can also distinguish between different lubricating conditions, i.e. wet or dry as shown in the work by Lingard et al. (Lingard, Yu, and Yau 1993). They discovered that AE outputs for lubricated tests were lower than that of unlubricated tests. In addition, the analysis

of frequency characteristics of the waves showed that material differences were evident in the frequencies observed (aluminium alloy slider at 140 kHz, steels at 195 kHz and brass slider at 131 & 202 kHz).

Although these studies were able to successfully use parameter-based analysis to interpret the AE waves generated, more recent studies have shown the advantage of using the frequency-based technique. Asamene & Sundaresan (Asamene and Sundaresan 2012) used the wavelet method to analyse the AE waves generated during wear and friction tests (reciprocating and oscillatory motions) of metallic specimens, and they were able to distinguish between friction-related AE and that of fatigue crack growth. CWT of friction-related AE waves exhibited more regions with AE impulses at different frequencies and at different times, indicating the continuous nature of the waves, while crack growth only had one significant frequency content at the start indicating the burst nature of the AE waves generated. This shows that CWT can be useful for the identification of the mode of AE generated, be it burst, continuous or a combination of both.

Studies by Baccar & Soffker (Baccar and Söffker 2015; Baccar and Soffker 2013) have shown the advantage of using the frequency-based technique in addition to parameter-based analysis. In the first study, they used STFT to represent raw AE waves captured during wear examination of a test rig and found that although parameter-based analysis of the AE energy indicated when each wear phase occurred, the same could not be said for the STFT results. When compared, STFT during the run-in phase looked similar to the wear-out phase, which did not allow for the identification of the two phases. They also found that the quantitative analysis of the frequency content of each phase was not possible with the STFT analysis (Baccar and Soffker 2013). In a further study by the same authors, they found that combining the STFT with CWT allowed them to distinguish between the wear phases and also identify the frequency components of each phase which was not possible in the previous study (Baccar

and Söffker 2015). This illustrates that although using the frequency-based technique provides more information than parameter-based analysis, a combination of the two techniques gives a more effective extraction of information from the raw AE waveforms. Other studies using frequency-based analysis include the use of the Window Fourier Transform technique to differentiate between plastic deformation and brittle fracture (Rubtsov et al. 2013); using frequency spectra to distinguish between abrasive and adhesive wear (frequency peak was at 1.1 MHz during adhesive wear and between 0.25 to 1 MHz during abrasive wear) (Hase, Mishina, and Wada 2012); the study of the relationship between sound frequency spectrum and polymeric gear materials (Hoskins et al. 2011) and the use of spectral analysis to monitor fretting wear in a poly-methyl-meth-acrylate/steel contact where the acoustic energy was found to be dependent on the contact load and sliding speed (Briscoe et al. 2001).

2.4.2 Tribology of Human Joints

Schwalbe et al. (Schwalbe, Bamfaste, and Franke 1999) and Franke et al. (Franke et al. 2004) used AE to study the *in vivo* friction and wear of knee joints. In both studies, a piezoelectric transducer was fixed directly to the skin over the medial femoral condyle of the knee joints of patients and measurements were taken while they carried out a variety of daily tasks, for example, climbing stairs. The AE waves from both studies were of the continuous type (continuous emission has been shown to represent friction-related event, (Asamene and Sundaresan 2012)), confirming the presence of friction processes during knee bending (Schwalbe, Bamfaste, and Franke 1999; Franke et al. 2004). Schwalbe et al. extended their study to investigate the probability of predicting bone fracture by applying combined bending and torsional loads to explanted human femur bones. The transducer was attached directly to the bone to collect and analyse AE waves generated. The AE signal was observed to be of two types: burst emission followed by continuous emission. The burst type indicated crack

initiation whilst the continuous type showed the relaxation phase following cracking. Agcaoglu & Akkus (Agcaoglu and Akkus 2013) conducted a similar study whereby they investigated the fatigue life of human tibia cortical bone by applying three-point bending loads to the bone with two R15-α sensors from Physical Acoustics attached to the sides of the support. They could predict failure at an average of 95±7% of the fatigue life on average. AE, in both studies, was used to distinguish between tribological processes and material fracture and predict fatigue failure of bones.

As humans age, the likelihood of developing osteoarthritis increases. Researchers have tribologically assessed the knee joint to diagnose osteoarthritis. It is understood that early diagnosis can lead to better management of diseases of this type, thereby reducing the likelihood of the knee degenerating to the point of requiring replacement. Shark et al. (L.-K. Shark, Chen, and Goodacre 2010) used a bespoke joint acoustic analysis system (JAAS) to assess knee osteoarthritis by identifying the differences between healthy and Osteoarthritic (OA) knee joints.

The JAAS (Figure 2.8) is a measurement system that combines a traditional AE system with an electronic angle measurement system to provide AE data based on the joint angle. Knee data was collected from 8 healthy subjects and 5 osteoarthritic subjects. To allow for comparison of the results, the mean age of the groups was kept in the same range - OA knees,

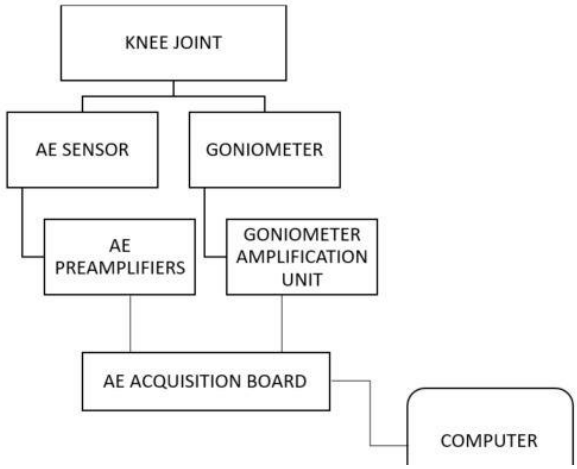


Figure 2.8: Illustration of the Joint Acoustic Analysis System

71.4 years and healthy knees, 71.5 years. They had the subjects perform sit-stand-sit movements in four phases: ascending-acceleration, ascending-deceleration, descending-acceleration, and descending-deceleration. AE data was acquired throughout via a wide band piezoelectric sensor attached to the knee joint and a goniometer attached to the lateral aspect of the knee for joint angle measurement. Recorded AE signals were analysed based on two AE features (Average Signal Level (ASL) and Peak Magnitude) and it was observed that OA knees produced AE events with higher peak magnitude than healthy knees and the maximum ASL of AE events for OA knees were higher than that of healthy knees (L.-K. Shark, Chen, and Goodacre 2010). There were able to successfully use AE to differentiate between healthy and OA knees.

In an extended study by the same authors, statistical analysis of the results identified further grouping of the OA knees based on different categories like age and progression of arthritis thereby indicating the possibility of a relationship between knee AE intensity and the progression of joint ageing and degeneration (L. K. Shark, Chen, and Goodacre 2011). A similar study by Khan & Yoho analysed the integrity of the knee joint between two groups of participants - young (average age of 23 years) and aged (average age of 42 years) (Khan and Yoho 2016) (refer to Table 2.2 for more information). They found that the aged knees generated AE signals with higher amplitude than the younger ones. This and the other studies show that there is a strong potential for knee AE to be used as a tool for the diagnosis of joint conditions, especially because it is a non-invasive procedure.

2.4.3 Implant condition monitoring

The promise shown by AE as a tool for measuring the tribology of human joints has prompted researchers to conduct studies of its use in arthroplasty research for implant condition monitoring as well as in the evaluation of implant failure mechanisms.

One of the first applications of AE in arthroplasty research was by Sugiyama et al. (cited in (Kapur 2016)) in 1989. Using AE, they established that hip implants were prone to failure at the implant-cement interface during torsional loading.

Rowland et al. (Rowland et al. 2004) and Rodgers et al. (Rodgers et al. 2014) used AE to monitor the health of hip implants. Rowland et al. attempted to correlate AE activity of metal-on-metal hip implants to out-of-line wear using a five-station wear rig. Pancom P15 sensors were attached to the upper and lower specimens (Figure 2.9). The presence of high attenuation made it possible for the upper and lower specimen to be isolated which in turn made zonal determination possible. They were able to show that there was a correlation between AE signals and out-of-line wear of metal-on-metal hip prostheses. How the AE signals related to the cup and head displacements and the load cycle was not given (Rowland et al. 2004).

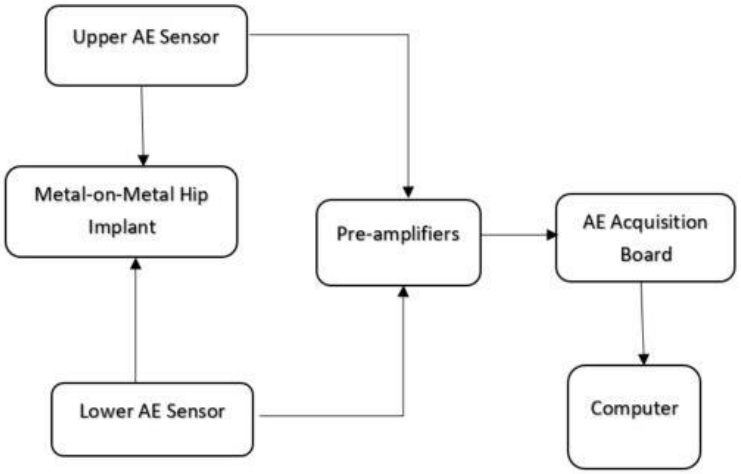


Figure 2.9: Illustration showing acquisition of AE signal from the hip implant

In contrast, Rodgers et al. (Rodgers et al. 2014) used their study to compare *in vivo* and *in vitro* data. For the *in vivo* tests, they used an AE system consisting of four sensors (so that the source of vibrations could be identified) placed against the skin of the patients to monitor the implants. The focus of the study was to characterize the squeaking of hard-on-hard bearing

surfaces so they chose patients who were known to have experienced audible squeaking, and had them perform a range of different motions including sit-to-stand and stand-to-squat. Squeaking is a characteristic of the stick-slip phenomenon which occurs during the transition between static and dynamic friction in a contact pair (Ferrer et al. 2010). Asperities are in greater contact during slip leading to the presence of more AE events, therefore, audible squeaking is an effect of the slip phase. They compared the *in vivo* results to the *in vitro* result from one retrieved hip implant and found that the *in vitro* tests had higher magnitudes than the *in vivo* tests but they both exhibited similar frequency characteristics suggesting that the *in vivo* results were lower due to attenuation effect. Another instance where AE has been successfully used for implant condition monitoring is in the monitoring of dental implants. These studies include the *in vitro* assessment of the bone-implant interface using AE (Ossi et al. 2012), the transmission of AE in bones, implants and dental materials (Ossi et al. 2013) and for dental biomechanical testing under cyclic loads (Huang, Chen, and Lin 2016).

It has been found that hip implants experience corrosive and fretting wear mechanisms at the stem-cement interface (Bryant et al. 2014; Howell et al. 2004; Zhang et al. 2009). Fretting will induce stress damage that has been shown to be detectable using AE (Briscoe et al. 2001). Hence, there is potential for using AE to detect fretting damage in implants.

2.4.4 Effect of Attenuation on AE Tests

Since bone within joints is surrounded by cartilage, muscle and fluids within the body, it is expected that there will be attenuation of the AE signal. Strantza et al (Strantza et al. 2014) carried out a study to highlight the complexity of wave propagation through bone tissue. They used AE sensors to conduct elastic stress wave measurements in human femur bones using the *pencil lead break method* for pulse excitation. This involves placing the lead of a pencil against the structure under test and applying a gentle load until it breaks. This break will allow the release of stress accumulated while force was applied which leads to a microscopic

displacement of the surface thereby causing the propagation of a pressure wave into the structure (Sause 2011). When compared to waveforms from a bulk and homogeneous metal block, it was found that the femur bones exhibited stronger dispersive and attenuative trends. This was due to the micro-structure and geometry of the bone. Increasing the propagation distance also led to strong changes in parameters such as rise time, central frequency and average frequency (Strantza et al. 2014).

Khan-Edmundson et al. (Khan-Edmundson et al. 2012) conducted a study to determine the range of frequencies observed on the skin surface and consequently developed a soft tissue attenuation model. They placed 4 ultrasonic sensors against the skin (between the greater trochanter and the mid femur) of a patient and recorded data as they undertook motions such as standing from sitting and walking up the stairs. Using statistical methods such as Fourier transforms and bode plots, the maximum frequencies present at the skin were around 20 kHz with majority below 10 kHz (Khan-Edmundson et al. 2012). During *in vitro* studies, there will be interferences such as background noise which could lead to misinterpretation of signals when relating to *in vivo* behaviour. Having an attenuation model makes it easier to relate *in vitro* bench tests to actual *in vivo* behaviour.

Mavrogordato *et al*. (Mavrogordato et al. 2011) proposed a possible solution by having the sensors embedded into the femoral stem of the hip implant before implantation. It was believed that this will not only reduce attenuation effects but also eliminate noise interference. They observed that specimens with embedded sensors produced a higher number of cumulative hit counts than those with external sensors only showing that the effect of attenuation had been reduced. The limitation to this method is that a failed sensor would not be replaceable as well as manufacturing difficulty as there are lots of factors (e.g., material compatibility, longevity and expense) to consider. Despite these limitations, an embedded sensor has a lot of potential as an effective method of minimising the attenuation effect of AE signals.

Table 2.2: Overview of selected biotribological applications of AE

AUTHOR	REFERENCE JOINT	TEST CONDITION	AE ACQUISITION SYSTEM	ANALYSIS METHOD	CONCLUSION
Shark et al. 2010 (L.-K. Shark, Chen, and Goodacre 2010)	Healthy and OA Knees	*In vivo*	JAAS	Parameter-based technique	JAAS showed potential as a tool for assessing knee joint conditions both in clinic and home settings.
Shark et al. 2011(L. K. Shark, Chen, and Goodacre 2011)	Healthy and OA Knees	*In vivo*	JAAS	Parameter-based technique	Wider test group confirmed result from previous studies.
Van Toen et al. 2012 (Van Toen et al. 2012)	Cadaveric Spine (Ligamentum flavum, LF and vertebral body, VB)	*In vitro*	Nano 30 sensor	Parameter and frequency-based techniques	There is potential for using AE to differentiate between failures of different spinal components.
Khan-Edmundson et al. 2012 (Khan-Edmundson et al. 2012)	Artificial Hip	*In vivo*	4 passive ultrasonic sensors	Frequency-based statistical analysis	Soft tissue attenuation influences AE signals.
Hua et al. 2014 (Hua, Fan, and Jin 2014)	Artificial Hip	*In vitro* tribo-acoustic rig test	Sound pressure was acquired with an omni-directional microphone	Parameter and frequency-based techniques	There is potential for tribo-acoustic tests for artificial joints.
Rodgers et al. 2014 (Rodgers et al. 2014)	Artificial Hip	*In vivo* and *in vitro* tests (retrieved hip implants)	4 passive ultrasonic sensors and National Instruments Compac DAQ	Frequency-based technique	Validates relating bench test result to implant performance and degradation within a patient.
Arun et al. 2014 (Arun et al. 2014)	Cadaveric Spine	*In vitro* compression test via impact loading	4 Nano 30 sensors	Frequency-based technique	Frequency-based analysis (STFT & spectral analysis) provides a complete understanding of tribological mechanisms present.
Khan and Yoho 2016 (Khan and Yoho 2016)	Young and aged knees	*In vivo*	4 R6α AE sensors and goniometers for joint angle measurements	Parameter-based technique	Onset of knee degeneration can be diagnosed using AE.
Fitzpatrick et al. 2017 (FitzPatrick et al. 2017)	Artificial Hip	*In vivo* and *in vitro* tests (retrieved implants)	4 passive ultrasonic sensors and National Instruments Compac DAQ	Frequency-based technique	AE has potential as a tool for monitoring the condition of artificial hips *in vivo*.

2.5 Bio-tribo-acoustic study – Aims and Objectives

In order to achieve the potential of using AE to identify and diagnose failed artificial joints, a test system that can combine biotribological study with AE testing is required. AE testing has already been shown to be capable of identifying tribological mechanisms of mechanical systems through tribo-acoustic testing. It has been established that artificial joints fail mostly through biotribological mechanisms and the literature has shown that AE can be advantageous for implant condition monitoring. This shows there is potential for using AE to monitor tribological mechanisms of artificial joints for the purpose of identifying and diagnosing their failure.

The overall aim of the research presented in this thesis is to diagnose tribological phenomenon in artificial joint materials using AE. This aim will be achieved by undertaking the following objectives:

1) Predict the coefficient of friction profile using time dependent AE features.
2) Investigate the progression of damage in a simulated artificial joint articulating surface using AE.
3) Undertake a fundamental study on the use of supervised and unsupervised learning to classify AE signals.
4) Identify the damage mode in a simulated artificial knee joint articulation using AE testing.

2.6 Summary

This chapter has presented a review of previous research relevant to this thesis including the aims and objectives. Coming up in chapter 3 is the general materials and methods used for the experimental work presented in this thesis.

Chapter 3

General Materials and Methods

3.1 Introduction

In the previous chapter, a review of previous research relevant to this thesis has been presented including the aims and objectives of the research project. This chapter presents the general materials and methods used in this thesis. The main materials used are presented in section 3.2. In section 3.3, the test methods, test parameters and data analysis that are employed in multiple chapters are presented. Methods and data analysis unique to each experimental work can be found in their individual chapters.

3.2 Materials

Materials used for the work reported in this thesis are those that have been used in the design and manufacture of artificial joints. They were chosen to represent a metal-on-polymer joint replacement articulation. The materials are presented in Table 3.1.

Table 3.1: Materials used in alphabetical order. Properties were either provided by the suppliers or from other public sources (Geetha et al. 2009; Xin, Shepherd, and Dearn 2012).

Material	**Young's Modulus, GPa**	**Poisson's Ratio**	**Density, g·cm^{-3}**
CoCrMo	240	0.29	8.4
PEEK	3.4	0.36	1.29
STEEL	210	0.29	8
UHMWPE	0.7	0.46	0.93

3.3 Methods

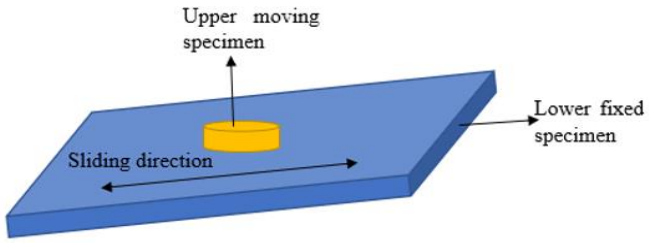
3.3.1. Tribo-Acoustic Test System

to allow for the acquisition of AE signals during tribological testing, a tribo-acoustic test system was designed. This consists of a tribometer for friction tests and an AE acquisition system for the acquisition and recording of AE signals.

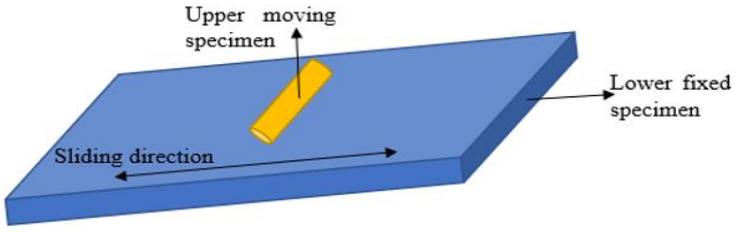
Friction Tests

In the design of artificial joint replacement devices, tribometers have been employed in material bench testing to investigate their suitability for the intended use before carrying out long term wear tests using joint simulators (Olorunlambe, Shepherd, and Dearn 2019). All tests reported in this thesis were carried out on the TE77 High-Frequency Friction Machine (Phoenix Tribology, Newbury, UK). It is a multi-function test machine for investigating materials' friction, lubrication, and wear behaviour. It can be loaded up to 1000 N and has a maximum stroke of 25 mm making it capable of simulating various sliding contact conditions for different applications. It has been shown to be capable of simulating biomechanical contacts such as hip replacement joints (Choudhury et al. 2013) making it a suitable choice of tribometer for the tests reported in this thesis.

It has a COMPEND 2000 software connected to a PC for sequence control of test parameters such as load and frequency and the acquisition of measured data such as coefficient of friction. The friction force is measured using a piezo-electric transducer and the coefficient of friction is calculated on-line using the measured load and the time-smoothed friction value.



(a) The TE77 face contact configuration, used in chapters 4 & 6.



(b) The TE77 line contact configuration, used in chapter 5.

Figure 3.1:The TE77 test configurations used in this thesis. (a) The face contact was used for tests in chapters 4 & 6; and (b) The line contact was used for tests in chapter 5.

The TE77 is capable of a number of several test configurations, but the two employed in this thesis are the line contact and the face contact (Figure 3.1). The experimental layout is shown in Figure 3.2.

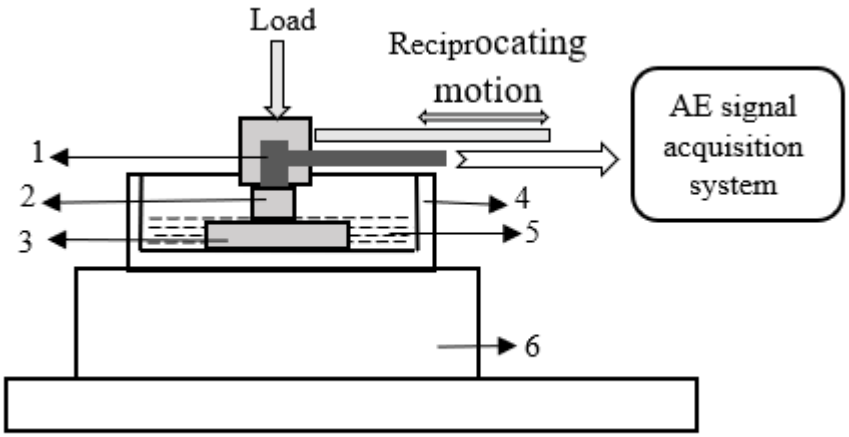


Figure 3.2: Experimental layout of the TE77 tribometer and the AE acquisition system. (1) AE sensor mounted on (2) polymeric upper specimen, (3) flat lower specimen, (4) lubricant bath, (5) lubricant, (6) heater block.

To simulate the *in vivo* joint environment, tests were conducted with Oxoid quarter strength Ringer's solution (ThermoFisher Scientific, Oxoid Limited, Hampshire, UK) serving as the lubricating medium at a temperature of $37\pm2^{\circ}$C (except for the adhesive tests in chapter 5 which were conducted without a lubricant to maximise the adhesive process). The solution was prepared by dissolving one tablet (~1.2 g) in 500 ml of distilled water and sterilised in an autoclave at 121°C for 15 minutes. The prepared Ringer's solution was drip-fed onto the contacting surface at 0.1 ml per minute. The Ringer's solution tablet was made from sodium chloride, potassium chloride, calcium chloride $6H_2O$ and sodium bicarbonate 0.05. Temperature has also been known to have an effect on AE signals (Boon et al. 2014) but since the temperature for all tests was kept constant, these effects were considered negligible.

Acquisition and Post-Processing of Acoustic Emission Signals

AE signals were acquired using an acquisition and recording system (Mistras Group, Cambridge, UK) comprising of a nano-30 AE sensor, a 2/4/6 preamplifier with a gain of 60 dB and the AEWin PCI2 software for the conditioning and storage of the detected signal.

The nano-30 miniature sensor was chosen due to its small size (8 mm outside diameter and a length of 8 mm, see Figure 3.3) so that it can be mounted directly on to the polymeric specimens (Figure 3.2). This will help reduce the effect of attenuation on the acquired signals. It has a resonant frequency of 300 kHz and a good frequency response over the range of 125 – 750 kHz. To ensure that significant AE signals are acquired, certain timing parameters need to be set by the user. Selected AE acquisition parameters that need to be defined before testing are presented in Table 3.2.

The acquired AE signals were post-processed using the NOESIS Advanced AE Analysis software (Mistras Group, Cambridge, UK) to compute new discrete AE features from the waveforms for further analyses. Post-processing of AE signals can produce 50+ features which can be computationally expensive during analysis. The number of features was minimised by focusing on the features most used in literature. The AE features used in this thesis and their definitions can be found in Table 3.3.

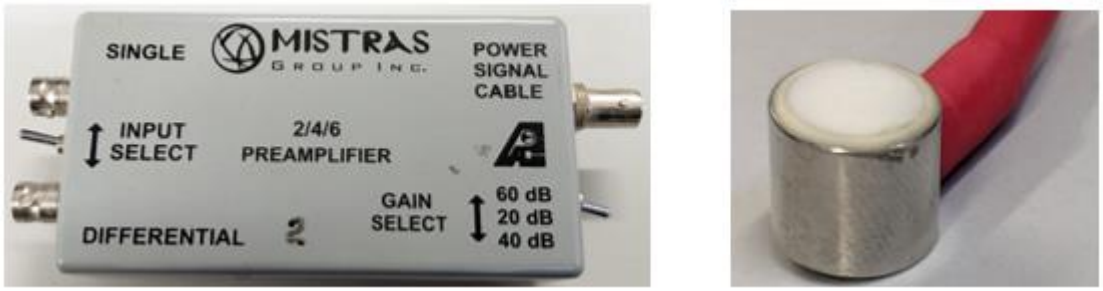


Figure 3.3: Pictures of the 2/4/6 Preamplifier (Left) and the Nano-30 AE Sensor (Right)

Table 3.2: AE Timing Parameters and their descriptions

AE TIMING PARAMETERS	DEFINITION
Peak Definition Time (PDT)	Ensures actual peak of the waveform is recorded.
Hit Definition Time (HDT)	Ensures that all counts in one hit are recorded.
Hit Lockout Time (HLT)	Used to decrease the negative effect of reflection in reverberant environment.

Table 3.3: Selected AE features and their definitions in alphabetical order.

AE FEATURES	DEFINITION
Absolute Energy (aJ)	The true energy of the signal on a 10 kohm resistor
AE Root Mean Square (V)	Root mean square of the area under the voltage curve
Amplitude (dB)	Maximum amplitude of the signal
Average Frequency (kHz)	Signal counts over signal duration
Average Signal Level (dB)	The RMS value converted to the dB scale
Counts to Peak (#)	Number of crossings from first threshold crossing to the point where maximum amplitude is reached
Duration (µs)	The time from first threshold crossing to the last threshold crossing
Peak Frequency (kHz)	Frequency corresponding to the peak value of the power spectrum of the FFT transform
RA Value (µs/dB)	Time per amplitude needed for signal to reach its peak value. Expressed as ratio of risetime to amplitude.
Risetime (µs)	The time from first threshold crossing to when the maximum amplitude is reached
Signal Strength (pVs)	The area under the signal envelope

3.3.2. Friction Test Parameters

TE77 Load

The TE77 test load used in the work presented in chapters 4 and 5 was determined using Hertzian contact mechanics (Johnson 1985) and the Charite Lumbar Spinal Implant (LSI) was used as an exemplar joint (Figure 3.4) . The contact pressure of a ball-and-socket Charite Lumbar Spinal Implant (with radii of 10 mm and 14 mm with a radial clearance of 14 0.35

mm) was calculated using the load and displacement conditions defined by BS ISO 18192-1 for wear of total intervertebral spinal disc prostheses (Moghadas et al. 2012; BS ISO 18192-1 2011). The calculated contact pressure was then used to find an equivalent load based on the specific test configuration (face contact in chapter 4 and line contact in chapter 5). The equations used for the calculations are presented in Equation 3.1 to Equation 3.4. The calculated maximum contact pressure can be found in Table 3.4.

$$\frac{1}{E^*} = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2}$$

Equation 3.1: Calculation for equivalent elastic modulus, E^*, where E_1 & E_2 represent the elastic moduli of the socket and ball materials, respectively, and ν_1 & ν_2 are the Poisson's ratios of the materials.

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$$

Equation 3.2: Calculation for equivalent radius, R, where R_1 & R_2 represent the radii of the socket and balls materials, respectively.

$$a = \left(\frac{3PR}{4E^*}\right)^{\frac{1}{3}}$$

Equation 3.3: Calculation for contact area radius, a, where P is the normal applied load.

$$p_0 = \frac{3P}{2\pi a^2}$$

Equation 3.4: Calculation for maximum contact pressure, p_0.

Table 3.4:Maximum contact pressure a ball-and-socket Charite Lumbar Spinal Implant is subjected to. Calculated using Hertzian contact mechanics.

Ball radius, mm	Load, N (BS ISO 18192-1 2011)	Maximum contact pressure, MPa
10	600	10.1
	2000	15.1
14	600	6.5
	2000	9.7

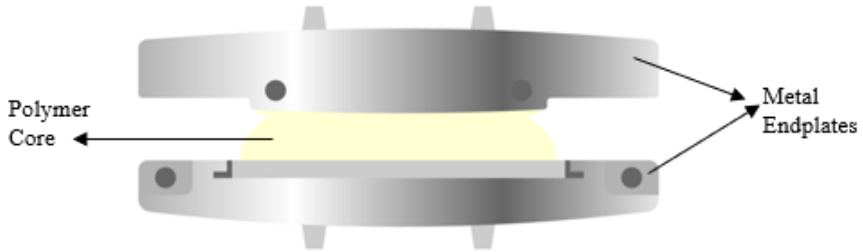


Figure 3.4: Image of a Lumbar Spinal Implant.

Choice of contact surface configuration

In order to closely simulate a ball and socket contact, a sphere-on-plane configuration was initially considered but calculations showed that a load of about 1 N would be required on the TE77 to simulate a contact pressure close to those achieved in a ball-on-socket Charite LSI. Such a low load could make for unstable friction calculation and a minimum sphere diameter of 20 mm would be required which is more than the 6 mm & 10 mm sphere diameters the TE77 is configured for. Hence, a disc-on-plate configuration (i.e., face contact) was chosen as an alternative for the proof-of-concept tests presented in chapter 4. The calculations can be found in the appendix.

Sliding Velocity

The maximum sliding velocity a lumbar spinal implant is subjected to during long-term wear testing on a spine simulator was calculated using simple harmonic motion equations and

motion values from the BS ISO 18192-1 standard (BS ISO 18192-1 2011). Using the maximum sliding velocity calculated (6.909 mm/s) and a frequency of 2 Hz, an equivalent stroke length for the TE77 was found to be 1.0996 mm. With such a low stroke length, only fretting wear would occur and there is an added risk of fluid entrapment due to the large diameter of the discs (10 mm). A stroke of 12.5 mm was chosen instead to obtain an average sliding velocity of 50 mm/s (when sliding at a frequency of 2 Hz) which is the recommended value for a linear reciprocating wear motion in the standard for wear testing of polymeric materials used in total joint prostheses (see Annex A1 of ASTM International 2017). This stroke value was kept constant for all tests presented in this thesis.

3.3.3. Preparation of Test Specimens

Before and after each test, polymeric test specimens were cleaned using the method described in appendix A6 of the Standard Test Method for Wear Testing of Polymeric Materials Used in Total Joint Prosthesis (ASTM International 2017). First, they were rinsed with tap water to remove bulk contaminants and then washed in an ultrasonic cleaner in a solution of 1% detergent for 15 minutes. They were then rinsed in a stream of distilled water before rinsing again in an ultrasonic cleaner in distilled water for a further 5 minutes. Afterwards, they were dried with lint-free tissue and then immerse in ethanol for 3 minutes. After immersion in ethanol, they were dried again with lint-free tissue and then air-dried at room temperature for 30 minutes. Metallic specimens were washed in ethanol before and after each test.

3.3.4. Optical Imaging

The Alicona InfiniteFocusG5 Optical 3D Measurement System (Alicona Imaging Gmbh, Raaba, Austria) was used to obtain and analyse images of the wear scars. Images were taken at a magnification of 20x with the polariser turned off to improve the quality of the image.

3.3.5. Data Analysis

After post-processing of acquired AE signals in NOESIS, further analyses were carried out using MATLAB versions R.2019a and R.2020a (MathWorks, Cambridge, UK).

The data analysis presented here was used in multiple chapters. The chapter-specific ones can be found in their respective chapters.

Clustering of AE data

Clustering is an unsupervised pattern recognition technique used to group data sets into one or more clusters based on similarities and differences detected between data points. AE signals from tests in chapters 4 & 5 were clustered using the k-means clustering method. This method clusters data by minimising the sum of Euclidean distances from all vectors of a cluster to its centre (Momon et al. 2012). The k-means clustering algorithm is as follows;

1. Sample data sets were defined as $X = \{x_i | i = 1,2,\cdots,n\}$, $C_j(j = 1,2,\cdots,k)$ where X denotes the k categories of clusters and $c_j(j = 1,2,\cdots,k)$ represents the initial cluster centre. The clusters $C_1, C_2, \cdots, C_k$ satisfy:
    a. $C_i \neq \emptyset, i = 1,2,\cdots,k$;
    b. $C_i \cap C_j = \emptyset, i,j = 1,2,\cdots,k; i \neq j$;
    c. $\sum_{i=1}^{k} C_i = \{x_1, x_2, \cdots, x_n\}$.
2. k samples were randomly selected, and $(c_1, c_2, \cdots, c_k)$ was defined as the initial clustering centre.
3. Using the squared Euclidean distance, each sample in the data sets $\{x_i\}$ was assigned to the k cluster centres c_i.
4. The centre of a new cluster $c_i(i = 1,2,\cdots,k)$, i.e., $c_i = \frac{1}{n}\sum_{i=S_i} x$, where n is the S_i cluster domain containing the number of samples, could then be calculated, if $c_i \neq$

$c_i(i = 1,2,\cdots,k)$ then step (3) was repeated. Otherwise, the algorithm converged, and the analysis ends.

To determine the optimal cluster number *k*, the Silhouette Index (SI) was plotted and the cluster number with the highest SI value was chosen as the optimal cluster number *k*.

Summary

This chapter has presented the general materials and methods employed in this thesis. Coming up in chapters 4, 5 and 6 are the presentation and discussion of the results from the three main areas of experimental work carried out to meet the aims and objectives of this thesis.

Chapter 4

Condition Monitoring of a Simulated Joint Articulation - A proof-of-concept study

The content of this chapter is based on the following publication:

Olorunlambe K.A.; Eckold D.G.; Shepherd D.E.T.; Dearn K.D. "**Bio-tribo-acoustic emissions: condition monitoring of a simulated joint articulation**" in Biotribology. – *Article in Press*

4.1. Introduction

Chapter 2 has shown that AE testing has potential as a tool for condition monitoring of artificial joints. To achieve this potential, the relationship between AE signals and the wear of artificial joints first needs to be established. This is a challenging prospect because the wear of artificial joints is traditionally diagnosed during post-retrieval analyses of failed implants. Since the causes of failure of implants can be related to wear and wear cannot happen without friction occurring within the articulating surfaces (Simon et al. 1975; Moghadas et al. 2012), an alternative is to establish a relationship between AE features and the coefficient of friction to interpret the wear behaviour. The study by Patzer and Woydt has shown that combining AE with the coefficient of friction tests can help improve the interpretation of wear behaviour (Patzer and Woydt 2021).

This chapter presents a proof-of-concept study on how AE features can be used to predict the coefficient of friction profile of simulated articulating joint replacement surfaces as a first step towards achieving the potential of AE as a tool for diagnosing joint pathologies. Time-dependent (TDD) AE features are used to predict the coefficient of friction profile during tribo-acoustic testing. The distribution of the different emission types from running-in to the prolonged sliding stage is also investigated to understand the evolution of surface damage.

4.2. Materials and Methods

4.2.1. Experimental Parameters

AE signals were acquired using the tribo-acoustic test set-up described in chapter 3. Tests were performed using ultra-high-molecular-weight polyethylene (UHMWPE) discs (supplied by Penta Precision, Portsmouth, UK) as the reciprocating specimen and medical grade cobalt chromium molybdenum alloy (CoCrMo) (supplied by Phoenix Tribology,

Newbury, UK), machined to a surface finish of less than 0.02 μm R_a, as the fixed specimen. For comparison purposes, poly-ether-ether-ketone (PEEK, supplied by Penta Precision, Portsmouth, UK) was chosen as an alternative material. Studies have shown PEEK to be a viable bio-material used in several biomedical applications such as in cervical TDR devices (Xin, Shepherd, and Dearn 2012). It has also shown promise for use in hip implants (Nakahara et al. 2013) making it an appropriate choice of an alternative material for testing. The PEEK and UHMWPE discs were of diameter 10 mm and 4 mm thickness and machined to a surface finish of 1.16 μm R_a and 3.25 μm R_a (measured with the Alicona InfiniteFocusG5 Optical 3D Measurement System from Alicona Imaging Gmbh), respectively.

Equivalent TE77 load was calculated using the maximum contact pressure reached by a Charite LSI with a 14 mm ball radius based on the loading parameters stated in BS ISO 18192-1 2011 (See Equations 3.1 to 3.4 and Equation 4.1). The final load is chosen and other test parameters are presented in Table 4.1. For an understanding of how AE features behave at different sliding velocities, tests were carried out at 2 frequencies (2 Hz and 4 Hz). The 2 Hz tests resulted in a sliding distance of 360 m at 0.05 $m \cdot s^{-1}$ while the 4 Hz tests resulted in a sliding distance of 720 m at 0.1 ms^{-1}. Both sliding velocities were within the operating envelope of the TE77.

$$P^* = p_0 a^*$$

Equation 4.1: Calculation for equivalent TE77 load, P, in a face contact where a* is the nominal contact area.*

During post-processing of AE signals using the NOESIS software, feature extraction settings were kept the same as the acquisition settings apart from the threshold which was increased to 40 dB as initial observation of the acquired signals (see Figure 4.1) show that some noise was still present in the data.

Table 4.1: Summary of Test Parameters

	Parameters	Values
TE77 PARAMETERS	Load	660 N
	Frequency	2 Hz, 4 Hz
	Stroke Length	12.5 mm
	Sliding Velocity	50 mm/s, 100 mm/s
	Test Duration	2 hours
AE ACQUISITION PARAMETERS	Preamplifier gain	60 dB
	Threshold	35 dB
	Sampling Rate	2 MHz
	Peak Definition Time (PDT)	400 µs
	Hit Definition Time (HDT)	400 µs
	Hit Lockout Time (HLT)	1000 µs
	Maximum Hit Duration	1000 ms
	Band Pass Filter	100 – 400 kHz

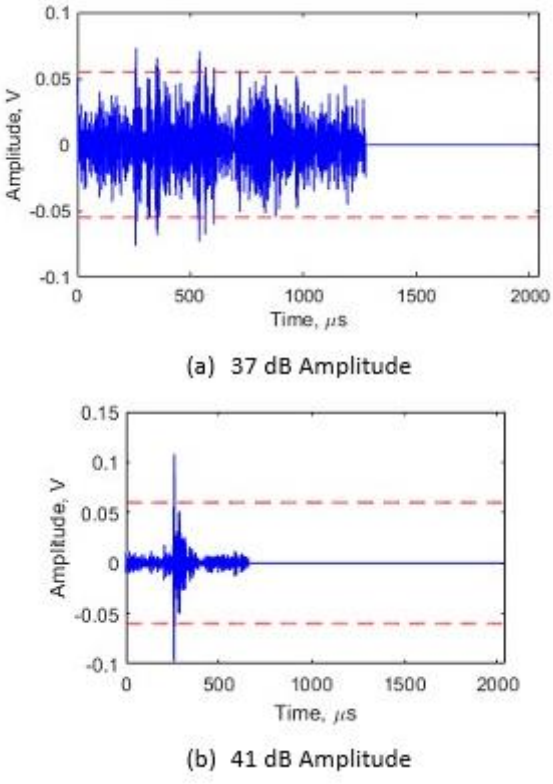


Figure 4.1: Raw AE signal at (a) 37 dB and (b) 41 dB amplitudes. The 37 dB signal is noisier than the 41 dB signal.

4.2.2. Data Analysis

Time Series Neural Network

Artificial neural networks (ANN) are computational architectures modelled after the brain's architecture (Curry and Rumelhart 1990). To further explore the relationship between time-dependent AE features and the coefficient of friction (CoF) profile, a nonlinear autoregressive neural network with external inputs (NARX) was deployed to predict CoF. NARX is a dynamic form of ANNs based on the linear ARX, commonly used in time series analysis (Narendra and Parthasarathy 1991; Chen, Billings, and Grant 1989; Ouyang 2017). It effectively solves nonlinear time series problems (Ouyang 2017), making it suitable for this study. The NARX model equation is shown in Equation 4.2, where y(t), the predicted value of the model, is predicted using its previous values, y(t-1) to y(t-n_y), and the corresponding previous values of an external input signal, x(t-1) to x(t-n_x). There are two forms of NARX architecture: series-parallel and parallel architecture (Ouyang 2017). The series-parallel NARX architecture was chosen for this study because the availability of the previous values of y(t) makes the network more accurate. The Neural Net Time Series App in MATLAB version R2021a was used for training and testing the NARX models.

$$y(t) = f(y(t-1), y(t-2), \dots, y(t-n_y), x(t-1), x(t-2), \dots, x(t-n_x))$$

Equation 4.2: NARX model equation

Herein, y(t) is the coefficient of friction recorded throughout the test and x(t) is a 3-element vector consisting of the three time-dependent AE features – absolute energy, average signal level, and AE root mean square value. The Neural Net Time Series App in MATLAB version R2021a was used for training and testing the NARX models. An illustration of the NARX network is shown in Figure 4.2. There will be differences in how the two polymeric materials behave when sliding against the CoCrMo plate due to different material properties; hence, two NARX models were built - one for PEEK and the other for UHMWPE. Both models

were trained using the Levenberg-Marquardt training function with ten neurons in the hidden layer. Training data were randomly divided into three splits - training data (70%), validation data (15%) and test data (15%). Performance was evaluated using mean squared error (MSE) and R^2 values. A low MSE and high R^2 value indicate a good training performance.

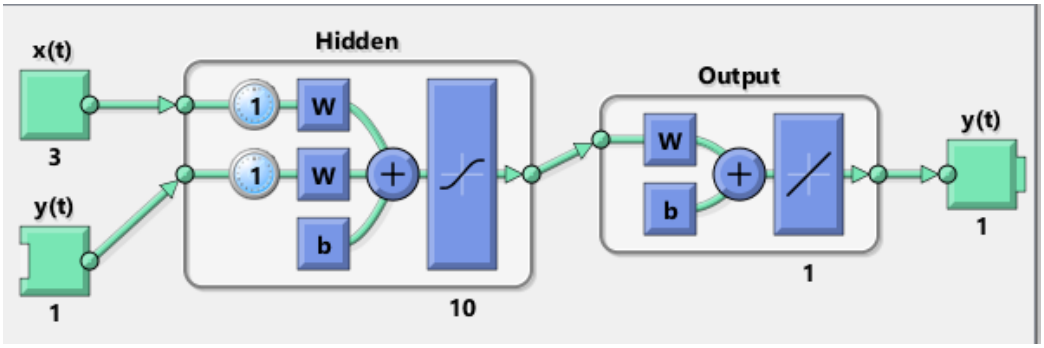


Figure 4.2: NARX Neural Network. Obtained from MATLAB.

K-means clustering

To understand the distribution of AE signals as the tests progressed from one stage to the other, acquired hits were clustered and then categorised into one of the three emission types – burst (hits with high amplitude and short duration), continuous (hits with low amplitude and long duration) and mixed mode (mixture of burst and continuous). The k-means clustering method was employed. The method is explained in detail in chapter 3 (section 3.2.5).

Since the three emission types can be distinguished using duration and amplitude values (Olorunlambe, Shepherd, and Dearn 2019), they were the AE features fed into the clustering algorithm and the initial number of k clusters was set to three emission types. The silhouette index (S.I.) plot was then used to determine if it was the optimal number of clusters.

4.3. Results and Discussion

4.3.1. Using AE to predict coefficient of friction profile

The three time-dependent AE features: root mean square value (RMS), average signal level (ASL) and absolute energy (AbsE), all exhibit similar transient response to the coefficient

of friction (CoF) (Figure 4.3 & Figure 4.4) for both polymeric specimens at the two test frequencies. This suggests a strong potential for using time-dependent AE features to predict CoF. Other studies investigating the potential of AE as a tool for predicting tribocorrosion processes in dental implants and for the early detection of failure modes in total hip replacements also found that the coefficient of friction and absolute energy of the AE signal exhibits similar transient responses (Lee et al. 2021; Barão et al. 2021). The predicted CoF can also be used to infer the wear behaviour of test specimens since previous studies have shown that the coefficient of friction of PEEK and UHMWPE can be directly related to the wear rate (Petrica et al. 2016; Kanaga Karuppiah et al. 2008).

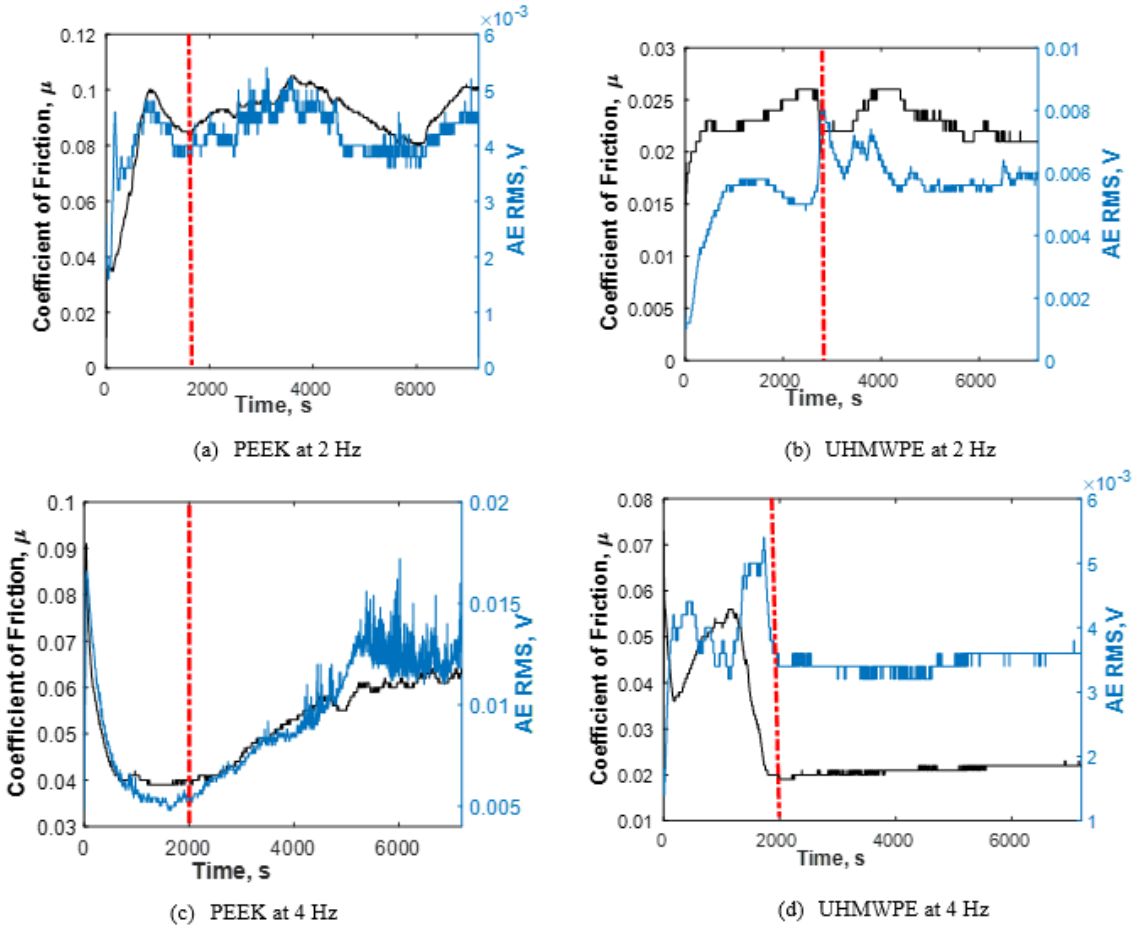


Figure 4.3: RMS (blue line) & CoF (black line) plots for (a) PEEK at 2 Hz, (b) UHMWPE at 2 Hz, (c) PEEK at 4 Hz and (d) UHMWPE at 4 Hz. Red dash line indicates transition from running-in (stage I, after lowest CoF is reached) to prolonged sliding.

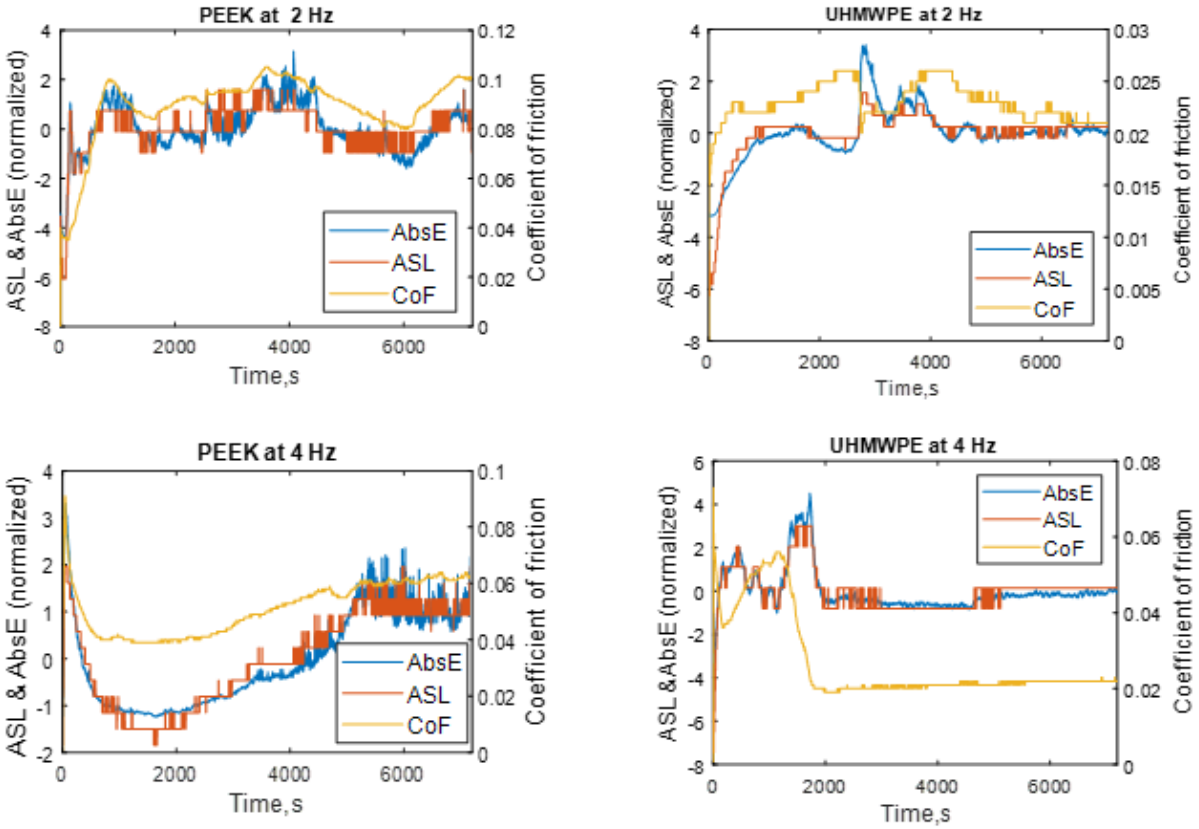


Figure 4.4:Plots of average signal level (ASL) and absolute energy (AbsE) in relation to coefficient of friction (CoF)

The correlation between all three time-dependent AE features and the coefficient of friction for both PEEK and UHMWPE (Figure 4.3 & Figure 4.4) make them suitable as the external input for the NARX neural network model. The training performance and training response is shown in Figure 4.5 and Figure 4.6.

Table 4.2 shows the mean square error (MSE) and R^2 values obtained after training both models. Both models have R^2 values greater than 90%, with PEEK having a slightly higher value than the UHMWPE model, implying that almost all variability is explained by both models indicating a good predictive capability. The trained network is then tested using data from the repeat bio-tribo-acoustic tests in open-loop feedback, and the results are presented in Table 4.3. The test response curves are presented in Figure 4.7 and Figure 4.8.

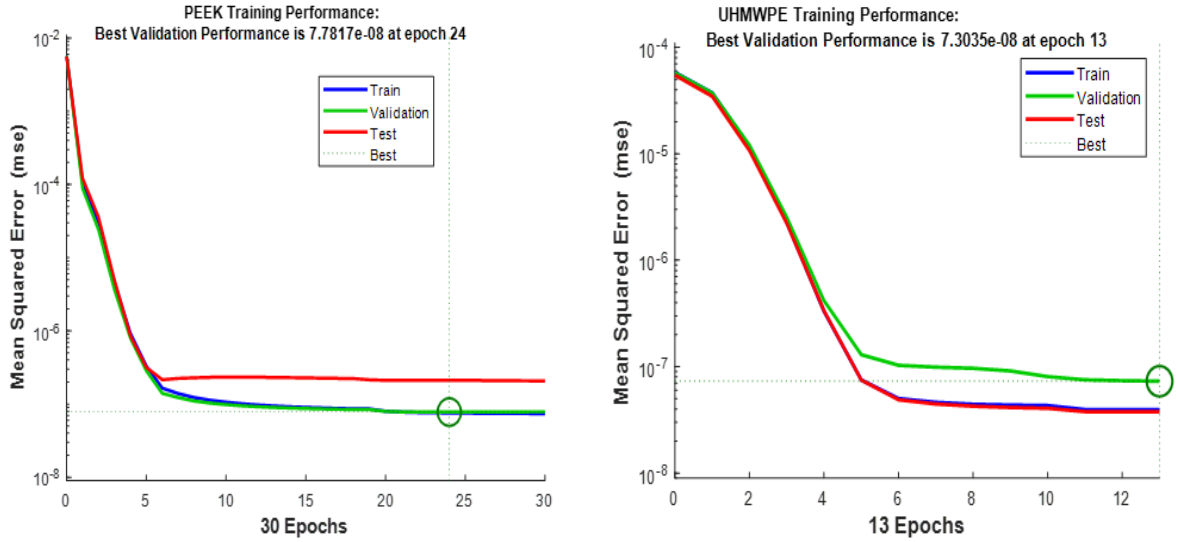


Figure 4.5: NARX Net Training Performance for PEEK (left) and UHMWPE (right)

Table 4.2: NARX Neural Network Training Result Summary

	PEEK	**UHMWPE**
MSE	7.78 x 10^{-8}	7.30 x 10^{-8}
R^2 value	0.9995	0.9858

Table 4.3: A summary of CoF prediction results with test data

	PEEK		**UHMWPE**	
	2 Hz	4 Hz	2 Hz	4 Hz
MSE	6.78 x 10^{-5}	6.39 x 10^{-7}	0.0011	1.08 x 10^{-6}
Regression, R	0.9718	0.9979	0.5976	0.8691
Mean CoF (Predicted) ± Standard Deviation	0.1079 ± 0.0150	0.0510 ± 0.0088	0.0251 ± 0.0062	0.0190 ± 0.0017
Mean CoF (Actual) ± Standard Deviation	0.1128 ± 0.0202	0.0513 ± 0.0093	0.0436 ± 0.0312	0.0187 ± 0.0020

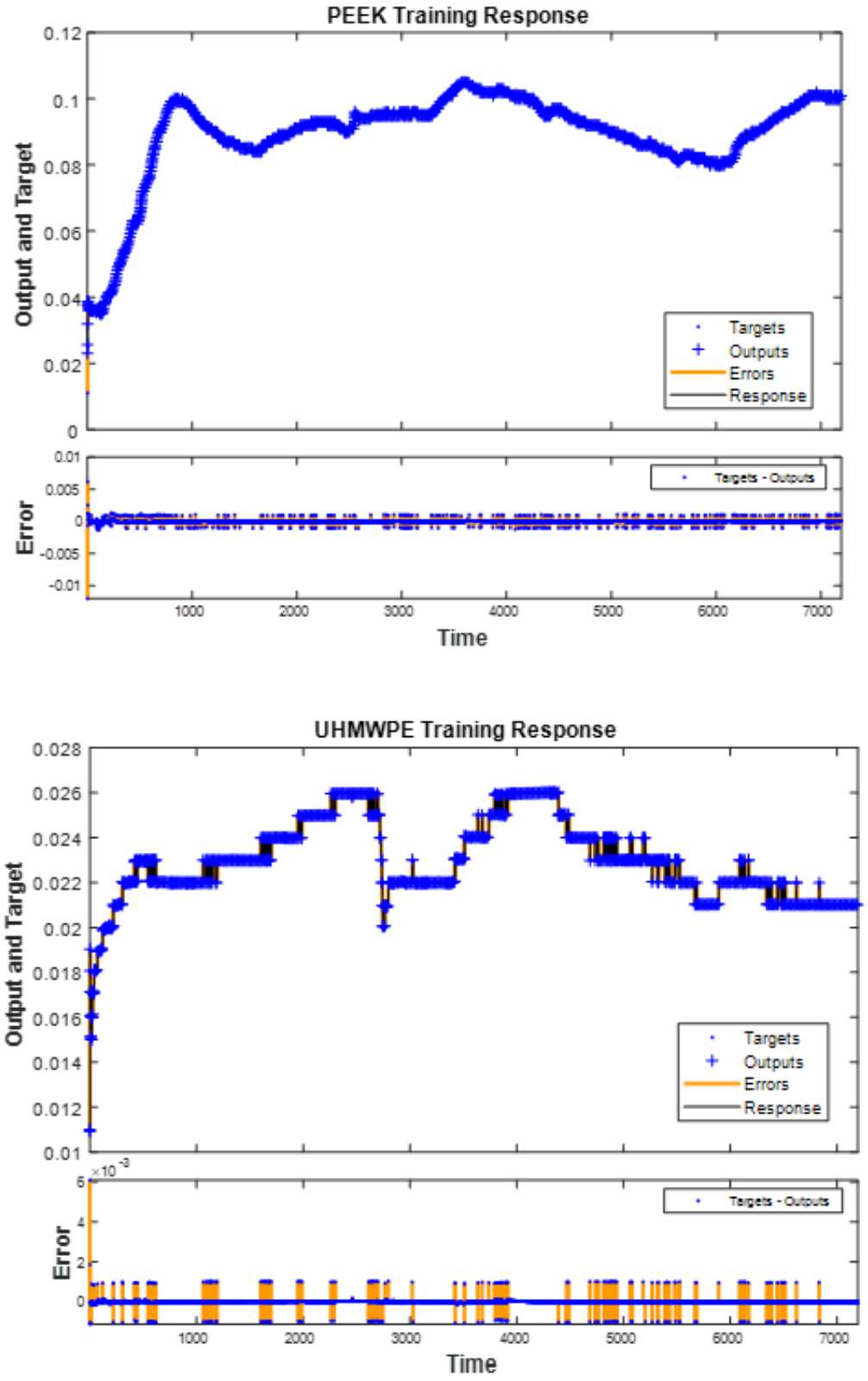


Figure 4.6: NARX Net Training Response for PEEK (top) and UHMWPE (bottom)

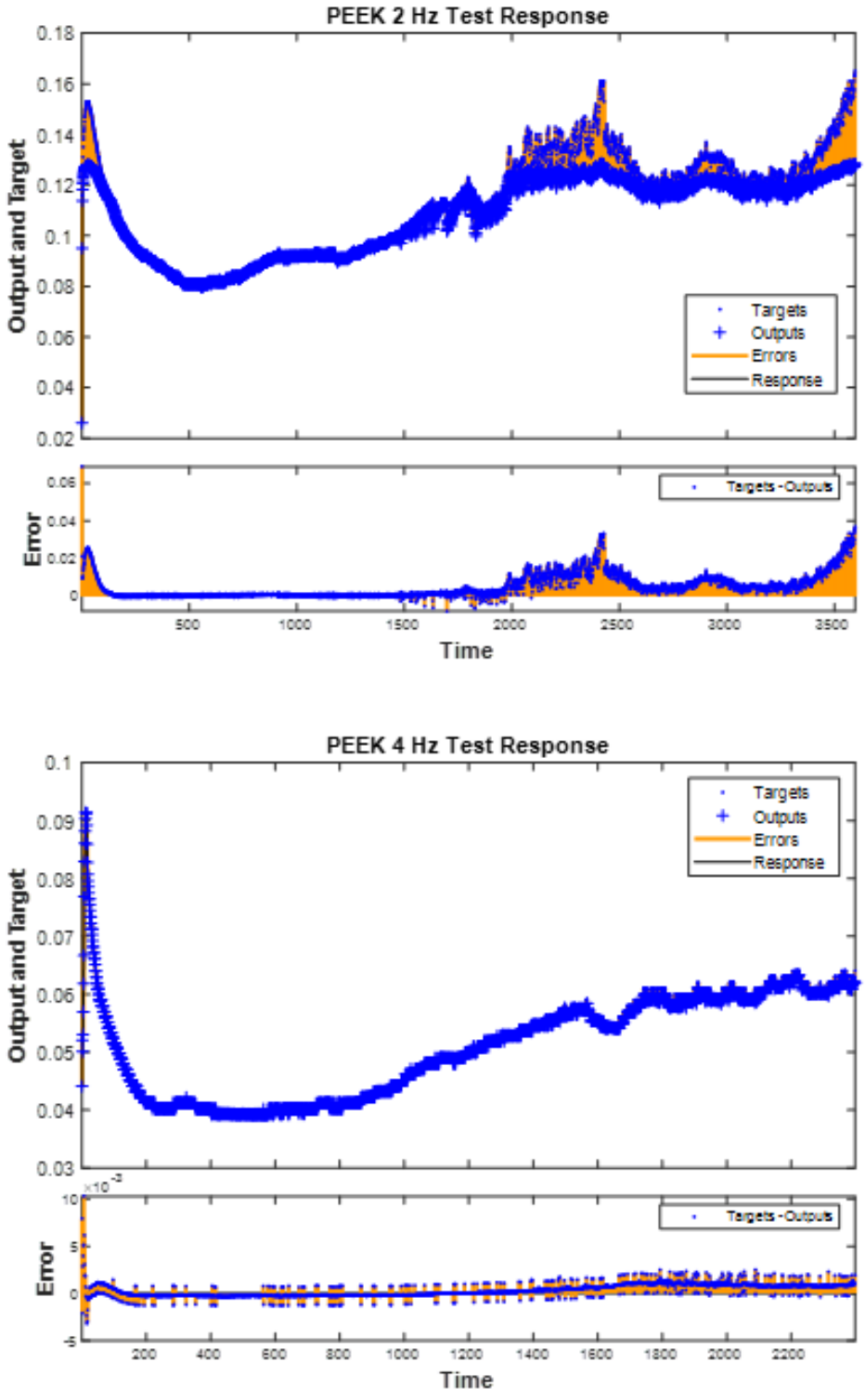


Figure 4.7: Test Response for PEEK Tests at 2 Hz (top) and 4 Hz (bottom)

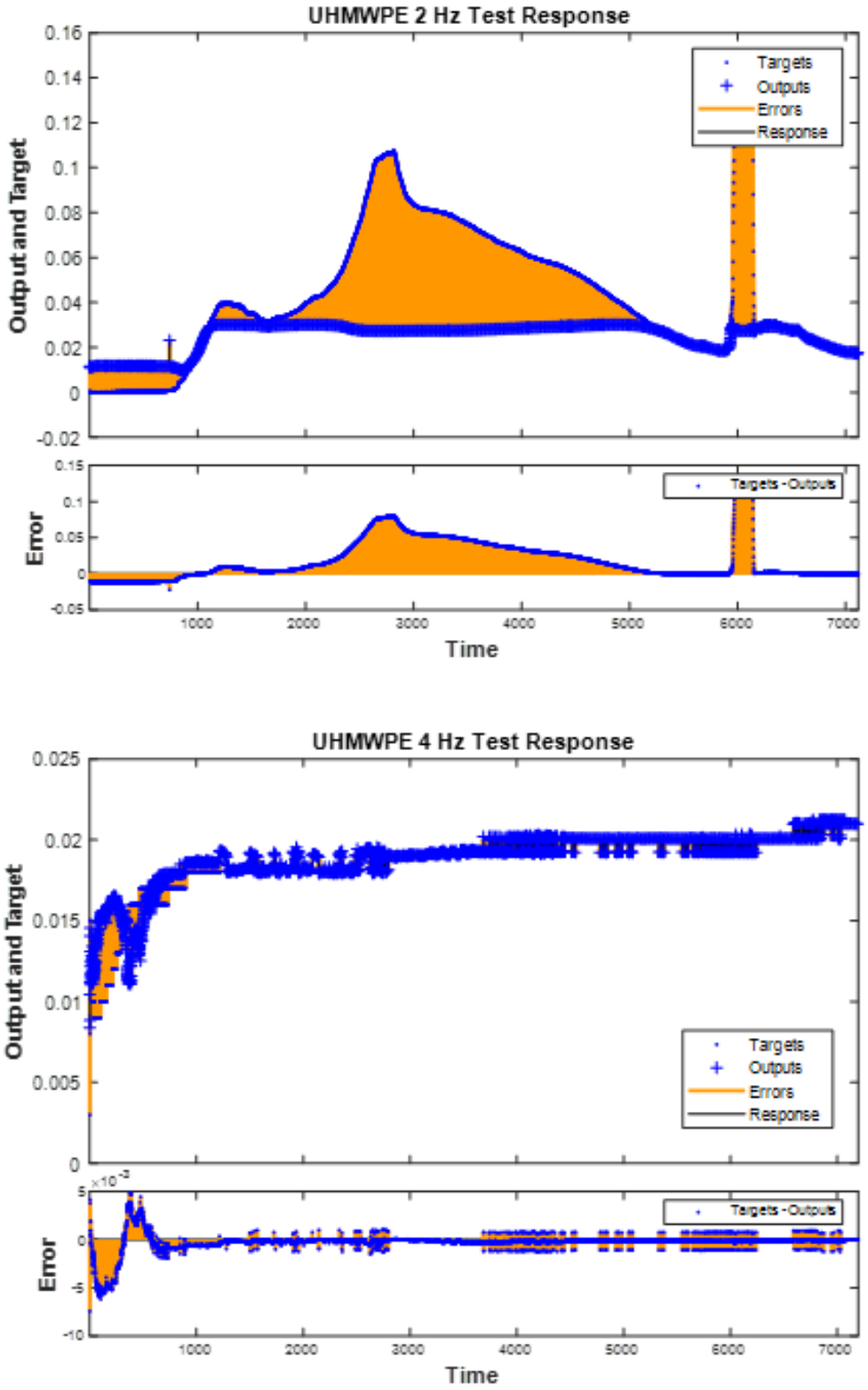


Figure 4.8: Test Response for UHMWPE Tests at 2 Hz (top) and 4 Hz (bottom)

The shape of the predicted CoF curve displays similar characteristics to the true curve for all tests (see Figure 4.7 and Figure 4.8). The difference between predicted and actual values is low as evidenced by the low MSE values (see Table 4.3). During prediction, it is assumed that the relationship between acquired AE data is uniform across all tests for each polymer at both test frequencies. Uniformity is not always guaranteed, as repeatability has been a limitation of AE testing (Hellier 2003). PEEK and UHMWPE have different physical and acoustic properties, such as Young's modulus and attenuation coefficients affecting the characteristics of AE signals acquired from both materials. The effect of these different material properties is explained further in Section 4.3.4. The strain waves generated in the polymer (due to loading) would differ for each test, causing acquired AE data to have different characteristics, hence the difference between predicted and actual CoF. NARX neural net test result for UHMWPE at 2 Hz exhibited the smallest R^2 value (Table 4.3), and is reflected in Figure 4.8, where the UHMWPE 2 Hz test curve has the most error between predicted and actual CoF values. There is a likelihood that the effect of the material and acoustic properties, in addition to the repeatability issues of AE testing, is more significant in the UHMWPE tests hence the low R^2 values obtained compared to the PEEK tests. Despite these limitations, there is still a close similarity between the predicted and actual CoF response curves (in addition to the R^2 value of about 75 % obtained for UHMWPE 4 Hz test predictions), thereby supporting the hypothesis that time-dependent AE features predict the coefficient of friction profile of a metal-on-polymer joint articulation surface during *in vitro* testing. This result has promising implications for the potential use of AE testing to evaluate the frictional (and, by extension, wear) behaviour of an artificial joint bearing surface *in vivo* when it will be impossible to carry out friction and wear tests.

4.3.2. AE parameters at different test stages

The bio-tribo-acoustic tests can be split into two stages as shown in Figure 4.3. At the onset of sliding, CoF and RMS values are seen to first reach an initial maximum followed by a minimum. This initial increase is due to the collision of surface asperities at the onset of sliding also known as the running-in stage (Boness and McBride 1991). The source of AE signals during running-in is the energy released during the initial collision and subsequent fracture of asperities in addition to the material deformation and crack formation that occurs within the contact region (Belyi, Kholodilov, and Sviridyonok 1981). Stage 2 is where prolonged sliding occurs as proven by the continued variation in CoF and RMS values observed. These variations are less severe than during running-in, indicative of a small change in friction over a large time span. Continued plastic deformation and crack propagation as sliding progresses is the predominant source of AE at this stage. There is also the ploughing action of deformed asperities (for the UHMWPE specimens) and the presence of wear particles entrapped within the contacting surfaces (for the PEEK specimens) (Jiaa and Dornfeld 1990), hence, the variation in underlying tribological processes and the recorded CoF and RMS values.

To better understand the distribution of AE signals across the two test stages, k-means clustering was used to cluster and categorise AE signals into three clusters. The choice of three clusters was affirmed by the silhouette index plot (Figure 4.9), where most of the hits in all three clusters had silhouette values greater than 70% proving that they belong to those clusters. It was assumed that the misclassified hits (hits with negative silhouette values) do not significantly affect the clustering results due to their small number. Cluster assignments for both polymeric materials at 2 Hz and 4 Hz are shown using the duration vs amplitude plots in Figure 4.10. After post-processing of AE signals, the UHMWPE 2 Hz tests had less hits than the other tests hence why its clusters are not as densely packed as the other tests.

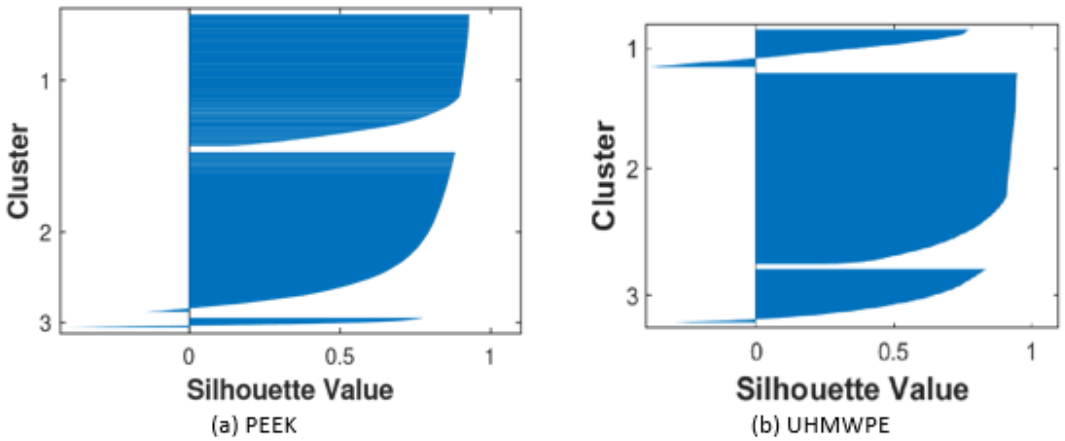


Figure 4.9: Silhouette plot showing silhouette values of hits in each cluster

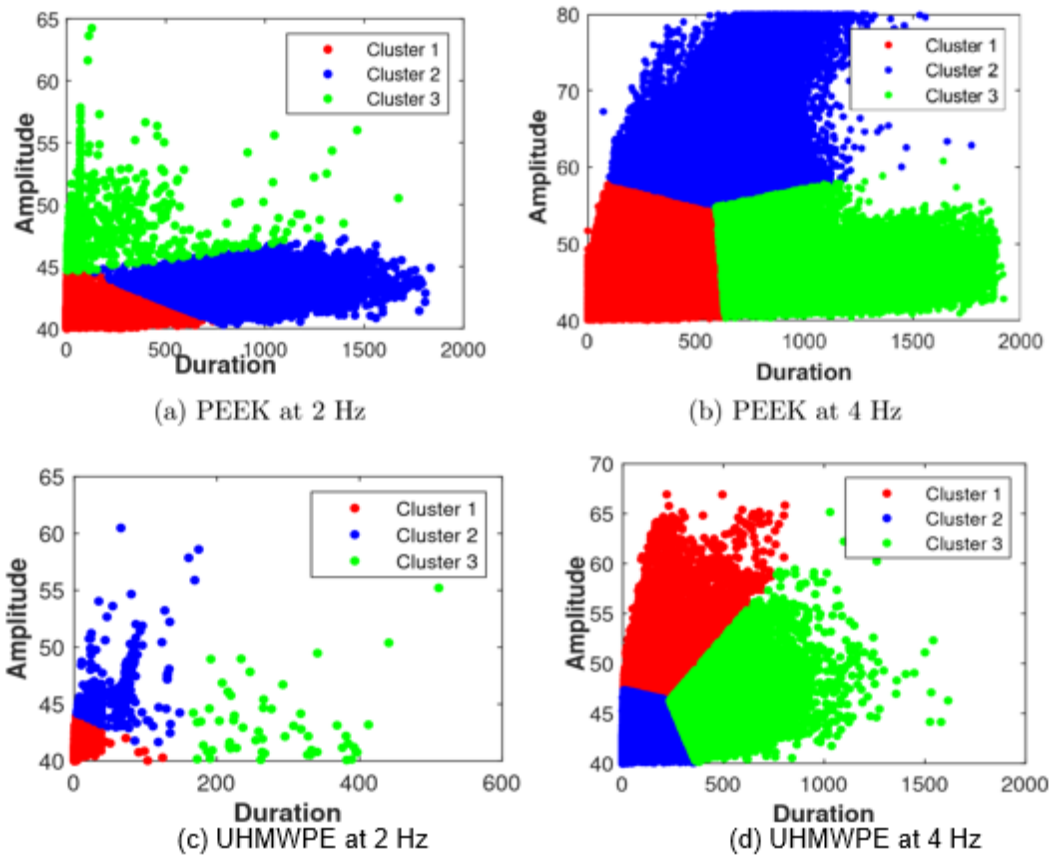


Figure 4.10: Cluster Assignments

Each cluster has similar properties for both materials at the two test frequencies. Based on observation of the cluster properties in conjunction with the features of the three emission types, the clusters are labelled as follows:

- Burst Emission - Clusters with high amplitude and short duration hits. These correspond to cluster 3 for the PEEK 2 Hz tests, cluster 2 for the PEEK 4 Hz tests, cluster 2 for the UHMWPE 2 Hz tests and cluster 1 for the UHMWPE 4 Hz tests.
- Continuous Emission - Clusters with low amplitude and long duration hits. These correspond to cluster 2 in PEEK 2 Hz tests and cluster 3 in PEEK 4 Hz, UHMWPE 2 Hz & UHMWPE 4 Hz tests.
- Mixed Emission - Clusters with mid-range amplitude and short to medium duration hits. These correspond to cluster 1 in PEEK 2 Hz, PEEK 4 Hz & UHMWPE 2 Hz tests and cluster 2 in UHMWPE 4 Hz tests.

In addition to duration and amplitude, the peak frequency values also exhibit significant differences for each emission type as shown in Table 4.4 to Table 4.7 where the mean value and 95% confidence interval (C.I.) of the three AE features are presented.

Table 4.4: Mean and 95% confidence intervals (C.I.) of AE features in each cluster for PEEK at 2 Hz

AE Feature		**Burst**		**Mixed**		**Continuous**
	Mean	C.I.	Mean	C.I.	Mean	C.I.
Duration, μs	208.73	187.92, 231.52	112.10	108.24, 115.27	901.20	896.32, 905.91
Amplitude, dB	48.22	48.05, 48.50	41.06	41.04, 41.08	43.11	43.09, 43.13
Peak Frequency, kHz	260.31	258.05, 262.21	222.82	221.71, 223.88	218.96	217.98, 219.89

Table 4.5: Mean and 95% confidence intervals (C.I.) of AE features in each cluster for PEEK at 4 Hz

AE Feature		**Burst**		**Mixed**		**Continuous**
	Mean	C.I.	Mean	C.I.	Mean	C.I.
Duration, μs	679.61	676.57, 682.22	100.99	100.20, 101.76	1134.9	1132.37, 1137.16
Amplitude, dB	67.83	67.73, 67.92	44.87	44.85, 44.89	45.69	45.66, 45.72
Peak Frequency, kHz	238.58	238.14, 239.00	246.28	246.15, 246.40	240.23	240.00, 240.46

Table 4.6: Mean and 95% confidence intervals (C.I.) of AE features in each cluster for UHMWPE at 2 Hz

AE Feature		**Burst**		**Mixed**		**Continuous**
	Mean	C.I.	Mean	C.I.	Mean	C.I.
Duration, μs	50.78	48.71, 52.91	3.88	3.62, 4.18	282.47	263.05, 306.03
Amplitude, dB	45.84	45.71, 46.00	41.15	41.11, 41.20	43.55	42.87, 44.51
Peak Frequency, kHz	231.45	230.14, 232.59	210.96	209.53, 212.56	218.99	209.02, 229.22

Table 4.7: Mean and 95% confidence intervals (C.I.) of AE features in each cluster for UHMWPE at 4 Hz

AE Feature		**Burst**		**Mixed**		**Continuous**
	Mean	C.I.	Mean	C.I.	Mean	C.I.
Duration, μs	154.68	151.66, 157.79	42.94	42.33, 43.63	503.98	500.81, 506.92
Amplitude, dB	51.24	51.15, 51.33	42.63	42.61, 42.65	46.04	45.99, 46.09
Peak Frequency, kHz	219.34	218.69, 220.03	214.49	214.04, 214.89	229.07	228.79, 229.32

Figure 4.11 shows the distribution of AE signal types across the two test stages for PEEK and UHMWPE at both test frequencies. All emission types are present in both test stages but in different proportions. Most of the burst emissions are present in stage I (i.e., running-in)

of the PEEK 2 Hz tests (Figure 4.11a), whilst a higher percentage of the continuous emissions are present in stage II (i.e., prolonged sliding) for all tests. Burst emissions are known to be generated due to damage formation such as asperity fractures (Asamene and Sundaresan 2012; Unnporsson 2013). Since micro-crack formations characterise the running-in stage due to contact and fracture of asperities at the onset of sliding which would cause an instantaneous release of high strain energy, it is no surprise that a higher percentage of the burst emissions are present during this stage of the PEEK 2 Hz tests. The force required to separate adhered particles or junctions as sliding progresses is another source of burst emissions during running-in.

Unlike the PEEK 2 Hz tests, a higher percentage of burst emissions were generated during the prolonged sliding stage not the running-in stage as expected. This results from ploughing causing either deformed asperities during UHMWPE/CoCrMo tests (Liu, Xiang, and Li 2004) or the generation of wear debris during PEEK/CoCrMo tests (Siskey et al. 2016). Both wear processes cause the release of sudden, instantaneous strain energy, thereby generating burst emissions. The tribological processes' definition softens when mixed emissions are measured, where both burst and continuous emissions cannot always be isolated. What is most common is for them to coincide, thereby generating mixed emissions.

The commonality between all tests is the higher percentage of the continuous emissions that are present in the prolonged sliding stage. Continuous emissions are generated when multiple signals overlap making them indistinguishable and the envelope of the signal amplitudes becomes constant (Asamene and Sundaresan 2012; Unnporsson 2013). The source of these emissions are predominantly steady friction and plastic deformation of surface asperities mostly during prolonged sliding.

The two test stages can be related to the everyday motion of the natural and artificial joints with the running-in stage representative of the static friction expected at the start of a

movement (e.g., starting to stand), during which more significant amount of force would be exerted whilst the prolonged sliding stage is representative of the period of the actual motion (e.g., walking).

These findings suggest that the occurrence of frictional events and plastic deformation that generates continuous emissions can be vital for the condition monitoring of the tribological behaviour of two contacting surfaces during sliding. The distribution of continuous emissions can be used to identify test stages and, by extension the most dominant surface damage occurring.

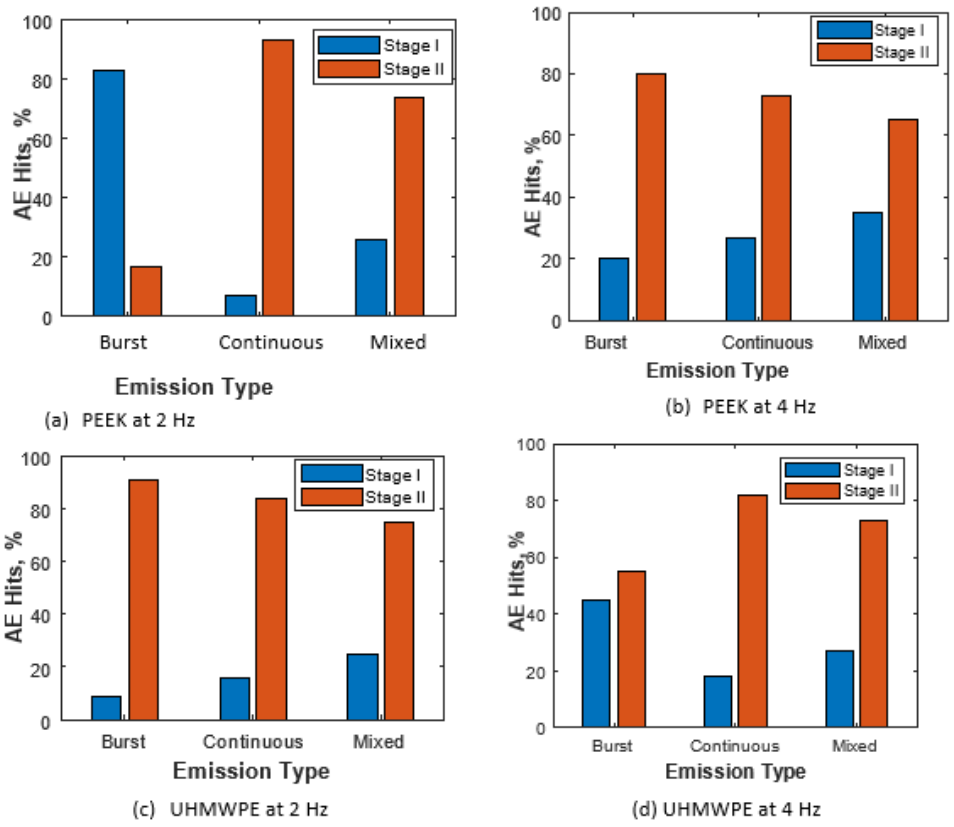


Figure 4.11: Distribution of AE signals across all test stages for (a) PEEK at 2 Hz, (b) PEEK at 4 Hz, (c) UHMWPE at 2 Hz and (d) UHMWPE at 4 Hz. The generation of more continuous emissions during prolonged sliding (stage II) is common to all tests.

4.3.3. Acoustic Emission and Wear

The wear mechanism experienced by the test specimens is predominantly adhesive with a polymer sliding against a smooth metal surface (Belyi, Kholodilov, and Sviridyonok 1981). This was confirmed by wear scar imaging (see Figure 4.12) whereby the peak regions (regions with high z value) presenting in the centre of the contact zone, indicate where adhesion has taken place. Studies have shown that the frequency of AE signals can also be used to infer wear behaviour (Ferrer et al. 2010). It has been reported that the AE frequency spectra due to sliding friction have a frequency band of 50 to 250 kHz (Ferrer et al. 2010) and the majority of hits from these tests had peak frequencies in the range of 150 to 275 kHz (Figure 4.13), proving that sliding friction is the most dominant source of detected AE hits. This is further evidenced by the fact that continuous emission hits, known to be mostly due to sliding friction, have average peak frequencies in the range of 220 to 240 kHz (see Table 4.4 to Table 4.7).

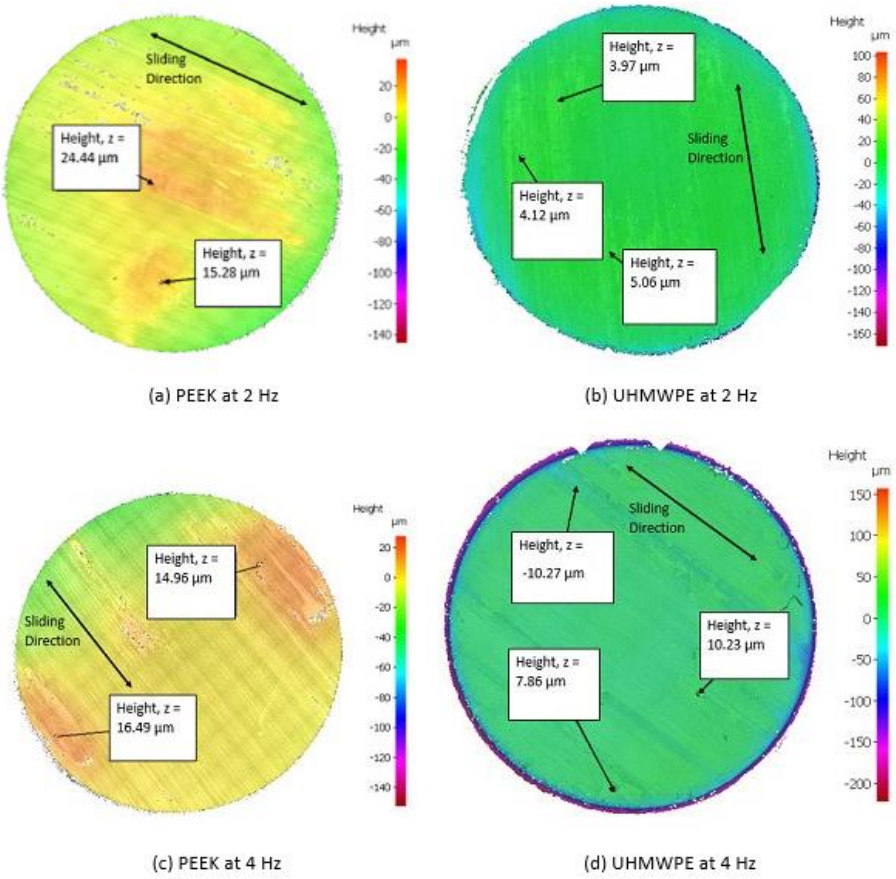


Figure 4.12: Polymeric specimens wear scar images for (a) PEEK at 2 Hz, (b) UHMWPE at 2 Hz, (c) PEEK at 4 Hz and (d) UHMWPE at 4 Hz. The regions of high z value indicate adhesion of displaced wear particles back onto the surface.

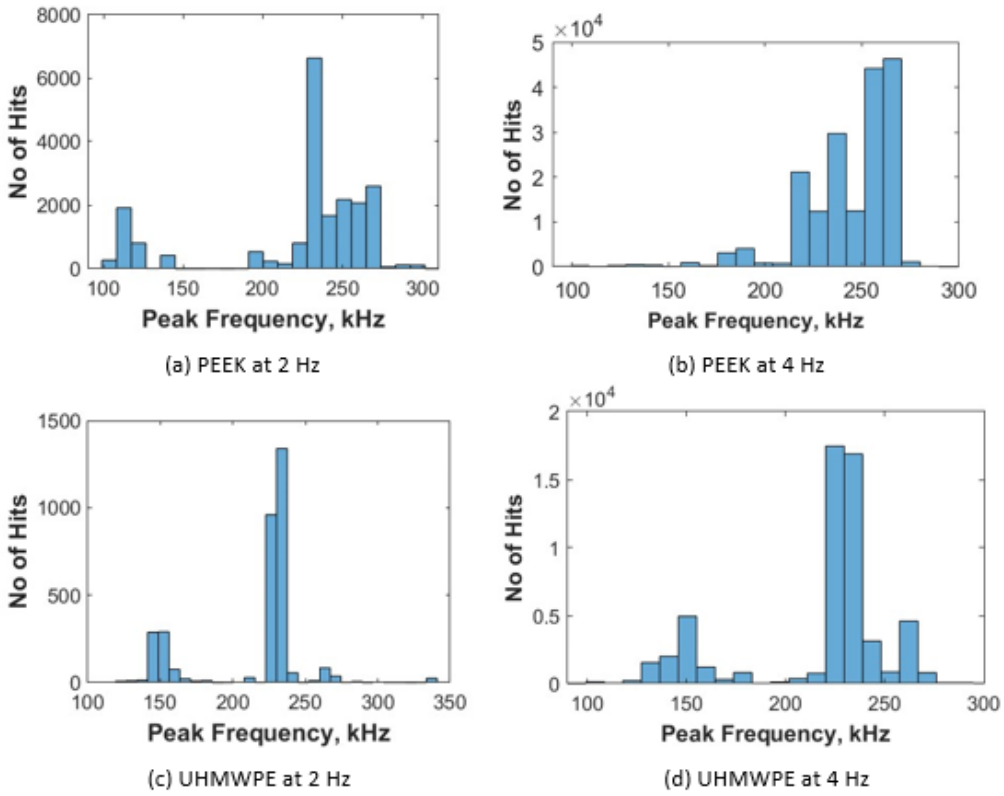


Figure 4.13: Peak Frequency Distribution Plots for (a) PEEK at 2 Hz, (b) PEEK at 4 Hz, (c) UHMWPE at 2 Hz and (d) UHMWPE at 4 Hz. Peak frequencies of the AE hits are mostly in the range 150 to 275 kHz.

4.3.4. Effect of differing material properties

It was observed that the PEEK tests produced more AE hits with higher intensities than the UHMWPE tests. This can be attributed to their differing mechanical properties, which govern their sliding characteristics.

PEEK has a higher Young's modulus and ultimate tensile strength compared to UHMWPE making it more resistant to deflection and plastic deformation. With the same load applied, it would require more force to break PEEK asperities compared to those on the UHMWPE surface. Hence, the generation of more strain energy at the contact zone, thereby producing AE hits with higher intensity as well as more hits overall. This is in line with the study by Belyi et al. (Belyi, Kholodilov, and Sviridyonok 1981), where it was found that AE intensity is higher for polymers with a higher Young's modulus.

Other factors that could explain the differences are the attenuation coefficient and speed of sound in both materials. Although the attenuation coefficient of PEEK (0.38 dB/mm) is higher than that of UHMWPE (0.24 dB/mm), sound moves faster in PEEK (2555 m/s) than in UHMWPE (1950 m/s) (Cadot, Saillant, and Dulmet 2016). The low speed of sound coupled with the higher deformability of UHMWPE could lead to more loss of signal levels as it travels from the source to the sensor, which explains why AE signals from UHMWPE have less intensity than AE signals from PEEK.

It is worth noting that despite the difference in intensity levels, AE signals from both materials behave similarly showing that differing material properties only affect the intensity of the signal. AE signals can be used to infer bio-tribological characteristics irrespective of the material type.

4.4. Conclusion

This chapter has shown that with the use of tribo-acoustic testing, AE signal features can be used to interpret frictional processes in a simulated metal-on-polymer joint articulation during *in vitro* testing. The following conclusions were reached.

- The shape of the coefficient of friction profile can be predicted using a NARX neural network with time-dependent AE features.
- The test stages can be distinguished using the percentage of continuous emissions generated, and this can be further used to infer the evolution of surface damage.
- Although AE intensity differed between the two polymeric specimens, the quality of the acquired AE signals was not affected. This shows that the concept would be applicable to different artificial joint articulations irrespective of their bearing surfaces.

Having shown how frictional properties of simulated articulating joint surfaces can be interpreted using AE features, the next chapter will show how AE signals can be classified using machine learning techniques.

Chapter 5

Diagnostic Capabilities of AE Testing – a fundamental study

The content of this chapter is based on the following publication:

Olorunlambe, Khadijat A., Zhe Hua, Duncan E. T. Shepherd, and Karl D. Dearn. 2021. "Towards a Diagnostic Tool for Diagnosing Joint Pathologies: Supervised Learning of Acoustic Emission Signals." *Sensors* 21 (23): 8091. https://doi.org/10.3390/s21238091.

5.1 Introduction

The previous chapter has shown the capability of AE testing as a tool for monitoring the tribological conditions of artificial joints further emphasising the potential of the AE technique in diagnosing joint pathologies such as osteoarthritis and artificial joint failure. To achieve this, a technique for differentiating AE signals based on wear mechanism must be developed. This is a challenging problem if AE signals are only analysed using parameter and frequency-based techniques. However, recent studies (Gutkin et al. 2011; Qiao, Weng, and Li 2019; Yao, Li, and Yuan 1999) have shown that machine learning techniques can be advantageous in the analysis of AE signals for damage recognition and identification of failure modes.

This chapter focuses on the application of machine learning techniques in the classification of AE signals according to adhesive and abrasive wear mechanisms initiated under controlled joint conditions using supervised and unsupervised learning methods.

5.2 Materials and Methods

5.2.1. Experimental Parameters

AE signals were acquired using the tribo-acoustic test set-up described in Chapter 2. Tests were performed using poly-ether-ether-ketone (PEEK) rods (6 mm diameter, 16 mm length – supplied by Direct Plastics, Sheffield, UK) as the reciprocating specimen and steel plates (38 mm by 33 mm by 4 mm) as the fixed lower specimen. For the abrasive wear tests, the steel plates were roughened with a belt sander fitted with P40 grade sandpaper.

Hertzian contact mechanics (see Section 3.2.2) were used to calculate the initial test load. First, the maximum contact pressure in a ball-and-socket Charite Lumbar Spinal Implant with 10 mm ball radius and 0.35 mm radial clearance (Moghadas et al. 2015) was calculated using the loading and displacement conditions defined by BS ISO 18192-1 for wear of total

intervertebral spinal disc prostheses (BS ISO 18192-1 2011). This was then used to calculate an initial equivalent TE77 load for a line contact using Equation 5.1. To simulate worst-case conditions, the calculated load was increased to 150 N. Final test parameters are given in Table 5.1.

AE signals were acquired at a sampling frequency of 2 MHz throughout the tests. After post-processing using the NOESIS Advanced AE analysis software, selected AE features were collated and loaded into MATLAB (R.2019a) for analysis using pattern recognition techniques from the machine learning toolbox and outlined in Section 5.2.2 below.

$$F = \frac{p_0^2 \pi l E^*}{2R^*}$$

Equation 5.1: TE77 Load calculation for a line contact where l is length of the cylindrical rod.

Table 5.1: TE77 and AE Acquisition Parameters

	PARAMETERS	VALUE
TE77 TEST PARAMETERS	Load	150 N
	Frequency	2 Hz
	Stroke	12.5 mm
	Test Duration	1 hour
	Lubricant	Dry (to induce adhesive wear) and Ringer's Solution (to induce abrasive wear)
AE ACQUISITION PARAMETERS	Threshold	40 dB
	Pre-amplifier Gain	60 dB
	Band Pass Filter	100 – 600 kHz
	Sampling Rate	2 MHz

5.2.2. Pattern Recognition Technique

The portfolio of pattern recognition techniques used to drive the learning and matching between recorded AE data and the information stored in the training database is outlined below.

Feature Selection and Extraction

AE features for clustering and classification were selected using hierarchical link clustering and principal component analysis (PCA). AE waveforms were defined using discrete parameters/features to allow pattern recognition techniques based on multi-parameter statistical analysis. However, post-processing of the AE signals produced 50+ AE features. Therefore, the input AE features were minimised to increase the speed and accuracy of classification. The feature selection method used is described in (Momon et al. 2012). First, the ten most common features used in previous studies were selected (Gutkin et al. 2011; Momon et al. 2012; Ech-Choudany et al. 2017) and then normalised to give a weighting using Equation 5.2. Next, the correlation matrix was calculated and subjected to complete link-hierarchical clustering. Eight features with Pearson correlation coefficients in the range $[-0.7: 0.7]$ were then selected. The eight selected features are shown in Table 5.2.

$$X' = \frac{X - \mu}{\sigma}$$

Equation 5.2: Equation for normalising the discrete AE features. μ represents the mean value of the descriptor and σ the standard deviation.

Table 5.2: Selected AE features after hierarchical clustering. Definitions can be found in Section 3.2.1.

No	AE FEATURES
1	Amplitude
2	Duration
3	Counts to Peak
4	RA Value
5	Average Frequency
6	Peak Frequency
7	AE Root Mean Square
8	Absolute Energy

Before supervised classification of the AE signals, PCA was used to determine new uncorrelated features by linear combinations of the features selected after hierarchical clustering. The principal components were chosen using the process outlined below:

Normalised AE features were composed into a matrix X of dimensions n by m (Equation 5.3):

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

Equation 5.3: Matrix, X, of normalised AE features.

The eigenvalues of the covariance matrix and the corresponding eigenvectors were then obtained (Equation 5.4 and Equation 5.5).

$$R = Cov(X) = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1k} \\ r_{21} & r_{22} & \cdots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k1} & r_{k2} & \cdots & r_{kk} \end{bmatrix} = \begin{bmatrix} 1 & r_{12} & \dots & r_{1k} \\ r_{21} & 1 & \cdots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k1} & r_{k2} & \cdots & 1 \end{bmatrix}$$

Equation 5.4: Covariance of Matrix X

$$\text{where } r_{ij} = \frac{\sum_{k=1}^{n}(X_{ki}-\bar{X}_i)(X_{kj}-\bar{X}_j)}{\sqrt{\sum_{k=1}^{n}(X_{ki}-\bar{X}_i)^2}\sqrt{\sum_{k=1}^{n}(X_{kj}-\bar{X}_j)^2}}$$

Equation 5.5: Equation for calculating covariance

Next, the eigenvectors corresponding to the larger p eigenvalues were used to form the projection matrix A.

P-eigenvalues, with a cumulative contribution rate of 95%, were selected for classification purposes.

The new matrix 'Y' after the reduction of dimension was obtained (Equation 5.6):

$$Y = XA$$

Equation 5.6: New matrix, Y.

Unsupervised Learning: K-means Clustering

The features listed in Table 5.2 were clustered using the k-means clustering method, which clusters data by minimising the sum of squared Euclidean distances from all cluster vectors to their centre. The algorithm can be found in Chapter 3 (see Section 3.2.5.)

Supervised Learning

AE data from all adhesive and abrasive wear tests were merged to create a library of labelled data for classification using supervised methods. First, hit vectors after a steady state was reached were selected. The merged dataset was then randomly split into two: training and test data at the ratio of 85% to 15%. Three classification models were employed, running each one 20 times. The performance of each model was evaluated using the average classification accuracy, the ratio of correctly predicted cases to a total number of cases and the average F-score (Equation 5.9), calculated from the precision (the ratio of the number of correct positive predictions to the total number of positive predictions, Equation 5.7) and recall (ratio of correct positive predictions to the number of actual positive cases, Equation 5.8), viz (Rousseeuw 1987):

$$Precision, P = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Equation 5.7: Formula for precision.

$$Recall, R = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

Equation 5.8: Formula for recall.

$$F - score = \frac{2PR}{P + R}$$

Equation 5.9: Formula for f-score.

Logistic Regression Classifier

Logistic regression, the most common method used in binary classification problems, employs a sigmoid function to compute the probability of a class as a function of the linear combination of multiple variables (Kleinbaum and Klein 2002). The sigmoid function allows mapping real numbers into binary form (0 and 1) and is represented by Equation 5.10 -Equation 5.12:

$$g(z) = \frac{1}{1 + e^{-z}}$$

Equation 5.10: Sigmoid function

Where z is an index that combines all the features of X.

$$z = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

Equation 5.11: z index calculation

Where α and β are unknown constant parameter. Hence, the logistic regression model can be written as:

$$P(\boldsymbol{X}) = \frac{1}{1 + e^{-(\alpha + \sum \beta_i X_i)}}$$

Equation 5.12: Logistic regression model

K-Nearest Neighbours (KNN) Classifier

The KNN classifier categorises unknown vectors based on their distance to their nearest neighbours in the training dataset, the distance measured being the weighted squared inverse Euclidean distance between the training set and the unknown vector. The classifier finds the nearest neighbours to the unknown vector and specifies the class with the most representation among those nearest neighbours as the predicted class (Momon et al. 2012). The optimum k-number was determined by training with different values and computing the average classification rate. A 5-fold cross-validation method measured classification rates, and the

optimum k-number maximised classification rate whilst minimising training time. An illustration of how the KNN classifier works is shown in Figure 5.1 below.

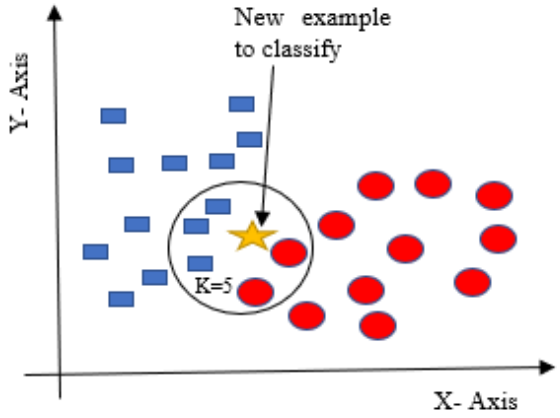


Figure 5.1: Illustration of the KNN Classifier

Neural Network Classifier

Lastly, backpropagation (BP) neural networks, computational architectures modelled after the brain's architecture (Curry and Rumelhart 1990), were used in the final classification process. The BP network selected was a multi-layer feed-forward neural network with the onward transmission of features and backpropagation of errors characteristics of this network. A three-layer pattern recognition feed-forward network comprising one input layer, one hidden layer (10 neurons) and one output layer was used in training (Figure 5.2). The weights and bias values were updated using the Bayesian Regularisation backpropagation training function. This training function uses the Levenberg-Marquardt algorithm to optimise the weights and bias (Foresee and Hagan 1997). The performance of each iteration was evaluated using a mean square error calculation.

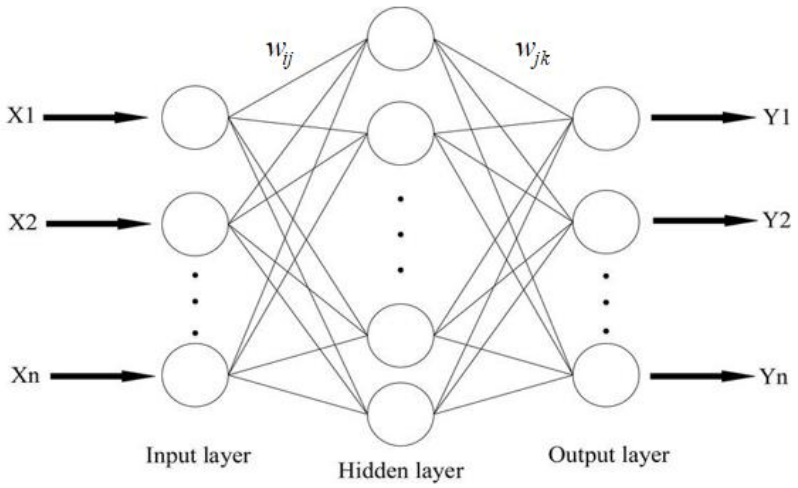


Figure 5.2: Illustration of a three-layer backpropagation neural network. The hidden layer has 10 neurons.

5.3. Results and Discussion

5.3.1. AE hits and wear mechanisms

There was a clear distinction between the recorded AE signals for the two wear mechanisms. Hits from the abrasive wear tests were approximately eight times more than those generated during the adhesive tests. The source of these hits is the tribological processes that each specimen is exposed to during testing. In addition to the friction profiles, microscopic examination of worn surfaces provides evidence of several underlying wear mechanisms present only in both tribological tests. These include micro-crack formation and deformation during sliding, and in abrasion, there is evidence of scouring and scratching, resulting in the generation of PEEK wear particles. Further analysis of the frictional data revealed that a quarter of the total AE hits were detected rapidly at the start of the adhesive test. In contrast, it took much longer before 25% of hits were detected (Figure 5.3) in the abrasive tests.

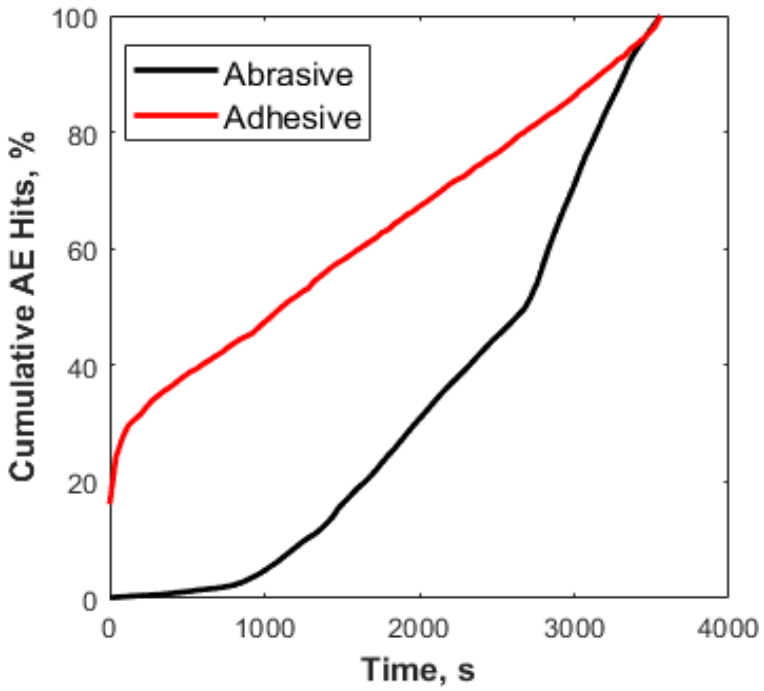


Figure 5.3:Cumulative AE hits vs time plot for adhesive (red) and abrasive (black) wear tests

Examination of the friction coefficient curves (Figure 5.4) provides a better understanding. The initial rapid rise of the friction coefficient in a short period for adhesive tests represents the initial sudden collision of asperities on contacting surfaces. This collision of asperities produces high strain energy, leading to many AE hits. The steady increase in hits correlates with the region of steady friction between contact surfaces, as shown in the friction coefficient (CoF) curve. Therefore, the plot of the cumulative hits for abrasion can also be related to the friction curve. There are three clear stages in the friction curve:

1. Running-in (initial collision of surface asperities and a slight decrease in CoF)
2. A second increase in CoF during prolonged sliding
3. Steady-state

These stages explain the three discernible sections in the plot of the cumulative hit.

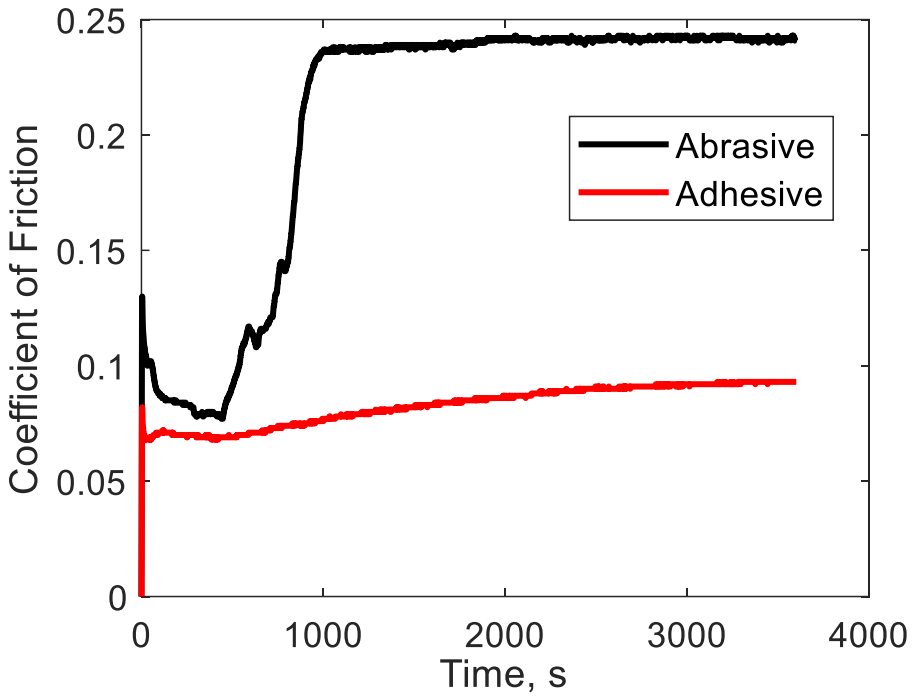
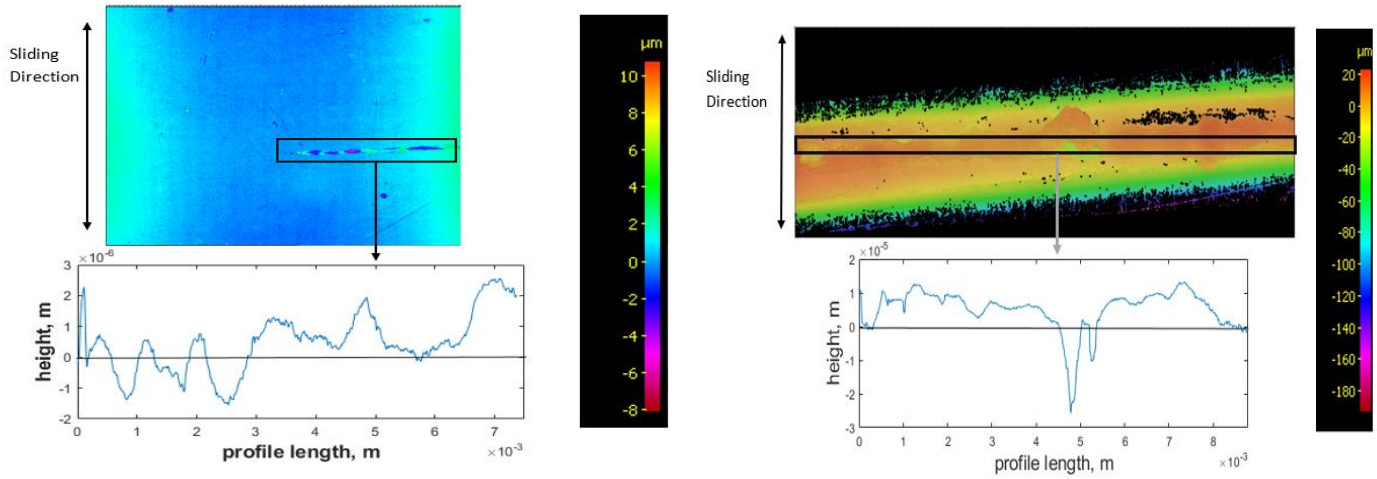


Figure 5.4:Coefficient of friction vs time curves for adhesive (red) and abrasive (black) wear tests

An Alicona Infinite focus optical microscope was used to analyse wear scars on the test specimens. Analysis of the wear scar on the steel plate used for the adhesive test (Figure 5.5a) shows regions of increase in height with a corresponding loss of height on the PEEK wear scar surfaces (Figure 5.5b), indicative of wear particles separating from the PEEK rod and adhering to the metal surface thereby confirming adhesive wear mechanism.



(a) (b)

Figure 5.5: Adhesive test specimens, contour image & corresponding graphical profiles of worn region after testing for (a) steel plate & (b) PEEK rods. Regions of high height shows wear particle transfer from PEEK unto steel plate indicating adhesion.

The wear scar on the steel plate for the abrasive test (Figure 5.6a) shows the grooves formed by the pre-test conditioning. The constant low topographical height (around $0\ \mu m$) observed in the PEEK and steel's graphical profiles (Figures 5.6a & 5.6b) indicates that the severity of the asperities has reduced due to material transfer proving abrasive wear has occurred. The smooth surface of the PEEK wear scar also implies a complete breakaway of material.

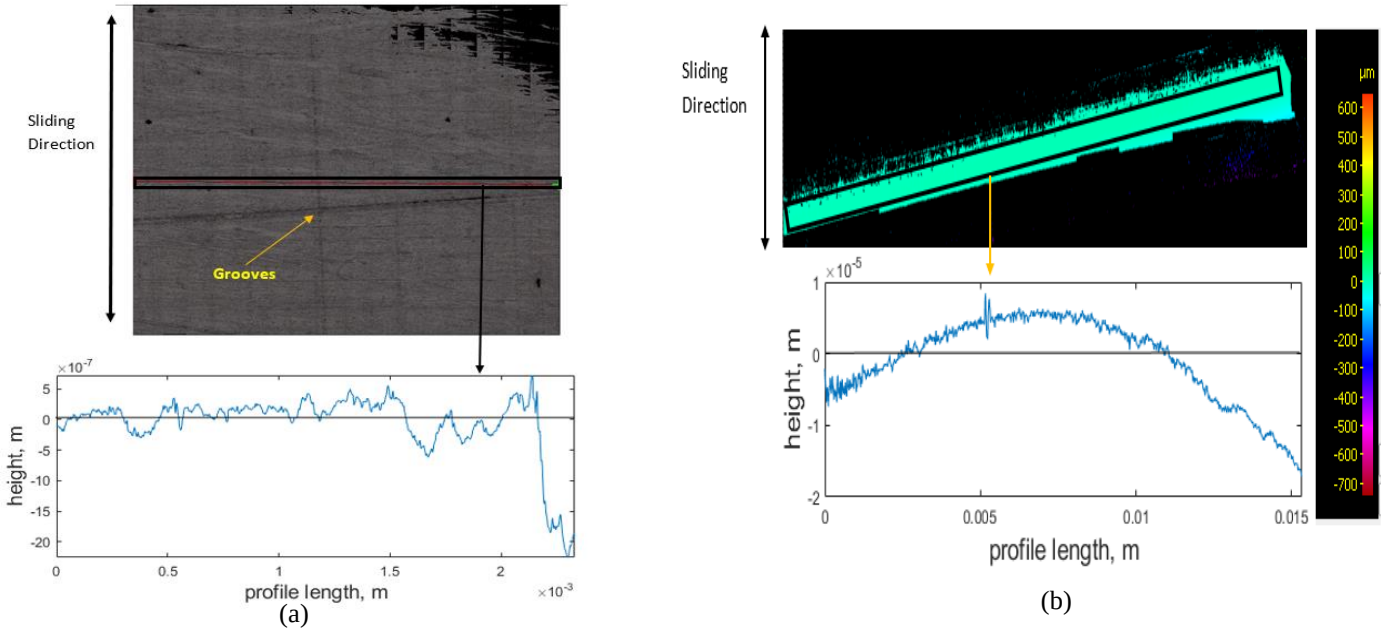


Figure 5.6: Abrasive test specimens, contour image & corresponding graphical profiles of the worn regions after testing for (a) steel plate & (b) PEEK rods. Low topographical height shows material has broken away indicating abrasive wear mechanism.

5.3.2. K-means clustering

Adhesive Wear

According to the Silhouette Index, optimal clustering was obtained with 2 clusters for the adhesive wear tests. Feature correlation plots are given in Figure 5.7 for duration vs amplitude and hits (events) vs amplitude and show minimum overlapping between clusters. The average values of five AE features for both clusters are shown in Table 5.3. Events in cluster 2 are of higher intensity than those in cluster 1, evidenced by cluster 2 having higher average values for the features except average frequency, for which cluster 1 has a broader range and a higher average. These events indicate the presence of more burst emission types (hits with high amplitude and short duration) in cluster 2 and more continuous emission types (hits with low amplitude and long duration) in cluster 1. The clustering output of all adhesive wear tests showed similar characteristics. A higher percentage of AE events were assigned to cluster 1 (~96%) than cluster 2 (~4%). This increase is expected as many of the cluster 2 events

were generated towards the end of the test. The raw waveforms of a sample event from both clusters can be seen in Figure 5.8. It shows how events from the two clusters differ (see Table 5.4 for feature data).

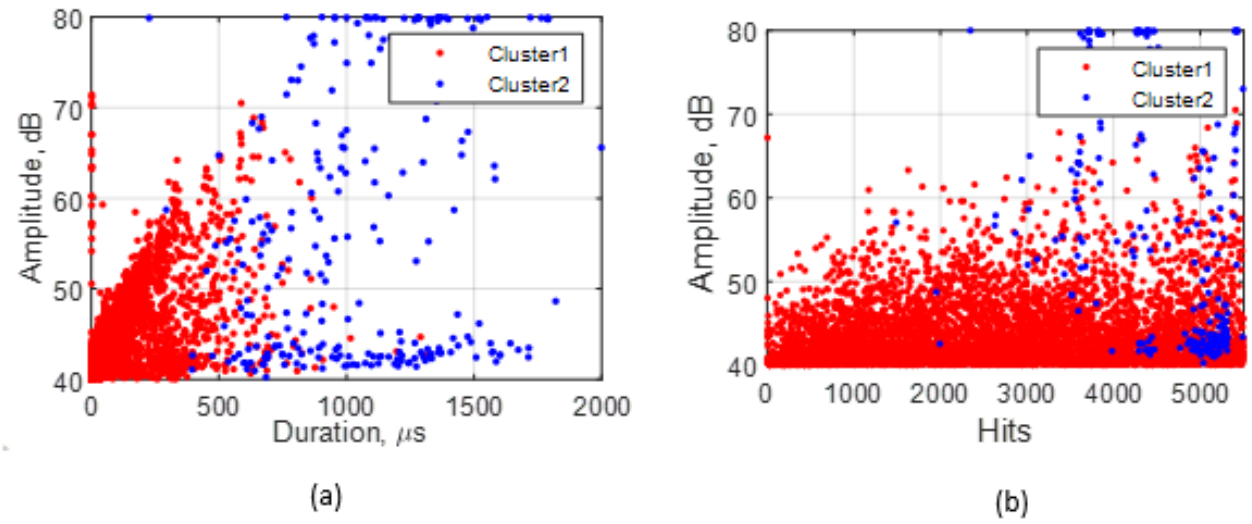


Figure 5.7: Clustering solutions for adhesive wear test showing (a) duration vs amplitude and (b) hits vs amplitude plots.

Table 5.3: Range and mean (including standard deviation) of five features per cluster for adhesive wear tests

AE FEATURES	**CLUSTER 1**		**CLUSTER 2**	
	Range	Mean (*std*)	Range	Mean (*std*)
Amplitude, dB	40 – 71.43	43.81 (*4.30*)	40.27 – 79.99	56.70 (*14.67*)
Duration, µs	0.5 – 1,308	85.71 *(133.16*)	226.5 – 1,998.5	1,048.60 (*347.05*)
RA Value, µs/dB	0 – 12.47	0.57 (*1.27*)	0.01 – 35.25	10.08 (*7.84*)
Average Frequency, kHz	0 - 1,000	333.73 *(378.23*)	1.46 – 194.26	60.36 *(54.67*)
Absolute Energy, attoJ	0.26 – 28,610	96.48 (*648.62*)	41.29 – 687,510	59,834 (*1.20 x* 10^9)

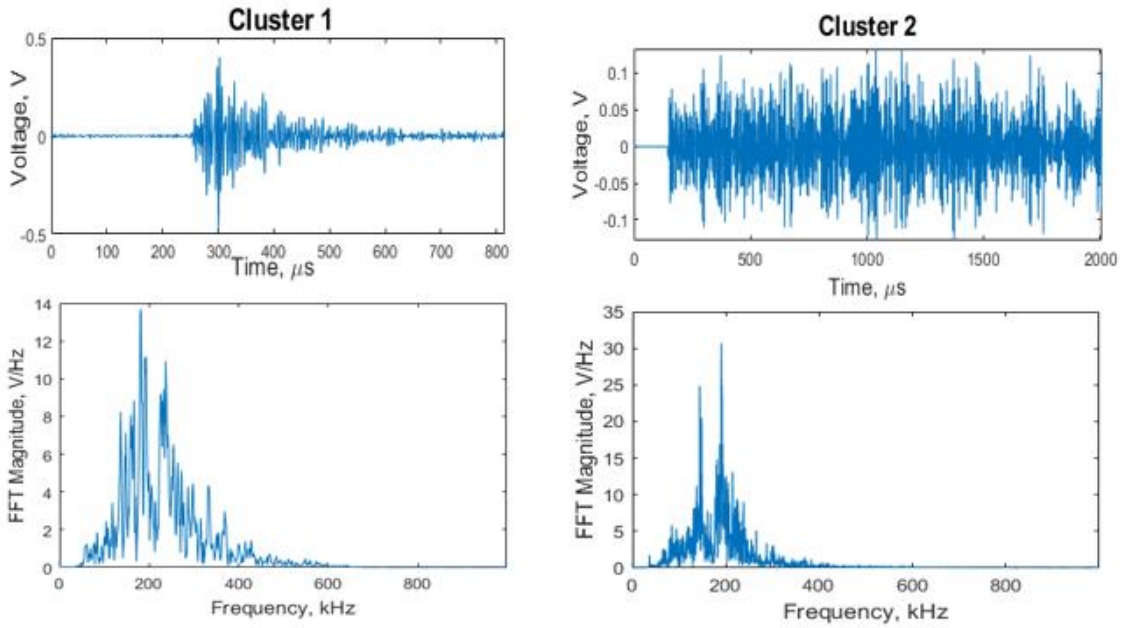


Figure 5.8: Waveforms of hits from adhesive wear test clusters in time domain (top) and frequency domain (bottom). Cluster 1 event is a burst emission while cluster 2 event is a continuous emission.

Table 5.4: Features of the sample AE waveforms (Figure 5.8) from adhesive wear tests clusters.

AE Features	**Sample AE Waveforms**	
	Cluster 1	Cluster 2
Amplitude, dB	53.47	42.48
Duration, µs	290.5	1855.5
RA Value, µs/dB	0.8	23.47
Average Frequency, kHz	141.14	75.99
Absolute Energy, attoJ	261.08	326.04

In general, wear tests tend towards three stages: running-in, steady-state and severe wear, often followed by rapid failure (Hase, Mishina, and Wada 2012; Mishina and Hase 2013). The running-in stage embodies initial contact between opposing micro-protrusions or asperities on the two contacting surfaces. As the wear process progresses, asperities gradually flatten (the so-called steady-state), and the actual contact area increases, leading to an initial rapid increase

in the amount of wear before reducing gradually (Hase, Mishina, and Wada 2012; Mishina and Hase 2013, 2019). Tests in this study were categorised into running-in and steady-state. Selected experimental conditions made it unlikely that severe wear would be reached due to the short duration of the tests (1 hour).

The most likely source of AE hits during running-in is the energy produced during asperity collisions and subsequent junction separation. Other sources could include micro-cracking and material deformation. Due to the high strain energy produced during running-in, the amplitude of the AE hits is high and of short duration (i.e., burst emissions). The steady friction experienced during steady-state should generate mostly continuous emissions (hits with low amplitude and long duration). As observed in Figure 5.4, the adhesive wear tests reached steady-state very early on in the tests. Reaching steady-state so early in the test makes it challenging to identify a clear distinction between the events recorded during running-in and steady-state. Cluster 1 events shown in Figure 5.7 were generated for much of the test duration. These AE events combine burst emission (running-in) and continuous emission (synonymous with steady-state tribological conditions). The low amplitude of the AE hits in cluster 1 indicates that most of the signals were generated after asperities have been removed or flattened and the tribological test has reached steady-state conditions. Other sources of AE events in cluster 1 include micro-cracking and fracture of surface asperity junctions.

The mean duration value of 85.71 µs is also an indication of burst emissions present from the initial running-in stage. As the sliding progresses further, the continuous contact between surface asperities and the lack of lubricating medium would cause the generation of PEEK wear particles which then adhere to the steel plate (Figure 5.5), leading to the emission of higher intensity AE events in cluster 2 compared to cluster 1. A similar observation was made by Hase et al. (Hase, Mishina, and Wada 2012), where they found that the generation of wear elements and transfer particles are the sources of burst emissions during adhesive wear

tests. The continuous adhesion process would require high strain energy, generating AE hits with high amplitude (burst emissions). As the test nears the 1-hour mark, other sources of AE hits would include steady friction between contacting surfaces and crack propagation. These are known to generate mostly continuous emissions; hence it is expected that hits generated would be mixed emissions (i.e., a combination of burst and continuous). Therefore, the duration and RA value of hits in cluster 2 are higher than in cluster 1.

Abrasive wear

Optimal clustering was also achieved with two clusters from the abrasive wear tests. Feature correlation plots (Figure 5.9) show that, unlike in adhesive wear, there is considerable overlap between clusters that can be attributed to the processes of abrasive wear, such as scratching, abrasion, and wear debris generation.

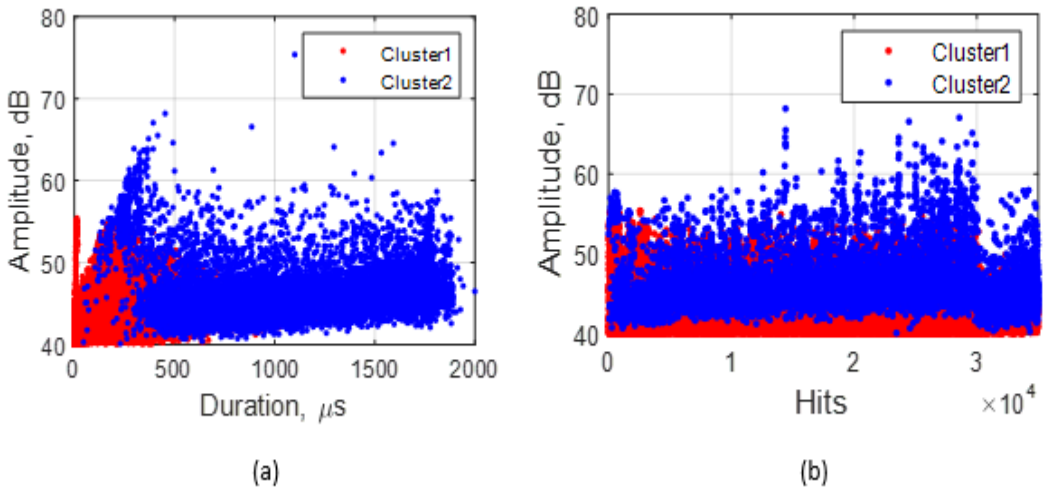


Figure 5.9: Clustering solutions for abrasive wear showing (a) duration vs amplitude and (b) hits vs amplitude plots.

All abrasive wear tests had similar clustering results (Table 5.5) to adhesive wear (Table 5.3), with AE events in cluster 2 having higher intensity than those in cluster 1. Although a higher percentage of AE events were assigned to cluster 1 than in cluster 2, the proportional difference was lower than that observed in the adhesive wear tests (~67% in cluster 1 and ~33% in cluster 2). Therefore, the average amplitude value cannot be relied upon to differentiate between the two clusters. However, the difference between the average values for the two clusters is low (43.76 dB in 1 and 46.22 dB in 2), and both clusters coincided, as shown in the hits vs amplitude plot (Figure 5.9b). The raw waveform plots in Figure 5.10 also shows how events from each cluster differ from each other (see Table 5.6 for the feature data).

Table 5.5: Range and mean (including standard deviation) values of five features per cluster for the abrasive wear test

AE FEATURES	**CLUSTER 1**		**CLUSTER 2**	
	Range	Mean (*std*)	Range	Mean (*std*)
Amplitude, dB	40.01 – 55.42	43.76 (*2.80*)	40.14 – 75.36	46.22 (*3.28*)
Duration, µs	0.5 – 1,200	116.84 (*157.93*)	45.5 – 2010	1,060 (*422.05*)
RA Value, µs/dB	0.01 – 14.42	0.87 (*1.75*)	0.01 – 42.55	13.54 (*9.40*)
Average Frequency, kHz	0 – 2000	212.22 (*298.36*)	1.57 – 193.80	27.72 (*23.44*)
Absolute Energy, attoJ	0.57 – 567.11	45.99 (*58.10*)	11.73 – 30,200	326.96 (*472.03*)

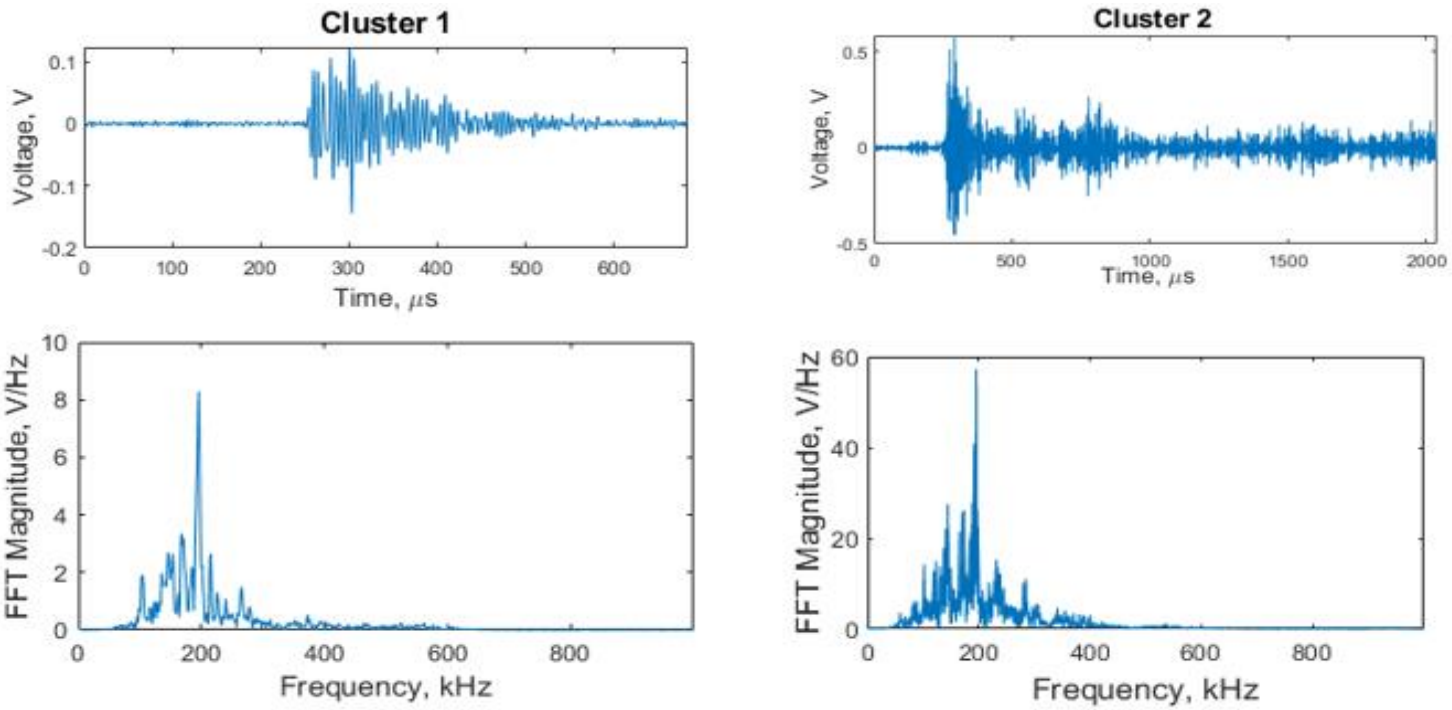


Figure 5.10: Waveform of hits from abrasive wear test clusters in time domain (top) and frequency domain (bottom). Event in cluster 1 is a burst emission while that of cluster 2 is a mixture of continuous and burst emissions.

Table 5.6: Features of the sample AE waveforms (Figure 5.10) from abrasive wear tests clusters.

AE Features	**Sample AE Waveforms**	
	Cluster 1	Cluster 2
Amplitude, dB	43.23	58.33
Duration, µs	108.5	1783.5
RA Value, µs/dB	1.03	0.64
Average Frequency, kHz	101.38	102.61
Absolute Energy, attoJ	25.55	1003.82

Abrasive wear is most commonly of two-body and three-body wear (Yang and Garrison 1989). Figure 5.4 shows that the running-in stage lasted for c.1000 seconds, during which time

there was a high number of burst emissions, i.e., events of short duration in cluster 1. The source radiated signals will include strain energy produced when surface asperities collide during running-in and the effects of third-body produced during contact (Boness and McBride 1991). Furthermore, the presence of wear debris increases micro-cutting of the polymer surface. This, in addition to continued plastic deformation and micro-crack and crack propagation (Yang and Garrison 1989), would cause more AE events with burst emission (Hase, Mishina, and Wada 2012), evidenced by the fact that cluster 1 accounts for a higher percentage of total AE events detected.

In addition to the processes mentioned previously, there is friction resulting from the two surfaces sliding against each other, another source of AE events (Belyi, Kholodilov, and Sviridyonok 1981; Asamene and Sundaresan 2012). As frictional events are known to generate continuous emissions (events with low amplitude and high duration), it can be concluded that cluster 2 events are primarily due to sliding friction between PEEK and steel. Continuous emissions would also have a long risetime hence the high RA value for events in cluster 2.

5.3.3. Supervised Learning

Hit (vectors) generated during steady-state phases were acquired for supervised classification using the three classification models discussed in Section 5.2.2. After selecting eight features using hierarchical link clustering, PCA extracted six new uncorrelated features that accounted for 95% of the variance between the eight features (Figure 5.11). After PCA, a 24,075 by six training data matrix corresponding to a 24,075 by one labelled output matrix was established (Table 5.7). Training examples for adhesive wear were labelled '0' and '1' for abrasive wear. Before training, data were randomly split into 85% for training and 15% for testing. The training data was used to train all three models and test data was used to evaluate

each model's performance on untrained data. Each classifier was used to train and test 20 times to examine the model robustness.

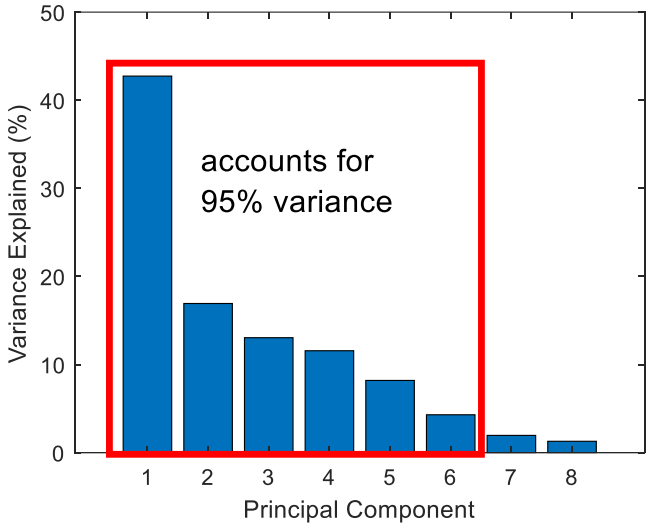


Figure 5.11: Principal Component Analysis Output. The principal components that accounted for 95% variance were selected for training.

Table 5.7: Extract of the data for supervised learning after PCA. Column seven is the known classification. 0 is for adhesive wear and 1 is for abrasive wear.

No	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6	Class
1	-0.5254	0.1712	1.5911	-0.0436	0.0650	-0.0137	**1**
2	-0.2013	1.0705	-0.1417	0.1585	0.0888	-1.5526	**0**
3	-0.2176	-0.2863	-1.1095	-0.7217	1.2744	-0.0381	**1**
4	-0.3134	-0.4454	-2.1773	-0.8948	1.6665	0.2197	**1**
5	-0.9043	0.1371	0.9589	0.0163	0.0178	0.4183	**1**
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
24071	1.7407	0.3037	0.2463	-1.7920	-1.4622	-0.6788	**0**
24072	-0.3685	0.1461	2.0166	0.0383	-0.0839	-0.0863	**1**
24073	0.0750	-0.9116	-1.1962	-0.5369	1.0159	0.1729	**1**
24074	-1.2619	0.9060	-1.8599	1.1055	0.0243	-0.5239	**0**
24075	-1.4791	0.9883	-0.9241	1.3757	-0.4782	-0.2947	**0**

The average classification accuracy and F-score for each classifier are presented in Table 5.8. Since F-scores are calculated assuming that classes are either positive or negative, two F-scores were calculated: adhesive wear as the positive class and abrasive wear as the positive class. An F-score of 1 is indicative of a perfect recall and precision.

Table 5.8: Summary of Classifiers' Performance

CLASSIFIER	**Average Training Accuracy (±std)**	**Average Test Accuracy (±std)**	**Average F-score (±std)**	
			Adhesive is positive	Abrasive is positive
Logistic Regression	0.73 (±*0.0025*)	0.72 (±*0.0048*)	0.66 (±*0.0096*)	0.77 (±*0.0036*)
Weighted k-Nearest Neighbours	0.96 (±*0.0007*)	0.97 (±*0.0026*)	0.96 (±*0.0029*)	0.97 (±*0.0025*)
Backpropagation Neural Network	0.98 (±*0.0014*)	0.98 (±*0.0024*)	0.98 (±*0.0024*)	0.98 (±*0.0023*)

All three classifiers have 70% and above accuracy, with logistic regression being the least accurate at 73% (training) and 72% (test) accuracy. However, both the KNN and BP neural network classifiers perform better, with BP neural network (98%) slightly outperforming KNN (97%) accuracy. The logistic regression model assumed a simple linear relationship between features, not recognising non-linear patterns hence the high misclassification rate compared to the other two classifiers. By classifying unknown data using distance to its five nearest neighbours, in the case of KNN, all patterns between features are always considered resulting in an optimised classification. In the case of the BP neural network classifier, using a training function to optimise the weights and bias values and minimise the squared errors help train a generalised and optimised model.

The classification accuracy achieved is comparable with other studies that have utilized machine learning algorithms in the analysis of AE signals. Qiao et al. (Qiao, Weng, and Li 2019) obtained a 93% average classification accuracy when using BP neural network to classify 8YSZ thermal barrier coatings according to the indentation failure experienced. In their use of

KNN to classify AE signals acquired during the fatigue testing of carbon fibre composites, Momon et al. (Momon et al. 2012) obtained an above 90% classification accuracy. McCrory et al. (McCrory et al. 2015) also found that using an artificial neural network can be advantageous for the classification of carbon fibre composites according to different damage mechanisms.

All three models perform better at recognising abrasive wear features as indicated by the F-scores (higher when abrasive wear is a positive class) and the confusion matrix plots (Figure 5.12). Wear processes such as micro-crack, plastic deformation and sliding friction are common to adhesive and abrasive wear mechanisms. AE events from these processes are likely to have similar feature characteristics hence the misclassification. The clustering results further emphasise these similarities between feature characteristics, where overlap can be found in AE feature values in cluster 2 for both wear mechanisms. There are several ways in which the misclassification rate can be further minimised.

- Increase the number of neurons in the hidden layer for the BP neural network (Yu et al. 2019).
- Increase features used for classification (Ray 2015) by choosing more features or deriving new features through the mathematical combination of original features. More features can help improve the model by recognising more patterns between features and improving the classification accuracy.
- Increase the number of training examples. Having more examples would help the model learn better, thereby improving classification accuracy.

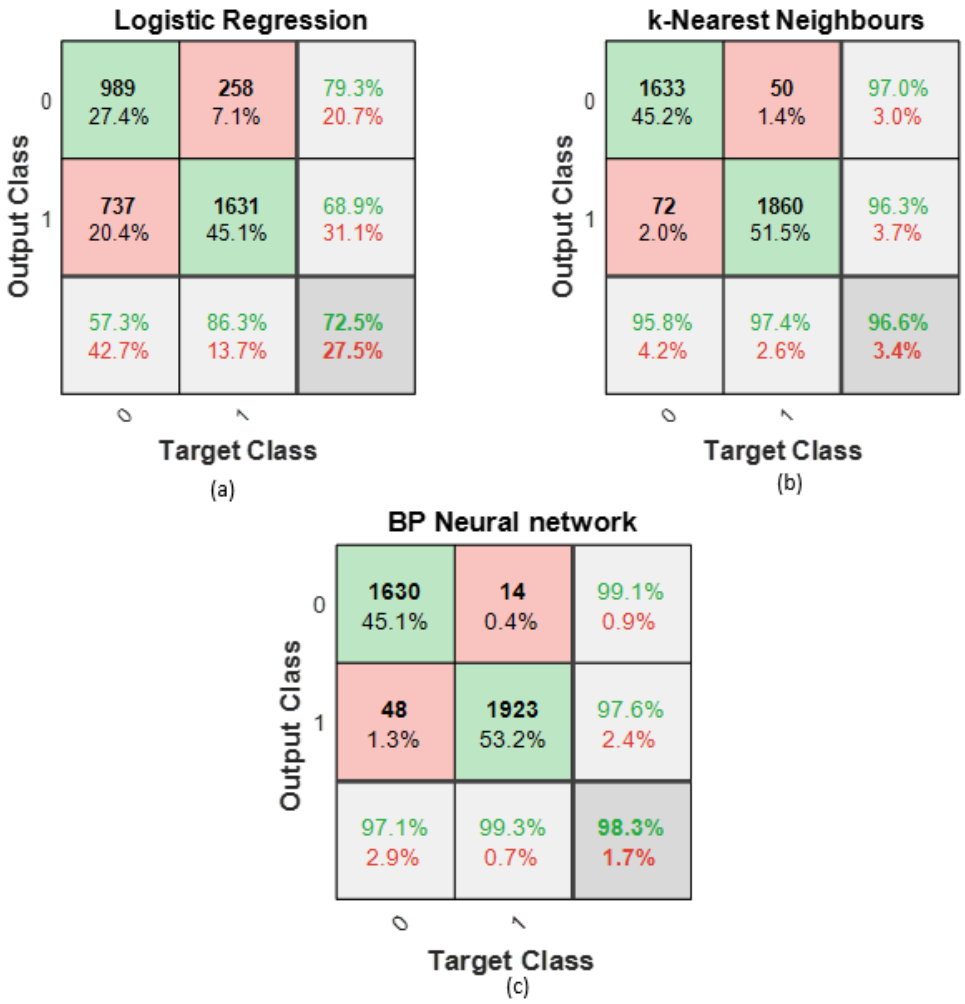


Figure 5.12:Confusion matrix for all three classifiers

KNN classifiers do not learn from the training data. Instead, training data is used to classify test data. Although the KNN classifier has a high classification accuracy, increasing training examples and/or features would increase the time required for new data to be classified. Also, finding the optimal k-value can be time-consuming and considering that the optimal k-value will be based on input features, the process would have to be repeated every time training data is updated (Lorena et al. 2011). The BP neural network has an advantage over the KNN classifier as it is unnecessary to learn the details of functions used in hidden layers. It is relatively simple to train with new data. Also, since the neural network learns from the training data, it is acceptable for training time to be high since test time is unlikely to be high (Crivelli, Guagliano, and Monici 2014). The ever-changing nature of AE signals would require a

classification model that is simple and easy to manipulate, making the neural network classifier the best choice.

5.4. Conclusions

In this chapter, pattern recognition techniques have been used to classify AE signals from tribological tests acquired under simulated artificial joint articulation and the following conclusions were reached.

- K-means clustering can be used to group AE hits detected during a tribo-acoustic test based on different tribological processes. This can be used to diagnose different stages of tribo-acoustic tests based on similarities and differences between AE features.
- Discrete features of AE signals acquired under controlled joint conditions have hidden relationships that can be identified using classification models such as KNN and BP neural networks.

A limitation of using traditional ANNs like the BP neural network is the complex feature engineering required to build a robust model. The next chapter will show how this limitation can be eliminated by using deep transfer learning of continuous wavelet transforms of AE signals for the classification of AE signals from a simulated joint articulating surface.

Chapter 6

Deep Transfer Learning of AE signals- using AE to identify the damage mechanism of a simulated artificial knee joint articulation

6.1 Introduction

The previous chapter has shown that using machine learning techniques, AE signals from biotribological contacts can be differentiated according to the wear mechanism involved. The BP neural network, a type of artificial neural network (ANN), emerged as the most intuitive machine learning technique.

One of the limitations of traditional machine learning techniques is the complex feature engineering required to build a classification model that will be robust to different conditions which can be computationally expensive and time-consuming. This is more significant in AE signal analysis, where the hits generated can be in the order of 1000s and potential features can be in the order of 100s depending on the duration and type of test. However, recent studies on the use of AE for damage recognition in civil engineering applications have shown that deep transfer learning and wavelet analysis can be used to classify AE signals from fracture processes and for wear monitoring in sliding bearing systems (Xin et al. 2020; Ren and Chen 2021; König et al. 2021).

This chapter focuses on the use of deep transfer learning and wavelet transforms to classify AE signals from a simulated artificial knee joint articulation based on two damage modes: burnishing (adhesive) and scratching (abrasive) wear. These two damage modes have been identified in several retrieved polyethylene (PE) inserts of failed knee implants as the cause of wear of the PE insert (Puloski et al. 2001; Engh et al. 2009; Grecu et al. 2016).

6.2 Materials and Methods

6.2.1. Experimental Parameters

AE signals were acquired using the tribo-acoustic test set-up described in Chapter 3. Tests were performed using ultra-high-molecular-weight polyethylene (UHMWPE) discs (supplied by PCS Instruments Ltd, London, UK) as the reciprocating specimen and medical

grade cobalt chromium molybdenum alloy (CoCrMo) (supplied by Phoenix Tribology, Newbury, UK and machined to a surface finish of less than 0.02 μm R_a) as the fixed specimen. The UHMWPE discs were of diameter 10 mm and 3 mm thickness and machined to a surface finish of 0.65 ± 0.17 μm R_a (measured with the Alicona InfiniteFocusG5 Optical 3D Measurement System from Alicona Imaging Gmbh), respectively.

Test conditions representing the linear motions of hinged knees were determined using the loading conditions stated in annex A1 of ASTM F732-17, the test method for linear reciprocating wear motion applications (ASTM International 2017). The recommended contact pressure of 3.54 MPa was doubled in order to simulate severe wear damage. The final test parameters, including AE acquisition parameters, are summarised in Table 6.1.

A burnishing wear mechanism was simulated simply by having a smooth UHMWPE sliding on a CoCrMo plate. To simulate the scratching wear mechanism, 45 mg of 80 grit size Silicon Carbide grinding grit was added to the contacting surface between the UHMWPE disc and the CoCrMo plate before testing. Five tests were carried out for each damage mechanism to test the reproducibility of the results. It is expected that damage mechanisms *in vivo* would be a combination of 2 or more identified mechanisms with overlap between characteristics of the damage processes and to represent this, a third set of tests were carried out to simulate progression from burnishing to scratching wear. After 20 minutes under burnishing test conditions, the test was paused and silicon carbide grinding grit was added to simulate scratching wear. The tests then ran for another 10 minutes before ending.

Table 6.1: TE 77 and AE Acquisition Parameters

	PARAMETERS	VALUE
TE77 TEST PARAMETERS	Load	556 N
	Frequency	5 Hz
	Stroke	12.4 mm
	Duration	Burnishing – 30 mins Scratching – 15 mins
	Lubricant	Ringer's Solution
AE ACQUISITION PARAMETERS	Threshold	35 dB
	Pre-amplifier Gain	60 dB
	Waveform streaming capture interval	Every 300 secs for 5 secs
	Sampling Rate	2 MHz
	Band Pass Filter	100 – 600 kHz

6.2.2. Analysis of Variance (ANOVA)

To check for any significant differences between discrete AE feature means of hits from tests under both damage mechanisms, a one-way analysis of variance was performed. Two thousand AE hits during the period after running-in were extracted for all tests and then averaged for each damage mechanism. The discrete features compared have been identified from previous tests (see Chapter 5) as significant AE features. A 95% confidence interval (CI) was employed and it was assumed that all variances were equal. This was used to find out which discrete AE feature is sensitive to different damage mechanisms.

6.2.3. Convolutional Neural Network and Wavelet Analysis

AE waveforms were classified using pretrained convolutional neural networks (CNN) and continuous wavelet transform (CWT) images. The proposed approach is outlined in Figure 6.1.

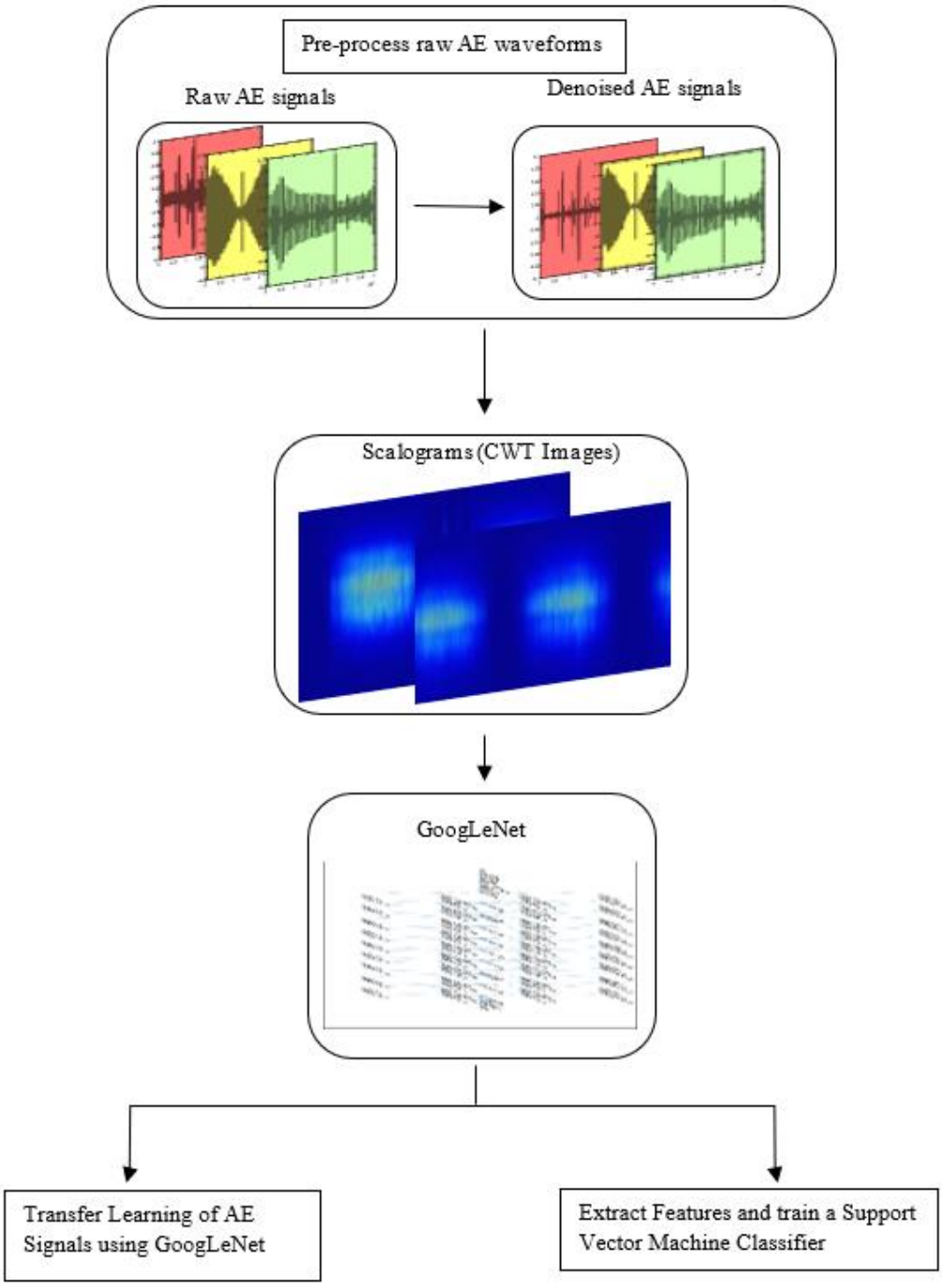


Figure 6.1: Illustration of classification of AE signals using a pre-trained CNN and CWT images

Extraction of Data for Training and Testing

The raw time series AE waveforms captured during testing were used for the purpose of classifying AE data based on the simulated damage mechanisms. For the individual simulated damage mechanism tests, 2 raw waveforms of 5 seconds duration each were selected

from each test repeat for both burnishing and scratching. Considering how damage evolves during testing, the waveforms were chosen from two different test stages- after running-in (after 1200 s for burnishing and 600 s for scratching tests) and just before the test ended (after 1800 s for burnishing tests and after 900 s for scratching tests). Each 5 second waveform was further subdivided into waveforms of 200 ms duration, each representing the time taken for one complete cycle during testing.

To make the classification models robust and generalisable to differing damage conditions and levels of damage, raw AE waveforms from the damage progression tests were also added to the data for training and testing. Just as for the individual damage tests, 10 s duration waveforms were selected from each damage stage (burnishing – after 480 s and 1020 s; scratching – after 1500 s and 1800 s), making a total of 20 s waveform from each test. The waveforms were also subdivided into smaller waveforms of 200 ms duration each.

All extracted waveforms were collated, labelled accordingly and then split into training and testing data at a ratio of 70:30. This resulted in a training data matrix of 672 by 400,000 and a test data matrix of dimension 288 by 400,000. Training and test data were then denoised and converted to CWT images before classification as outlined in the following sections.

Pre-processing of AE signals

Raw AE signals were denoised before the continuous wavelet transform process to ensure that features learnt during classification are not contaminated by the presence of noise using wavelet denoising. Wavelets are functions that are used to divide a signal into different frequency components for more in-depth analysis of the original signal (Polat and Siraç 2018). The denoising process is summarised as follows (Polat and Siraç 2018; Ai et al. 2018):

- Wavelet decomposition of a signal into different components.

- Calculation of thresholds. The thresholds were selected based on a mixture of Stein's Unbiased Risk Estimate (SURE) principle and the Fixed Form threshold.
- Reconstruction of signal with the coefficients after denoising.

To determine the best wavelet to use for the denoising process, different wavelets (at 3 decomposition levels) were used to denoise a sample raw AE signal, and their performance was evaluated using the signal-to-noise ratio (SNR) (Equation 6.1) and the mean square error (MSE) (Equation 6.2) of the variance between the original signal and the denoised signal (Ai et al. 2018). The best performing denoising process would maximise SNR whilst minimising the MSE value. This was found to be the sym6 wavelet with 4 decomposition levels (See Table 6.2)

$$SNR = 10\ log\left(\frac{\sum_{n=1}^{N} f(n)^2}{\sum_{n=1}^{N}\left[f(n) - \hat{f}(n)\right]^2}\right)$$

Equation 6.1: Formula for computing SNR where $f(n)$ is the original signal, $\hat{f}(n)$ is the denoised signal and N is the length of the signal.

$$MSE = \sqrt{\frac{\sum_{n=1}^{N}\left[f(n) - \hat{f}(n)\right]^2}{N}}$$

Equation 6.2: Formula for computing MSE where f(n) is the original signal, f̂(n) is the denoised signal and N is the length of the signal.

Table 6.2: The signal-to-noise ratio and mean square error of different wavelets. The wavelet with the best denoising effect is shown in the red font.

Wavelet/ Decomposition Levels	**4**		**6**		**8**	
	SNR	MSE	SNR	MSE	SNR	MSE
db6	9.5140	0.2313	9.5129	0.2314	9.5129	0.2314
db8	9.7912	0.2241	9.7909	0.2241	9.7909	0.2241
db10	9.1730	0.2406	9.1729	0.2406	9.1729	0.2406
bior2.6	9.1631	0.2409	9.1591	0.2410	9.1591	0.2410
coif4	9.4683	0.2326	9.4681	0.2326	9.4681	0.2326
coif5	9.4725	0.2324	9.4724	0.2324	9.4724	0.2324
sym6	**9.8435**	**0.2227**	**9.8425**	**0.2228**	**9.8425**	**0.2228**
sym7	9.3565	0.2356	9.3561	0.2356	9.3561	0.2356
sym8	9.7514	0.2251	9.7512	0.2251	9.7512	0.2251

Continuous Wavelet Transform of denoised signals

Continuous wavelet transforms (CWT) have proven useful for in-depth analysis of raw AE signals by decomposing them into their frequency components whilst keeping the time localisation (Asamene and Sundaresan 2012). CWT's integral is shown in Equation 6.3 (Polat and Siraç 2018). By changing the scale parameter and the position parameter, CWT allows us to shift and stretch/compress the mother wavelet thereby providing us with the frequency and time information simultaneously (Polat and Siraç 2018).

$$CWT = \int_{-\infty}^{\infty} \left[x(t) \frac{1}{a} \psi^* \left(\frac{t-b}{a} \right) \right] dt$$

Equation 6.3: CWT Integral where x(t) is the original signal, $\psi(t)$ is the analysing function (wavelet), a is a scale parameter and b is a position in time.

The Morlet wavelet (EQ) was chosen as the analysing wavelet due to its capability of extracting features with equal variance in time and frequency. This makes the time-frequency resolution adaptable to different signals thereby guaranteeing the extraction of temporal features (Zhao et al. 2019). The extracted time-frequency features are represented as scalograms, in the form of RGB images, which are then used for training and testing the CNN models. The time-frequency representations were created using the Signal Processing toolbox in MATLAB version 2021a. The script can be found in the appendix.

$$\psi(t) = e^{-\frac{t^2}{2}} \cos(5t)$$

Equation 6.4: Mathematical representation of the Morlet wavelet

Transfer Learning using Convolutional Neural Network

A CNN is a deep learning technique that is an extension of artificial neural networks (ANNs) which are computational architectures modelled after the brain's architecture (Curry and Rumelhart 1990). Just like ANNs, CNNs have an input layer, hidden layer and an output layer. The difference is that CNNs can be used for image classification unlike ANNs.

CNN consists of several hidden layers which includes convolutional (CONV) layers with activation function (ReLu) and pooling (POOL) layers followed by an optional fully connected (FC) layer (Xin et al. 2020). CNNs can have hundreds of layers, making them very accurate, more robust and adaptable (Xin et al. 2020). Training a new CNN can be very computationally expensive and time-consuming, hence, transfer learning was employed in this study. Deep transfer learning is when a pre-trained CNN is modified and used to train new images thereby saving time and computing power. The pre-trained network chosen for this study is GoogLeNet. GoogLeNet is a 22-layer deep network that has been trained on images in the ImageNet database and classified into 1000 categories (Szegedy et al. 2015).

The images fed into the modified GoogLeNet for retraining are the RGB cwt scalograms of the denoised AE waveforms. Before training on the RGB scalograms, the last 4 layers of the network were modified to fit the new training and response data (Figure 6.2). The RGB scalograms were also resized to [224 224], the required size for inputs into the network. All training and test images were stored in an image datastore on MATLAB. The training data was further split at a ratio of 80:20 to create the network training and validation datasets. The full training options are presented in Table 6.3. The MATLAB script used for CNN training and testing can be found in the appendix.

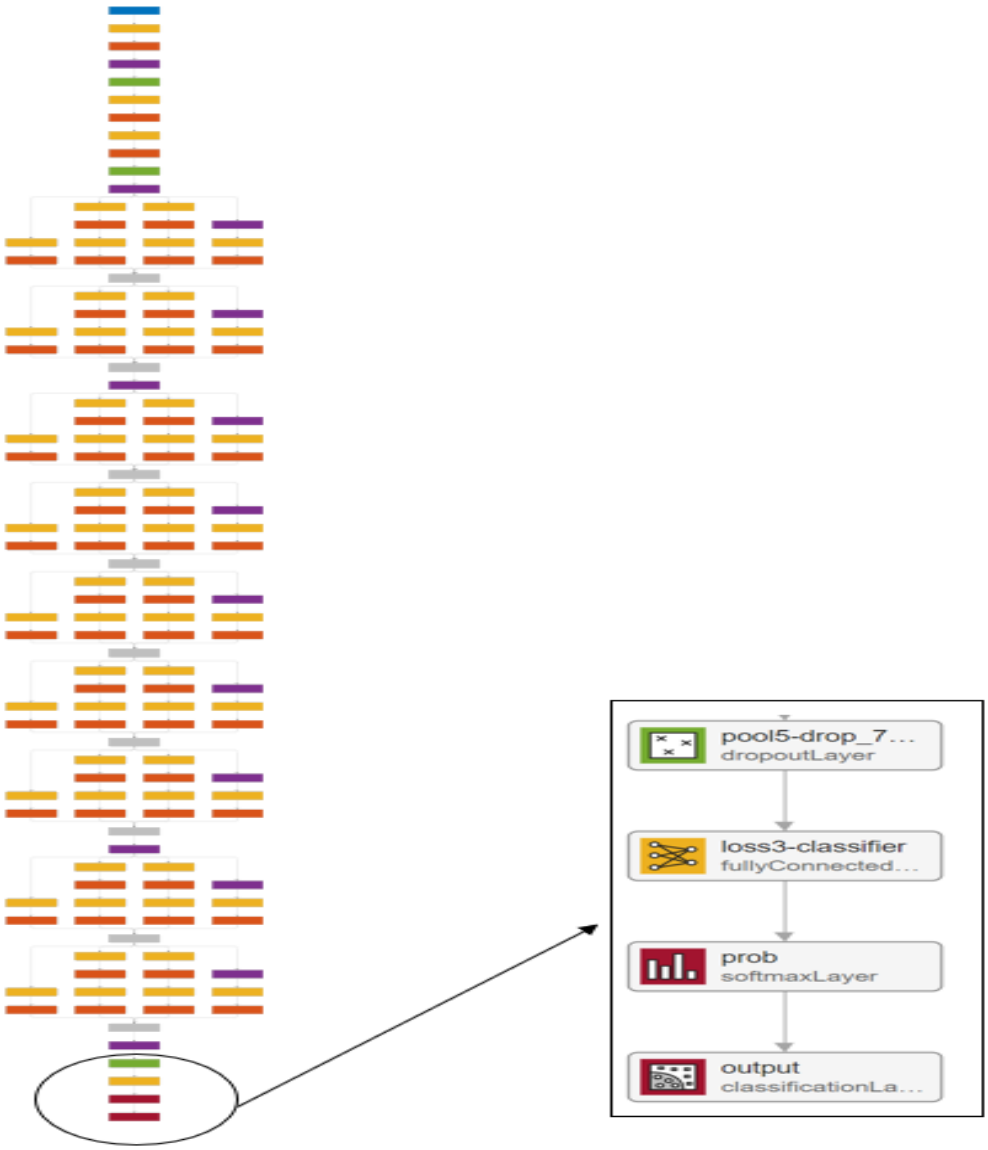


Figure 6.2: GoogLeNet Framework. The expanded view shows the layers that are finetuned before retraining on new dataset (Image recreated from MATLAB).

Table 6.3: Transfer Learning Training Options

Training Options	**Value**
Solver	Stochastic Gradient Descent with Momentum (sgdm)
Mini Batch Size	28
Maximum Epochs	30
Initial Learning Rate	0.0001
Validation frequency	20

Combining Convolutional Neural Network and Support Vector Machine

For comparison purposes, a combined CNN-SVM classifier was also used to classify the raw AE waveforms. This approach has been found to be as good as and, in some cases, having higher classification accuracy than ordinary CNN models (Tang 2013; Niu and Suen 2012; Agarap 2017).

The main purpose of Support Vector Machines (SVMs) is to find the optimal hyperplane that can be used to separate two classes in a dataset (Agarap 2017). This is done using different kernel functions to transform the non-linearly separable data into a linearly separable form by projecting the data into the feature space (Niu and Suen 2012).

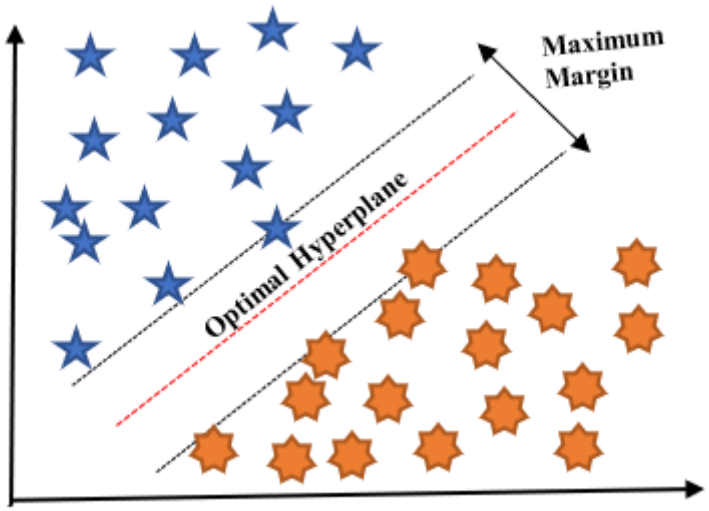


Figure 6.3: Illustration of the Support Vector Machine

Using CNN in combination with SVM can be advantageous because there is no need to retrain the network since the images just need to be passed over the network once thereby saving some time. This is particularly helpful in situations when there is no access to a GPU to speed up training. The features were extracted by running activations over the last global pooling layer. This is the first layer with no more learnable parameters and it is also robust to spatial translations of the input data. Extracting features from the deeper layers also provides higher-level features. Features for training and testing were extracted in this way, and the

features were then trained on a SVM network using the classification learner app in MATLAB. Two forms of CNN-SVM classifications were compared for this study. In the first one, features were extracted from the last global pooling layer of the un-trained network and then fed into the SVM model as the input. In the second form, features were extracted from the modified and retrained CNN network and then fed into the SVM for training and testing. The MATLAB script used for feature extraction can be found in the appendix.

Kernel functions in SVMs aren't always linear, they can be quadratic, polynomial or gaussian with different hyperparameters. The trained SVM's accuracy will depend on how the different hyperparameters relate to each other making, hyperparameter optimisation a necessity when training SVMs. The optimisable SVM was first trained to find the best hyperparameter values with the least classification error. These were then used to train and test the SVM network using the features extracted from the GoogLeNet CNN. The same training examples used for the modified GoogLeNet network were used for the two CNN-SVM models. Training was carried out using 6-fold cross validation.

6.3. Results and Discussion

6.3.1. Damage Mode Validation

Table 6.4 shows that the burnishing tests have lower coefficient of friction (CoF) and frictional work done (FWd) in comparison to the scratching tests. Both CoF and FWd are dependent on the friction force generated between the contacting surfaces and it is expected that scratching tests would generate higher friction force. In addition to the surface asperities coming into contact, the presence of the silicon grinding grit in the scratching tests provides more resistance to sliding hence the frictional force required to break the asperities would be higher, leading to the higher CoF and FWd values. In his pin-on-disk wear test of UHMWPE

sliding against a steel counterface, Vesa Saikko also obtained average CoF values of about 0.03 (Saikko 1998). The study by Fisher et al has shown that increased counterface roughness can lead to an increase in the CoF value (Fisher et al. 1994) further validating the friction values obtained for the scratching tests where the presence of the silicon particles would have caused an increase in the surface roughness of the counterface and this would have contributed to the increase in CoF value.

Table 6.4: Average Coefficient of Friction and Frictional Work Done

Damage Mechanism	Average CoF (±std)	Average FWd (±std), J
Burnishing	0.031 (±0.007)	3,801.4 (±876.23)
Scratching	0.171 (±0.055)	10,692 (±3,447)

Wear Scar Analysis

After each test, the images of the wear surface were taken and analysed using the Alicona InfiniteFocusG5 Optical 3D Measurement System. Image field captures of the UHMWPE wear surface was taken at a magnification of 50X (Figure 6.4 and Figure 6.5).

There is evidence of displaced wear particles adhering back on to the UHMWPE surface (See region A in Figure 6.4) confirming burnishing wear mode since burnishing wear is a form of adhesive wear (Puloski et al. 2001). Burnishing wear is also known to lead to a smooth and shiny surface, and this is visible in the wear scar image. Comparison between the pre-test and post-test surface roughness measurements (pre-test: 0.835 ± 0.046 µm and post-test: 0.486 ± 0.062 µm) also shows that the UHMWPE surface is smoother post-test thereby confirming burnishing wear mode has occurred.

Inspection of the UHMWPE surfaces post-scratching tests show the presence of grooves indicative of scratching having taken place during sliding (See Figure 6.5). Further

analysis of the CoCrMo counterface (See Figure 6.6) also shows evidence of scratching in the presence of the silicon carbide grinding grit between the two contacting surfaces. It is expected that the scratching process would make the surfaces, rougher and the comparison of the pre-test and post-test surface roughness values of the UHMWPE supports this (pre-test: 0.863 ± 0.027 µm and post-test: 1.194 ± 0.067 µm) further confirming that scratching has taken place.

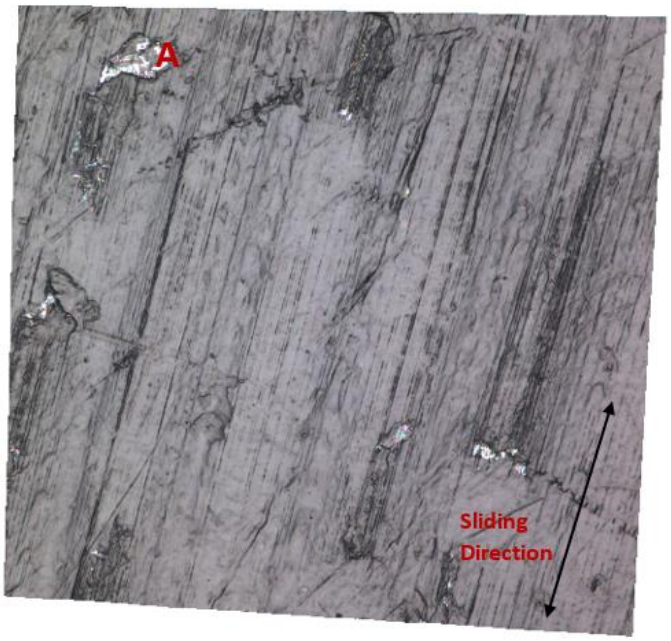


Figure 6.4:Image field capture of UHMWPE wear scar for burnishing tests. Region labelled A shows adhesion of displaced wear particles back unto the UHMWPE surface.

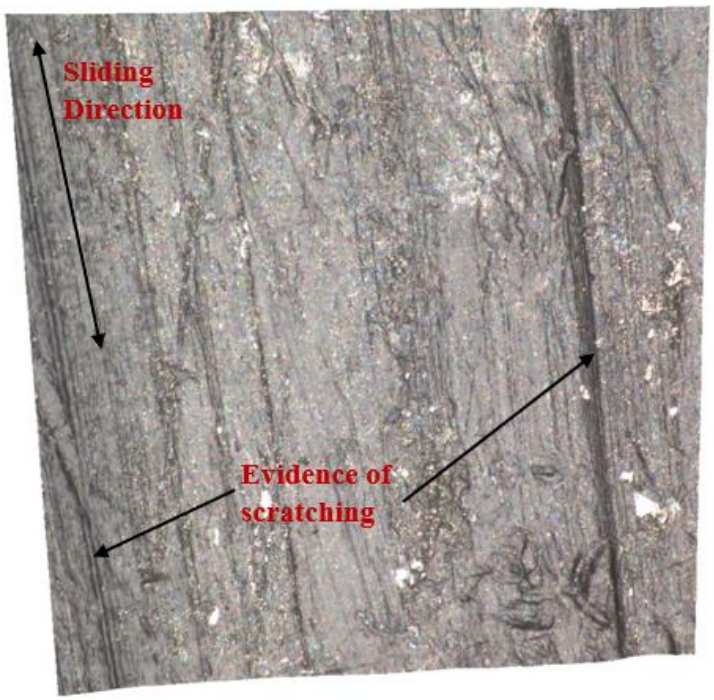


Figure 6.5:Image field capture of UHMWPE wear scar for scratching tests. There are grooves on the surface caused by the scratching of the grinding grits on the surface during sliding.

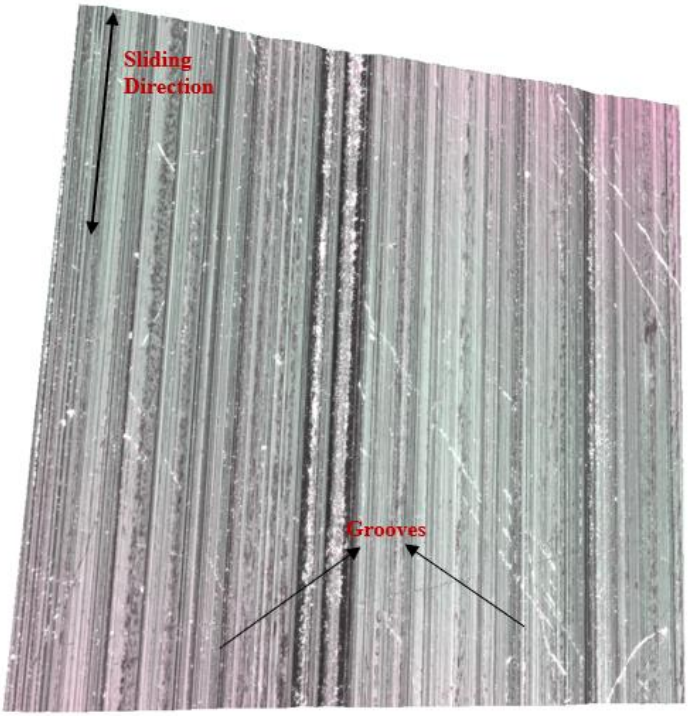


Figure 6.6: Image field capture of the CoCrMo counterface post-scratching tests. The grooves indicate scratching wear having taken place.

6.3.2. Relationship between AE hit features and damage mechanism

Coefficient of Friction (CoF) and AE Root Mean Square (RMS) value

The average CoF and average AE root mean square (RMS) values for the two simulated damage modes exhibit the similar transient response as shown in Figure 6.7 and Figure 6.8. This is in line with the result obtained from the proof-of-concept tests in chapter 4 where it was shown that the coefficient of friction curve can be predicted using time-dependent AE features of which AE RMS value is one of them.

Since CoF values are based on the friction force generated during sliding and the friction force is a representation of the energy required to break contacting surface asperities during sliding, it is expected that time-dependent AE features such as AE RMS values will mirror the CoF values obtained as well. A high friction force for example implies a high amount of energy is required to break surface asperities which would in turn generate high CoF values as well as AE signals with high energy and thus a high RMS value. This is reflected in the scratching tests where the presence of the silicon carbide grinding grit would cause an increase in the friction force generated hence why the CoF and AE RMS values are higher than those observed in the burnishing tests. In their work on how AE can be used for the early detection of failure modes in total hip replacements, Lee et al also found that CoF and AE energy show a similar transient response (Lee et al. 2021).

In contrast to the single simulated damage tests, the CoF and AE RMS values of the damage progression tests do not have the same transient response from start to finish. Upon transitioning from burnishing to scratching wear mode (due to the addition of the silicon carbide grinding grit), the CoF reduced whilst the AE RMS value increased and was the case for all the repeat tests (See Figure 6.9). This contrasts with the initial hypothesis where it was expected that the presence of the grinding grit would cause an increase in both CoF and AE

RMS values. There are two possible reasons for the reduced CoF after grinding grit was introduced to the test bath. First, since the UHMWPE discs and the CoCrMo were already in contact, it was difficult to ensure the grinding grit added made its way between the contacting surfaces. The alternative was to stop the test entirely and remove the UHMWPE disc before adding the grinding grit on to the CoCrMo counterface to ensure it stayed within the contact area but this would cause the tribofilm to break which would change the test and comparison between the two test regions would not be possible. Second, it is suspected that after pausing to add the grinding grit, the tribofilm between the contacting surfaces creates a smooth barrier leading to a decrease in friction force hence reduced CoF values.

On the other hand, there was an increase in the AE RMS value after the grinding grit was added. Despite not being perfectly within the contact face, the sliding motion would cause a dispersal of the grinding grit on the CoCrMo counterface. As sliding progresses upon test re-start, the scratching of the grinding grit on the surfaces would produce strain energy that is picked up by the AE sensors in addition to the ones due to the UHMWPE sliding on the CoCrMo hence the increase in the AE RMS value. This highlights the advantage of the AE technology over conventional monitoring techniques and shows the potential of AE testing being highly sensitive to changes that might not be picked up by other conventional monitoring systems. In their study on the relationship between AE and the wear phenomena in severe-mild wear transition, Alan Hase et al also found that using the AE mean value was more reliable than CoF in distinguishing the differences between the wear mode of different sliding conditions (Hase, Wada, and Mishina 2008).

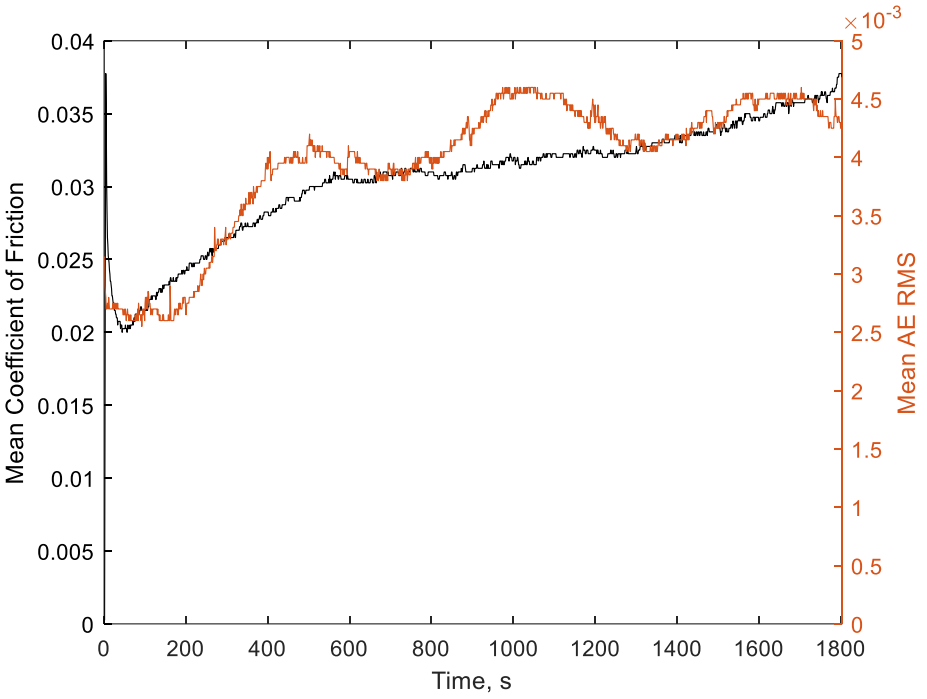


Figure 6.7: Mean CoF and Mean AE RMS plots for burnishing test

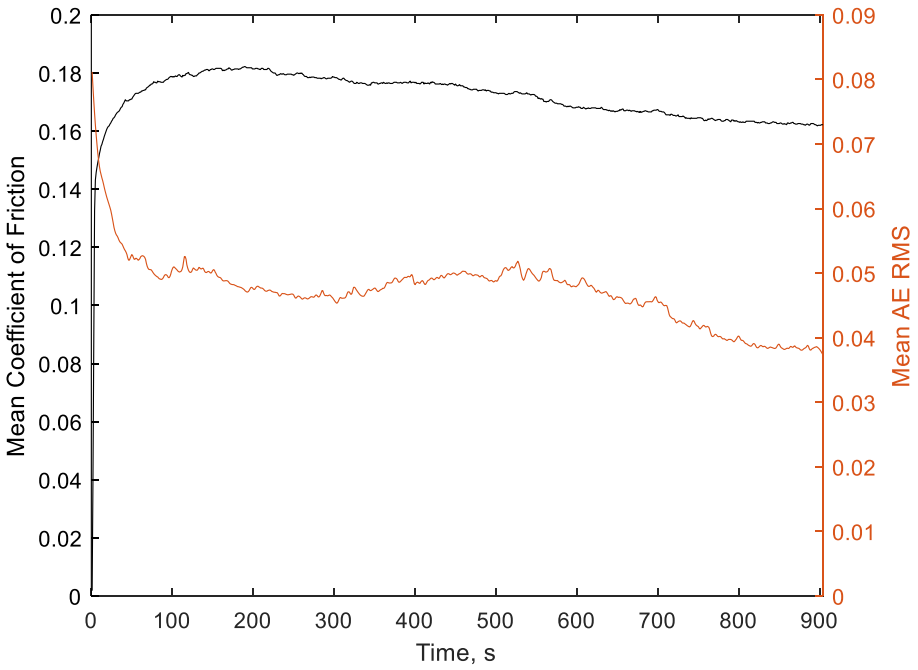


Figure 6.8: Mean CoF and Mean AE RMS for scratching tests

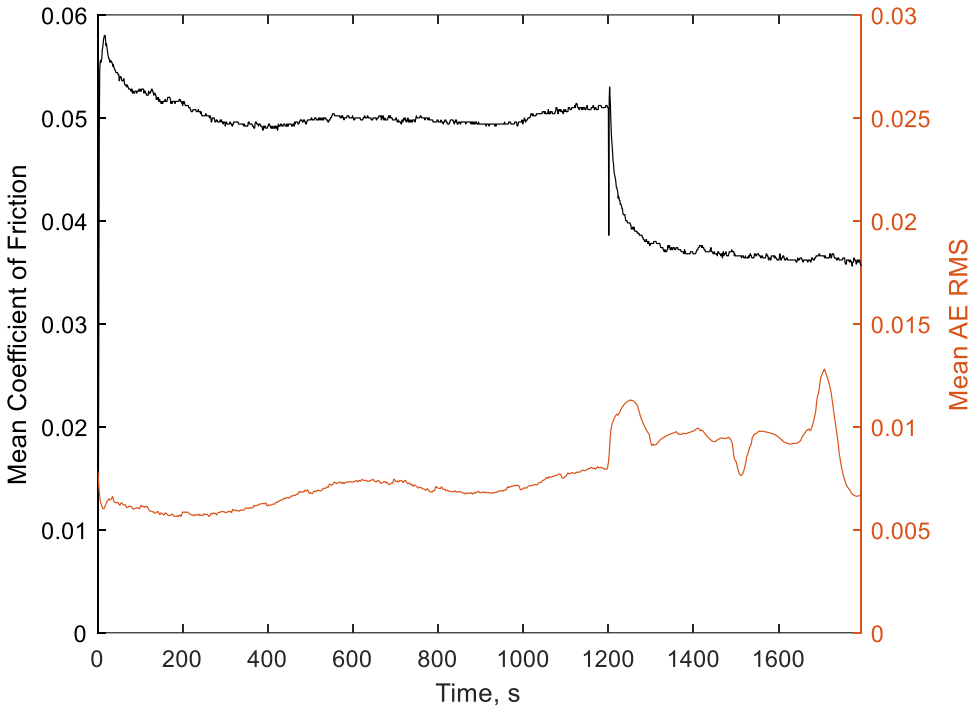


Figure 6.9: Mean CoF and Mean AE RMS plots for wear mode progression tests

One-way Analysis of Variance

The one-way ANOVA test showed that there are significant differences between discrete AE features of the burnishing and scratching tests with the p-value being less than 0.05 for all tests (See Table 6.5). For all features tested, scratching tests have higher values than burnishing tests apart from the average frequency value. Considering the more severe wear mechanisms associated with the scratching tests, it is expected that AE hits from scratching tests would have higher intensities than those from the burnishing tests.

Further analysis of the ANOVA test result shows absolute energy, amplitude, duration and peak frequency as features with the least p-values. Of these, peak frequency and amplitude emerge as the best features to distinguish between burnishing and scratching as reflected in the box plots of the feature data shown in Figure 6.10. In the context of using AE to differentiate between damage mechanisms, this finding implies that peak frequency and amplitude are two

significant discrete AE features to be considered during analysis. This finding also aligns with what has been reported in literature regarding the correlation of AE features and wear mechanisms. Hase et al found that amplitude and peak frequency can be used to differentiate between adhesive and abrasive wear mechanisms (Hase, Mishina, and Wada 2012) in pin-on-block metal contacts. In addition to this, they also created a correlation map showing that amplitude and AE frequency (peak frequency in particular) is significant for identifying wear mechanisms involving deformation and fracture. Amplitude was also found to be significant in classifying between damage mechanisms found in laminated material (Masmoudi, El Mahi, and Turki 2015).

Observing a similar relationship between AE discrete features and damage mode identification to those reported in literature gives further credence to the assumption that AE can be advantageous in the monitoring and identification of tribological mechanisms of artificial joint materials with the potential to be extended to the diagnosis of failed artificial joints.

Table 6.5: Summary of ANOVA test results. StDev is standard deviation and CI is confidence interval.

Features	**Burnishing Tests**			**Scratching Tests**			**P value**
	Mean	StDev	95% CI	Mean	StDev	95% CI	
Absolute Energy, aJ	13.43	8.54	-63.30, 90.16	1418.9	2475.7	1342.2, 1495.6	9.57e-132
Amplitude, dB	37.52	0.85	37.42, 37.62	45.97	3.17	45.87, 46.07	0
Average Frequency, kHz	223.10	167.08	217.14, 229.05	185.09	95	179.13, 191.04	1.43e-18
Duration, µs	527.70	291.74	513.64, 541.76	762.71	347.34	748.65, 776.77	1.16e-111
Peak Frequency, kHz	210.58	13.23	209.79, 211.37	232.81	21.73	232.02, 233.60	4.05e-283
RA Value, µs/dB	7.48	5.37	7.25, 7.72	8.74	5.30	8.51, 8.98	1.10e-13

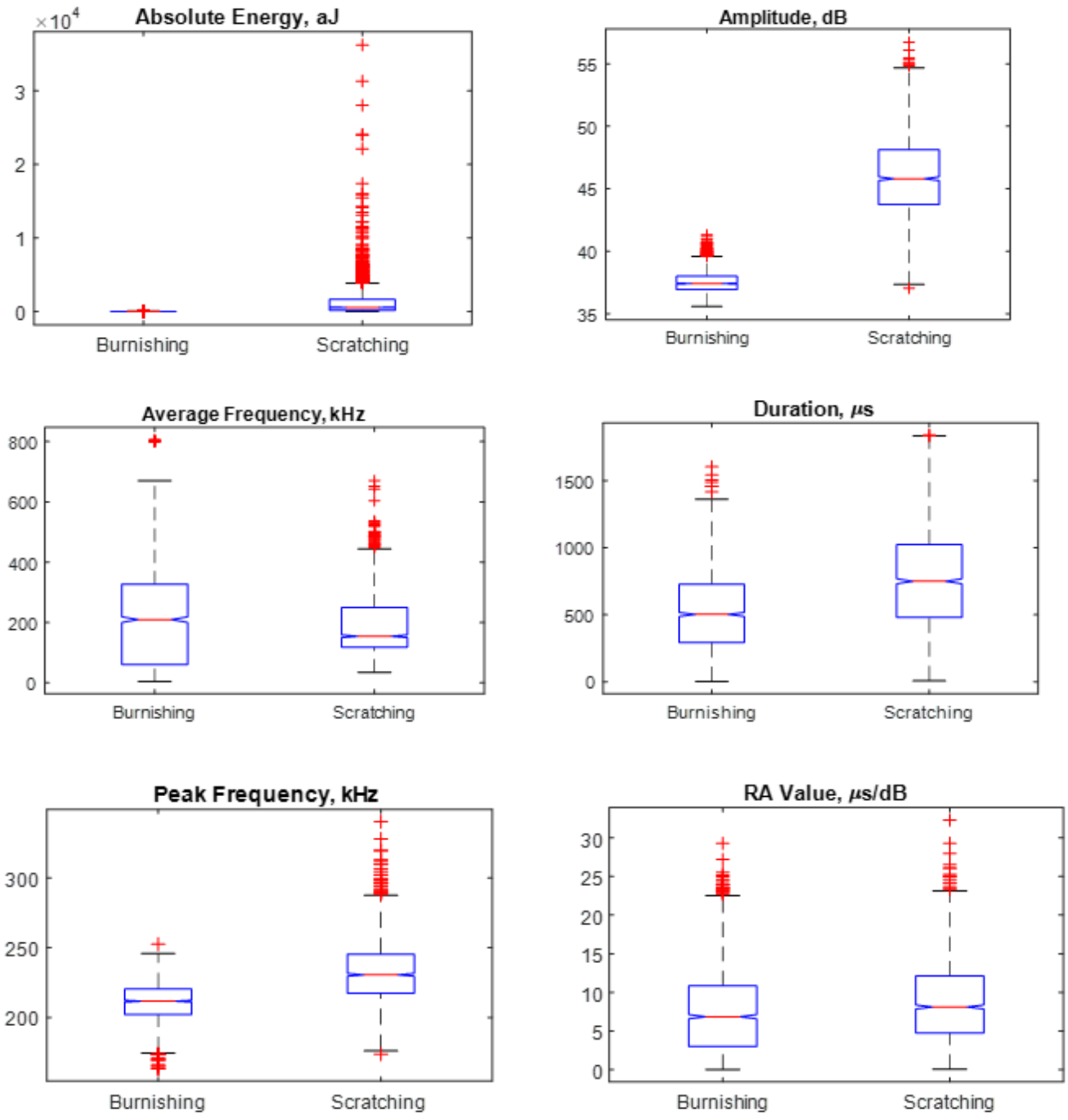


Figure 6.10: Box Plots of Average AE Features showing result of ANOVA test.

6.3.3. Classification of AE signals based on damage mode

Raw AE waves of 10 seconds length each from the burnished, scratched and the progression tests were acquired for the CNN classification. Before denoising, the waveforms were split into smaller waveforms of 200 ms duration, which is equivalent to one cycle on the TE77, resulting in a total of 960 waveforms. These were then denoised and split for training and testing at a ratio of 70:30 resulting in 672 images for training and 288 for testing. The denoised AE waveforms were then transformed into CWT images (Figure 6.11 and Figure 6.12) for classification using the three convolutional neural network models stated in section 6.2.3 – GoogLeNet transfer learning, CNN-SVM using features from un-trained GoogLeNet and CNN-SVM using features from the trained GoogLeNet.

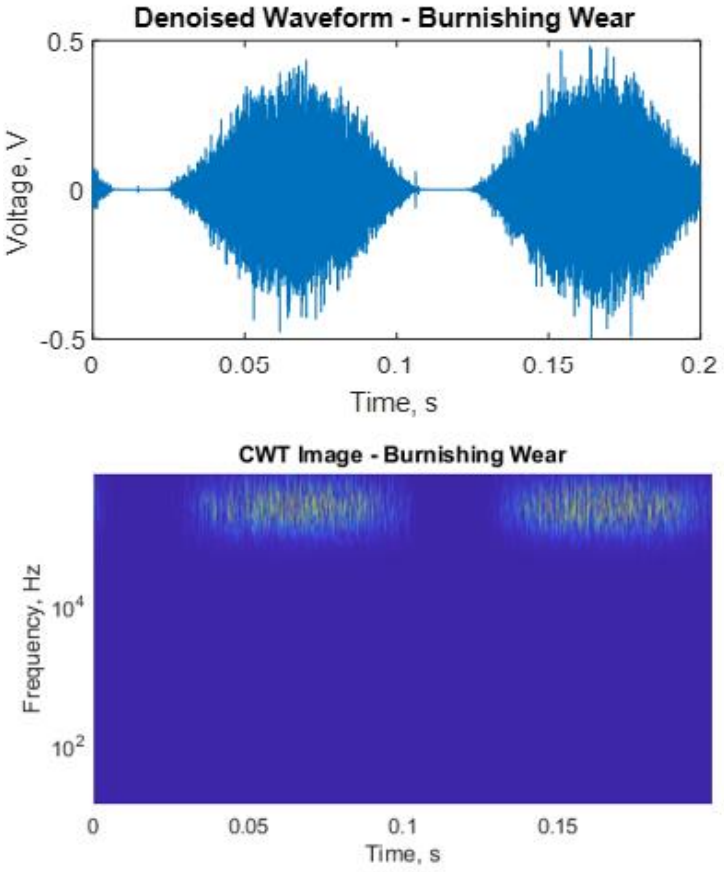


Figure 6.11: Denoised waveform (top) and the corresponding CWT Image (bottom) of a sample signal for Burnishing Wear tests.

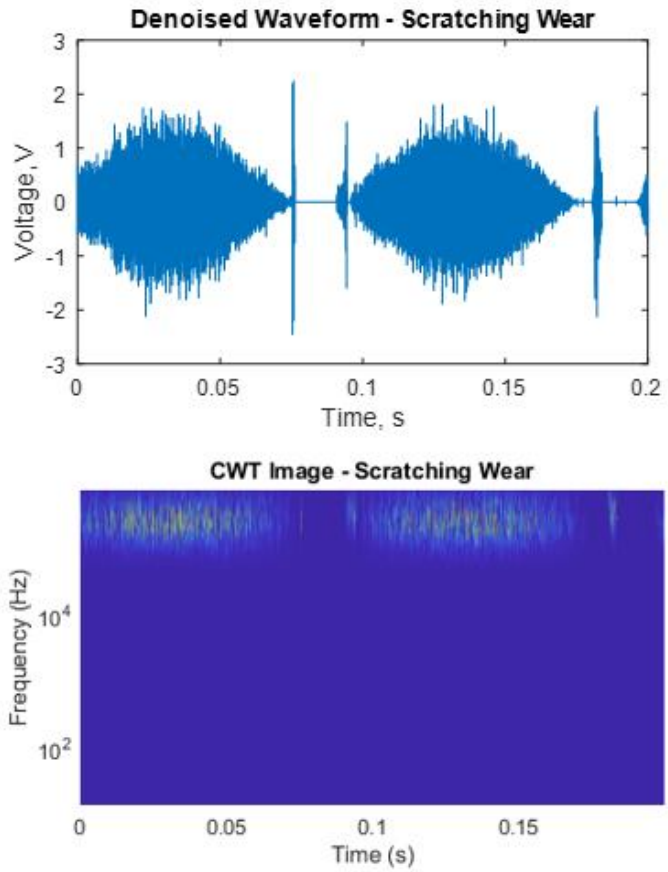


Figure 6.12:Denoised waveform (top) and the corresponding CWT Image (bottom) of a sample signal for Scratching Wear tests.

The training progress for the GoogLeNet transfer learning model (See Figure 6.13) shows that as the number of iterations increases, the training & validation accuracy increases whilst the loss value decreases. With each new iteration, the network finds new relationship between the features of the CWT images thereby making better predictions.

As previously mentioned in section 6.2.3, minimum classification error plots were used to find the best hyperparameters to use for the CNN-SVM models. This turned out to be linear SVM for the model using features from the un-trained GoogLeNet network (See Figure 6.14) and Gaussian SVM for the model using features from the trained GoogLeNet network (See Figure 6.15). Feature extraction from the last global pooling layer of the GoogLeNet network resulted in 1024 features. Hence, there were 672 by 1024 training data matrix and 288 by 1024 test data matrix.

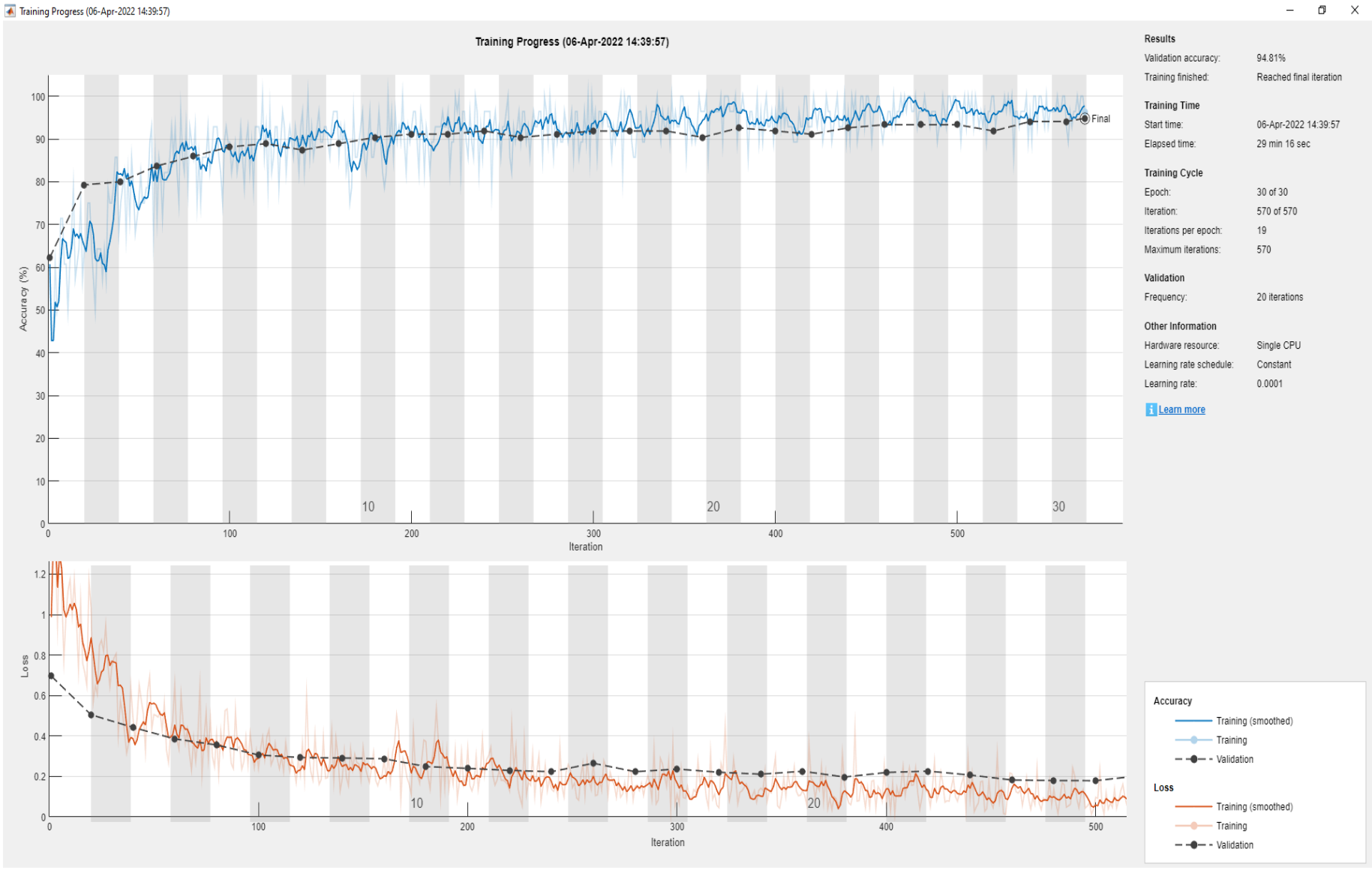


Figure 6.13: CNN Transfer Learning (based on GoogLeNet) Training Progress. The top image is the accuracy plot and the bottom image is the corresponding loss plot.

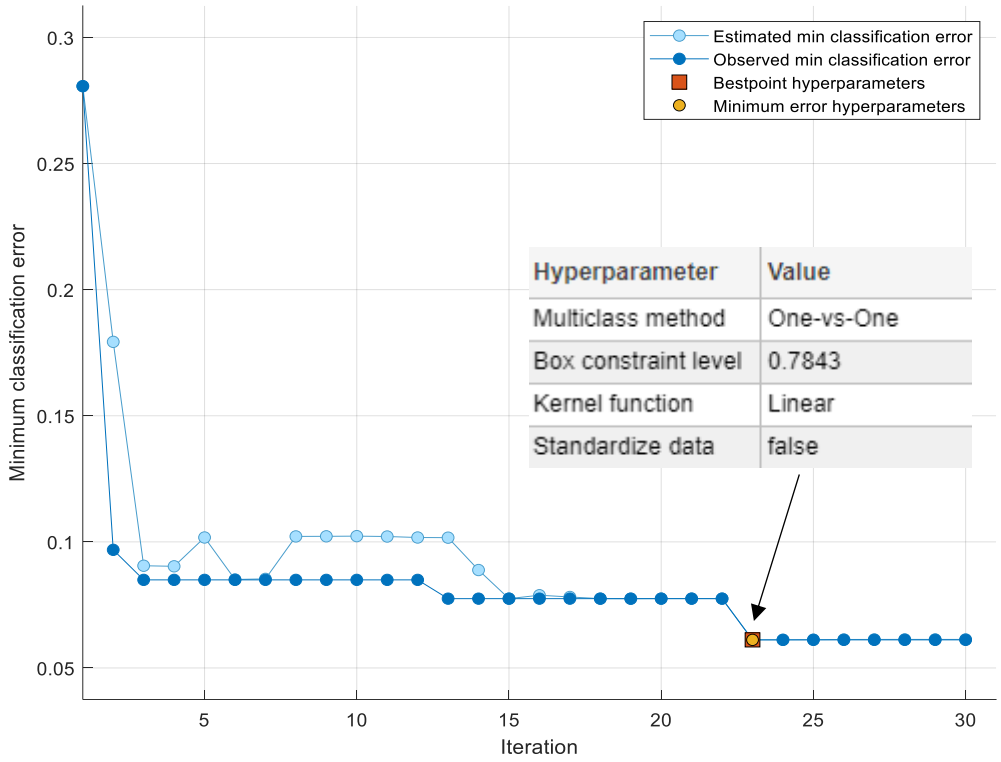


Figure 6.14:Minimum classification error output for CNN-SVM using un-trained GoogLeNet features

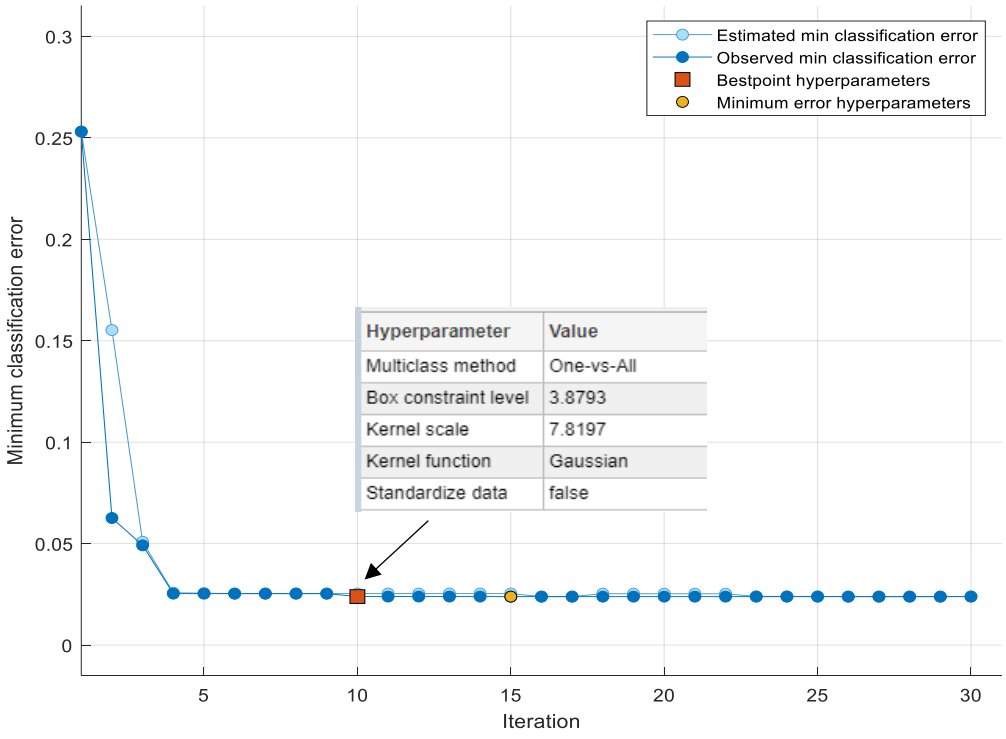


Figure 6.15: Minimum Classification Error Plot for CNN-SVM using trained GoogLeNet features.

After training, all three models were tested using the test data. The training and test results are presented in Table 6.6. All three deep learning models achieved a test accuracy of above 90% with the CNN-SVM model with features from the trained GoogLeNet network performing the best with a test accuracy of 99.3%. Quite unsurprisingly, the CNN-SVM model using untrained GoogLeNet features has the worst training and test accuracy of 93.5% and 95.1%, respectively. As shown in Figure 6.13, the training examples must be passed over the deep network for several iterations for the network to fully learn the relationship between the features for improved predictive capability. When extracting features from the last global pooling layer of the unmodified and untrained network, the CWT images are only passed over the network once and the network is yet to fully learn and extract all the important features of the CWT images. On the other hand, features extracted from the trained network would have better predictive capabilities because the network has been able to fully learn how the training examples of each class are related to each other thereby producing features that are more representative of the class they belong to.

Table 6.6: Training and Test accuracy of all three CNN models.

CNN Models	Training Validation Accuracy	Test Accuracy
GoogLeNet Transfer Learning	0.9482	0.9653
CNN-SVM (using un-trained GoogLeNet features)	0.9345	0.9514
CNN-SVM (using trained GoogLeNet features)	0.9688	0.9931

Of the three deep learning models tested, the two models with the higher classification accuracy both perform better at predicting burnishing wear modes as opposed to scratching wear modes. The confusion matrix plots (Figure 6.16 to Figure 6.18) confirm this. There are some wear processes, e.g., sliding friction, that are common to both burnishing and scratching

damage modes and the CWT images of the signals are likely to have similar feature characteristics thereby causing the small number of misclassifications observed to occur. Although the misclassification errors could possibly be further minimised by undertaking more signal pre-processing steps before wavelet transforms and by building a new deep learning model from scratch. As only a small percentage of the examples are misclassified, it is not beneficial to spend the extra computing power required for building a deep learning from scratch which can take hours to days before a fully optimised network is built. The above 95% classification accuracy achieved with deep transfer learning and wavelet transforms shows that there is huge potential for using AE testing to identify and possibly diagnose natural and artificial joint pathologies.

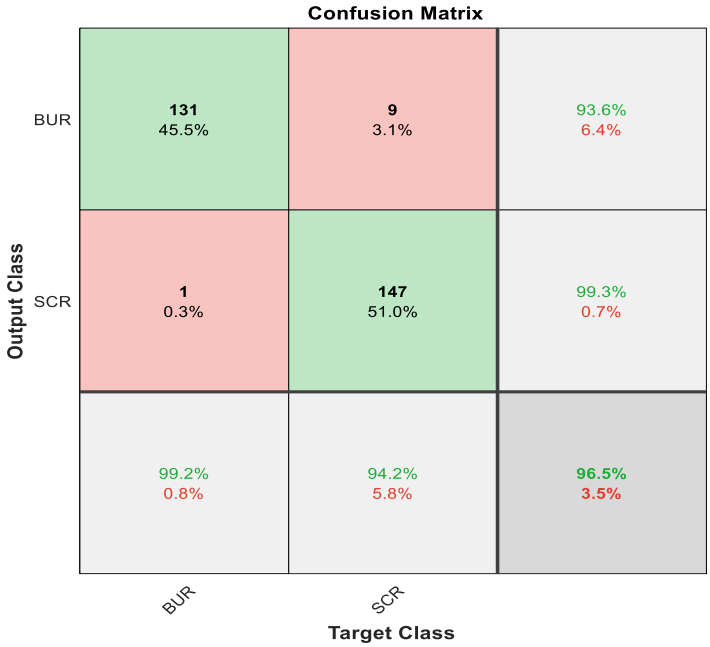


Figure 6.16: GoogLeNet Transfer Learning Test Confusion Matrix.

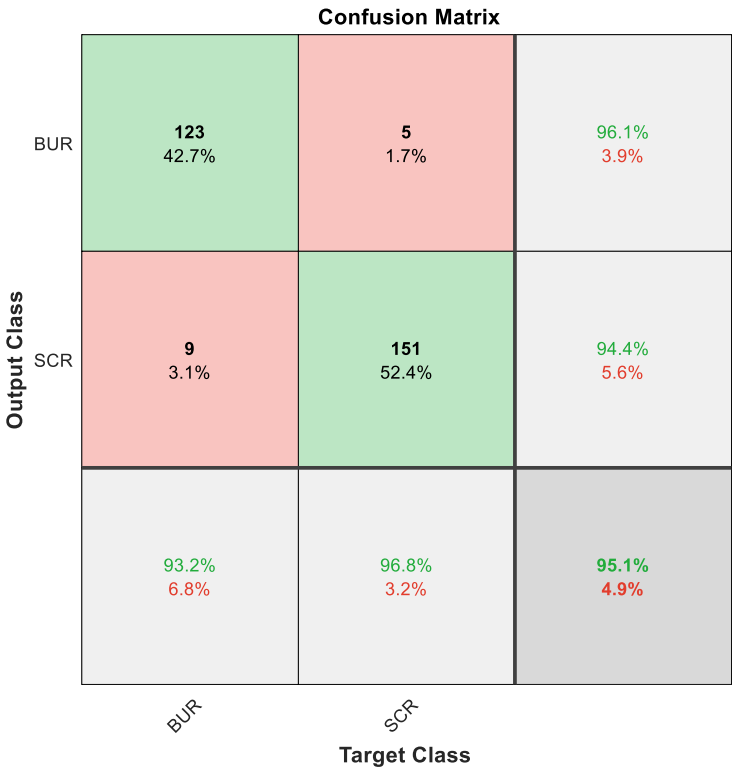


Figure 6.17: Test confusion matrix for CNN-SVM using un-trained GoogLeNet features

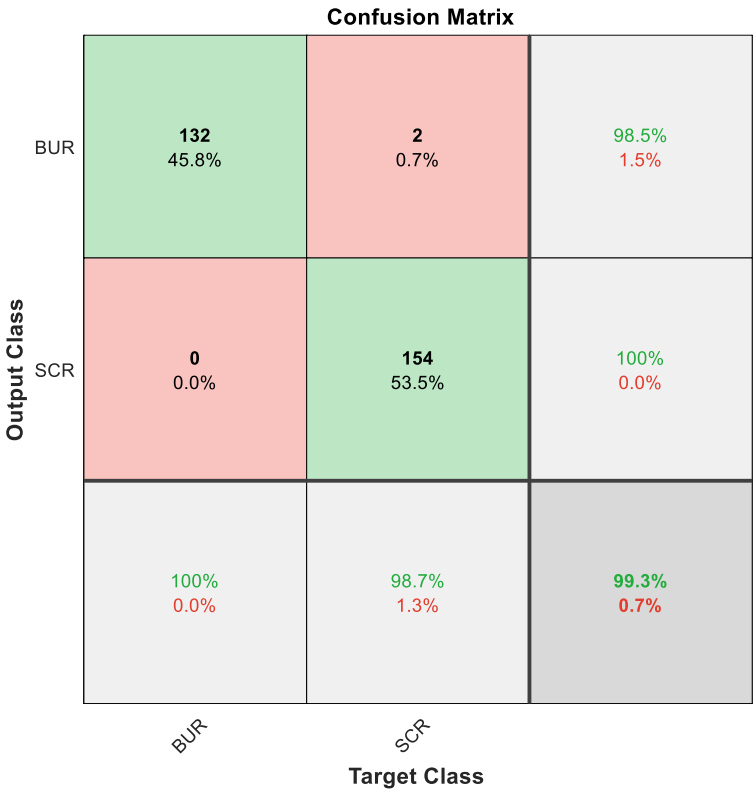


Figure 6.18: Test confusion matrix for CNN-SVM using trained GoogLeNet features

The classification accuracy achieved here is comparable with other studies where AE signals are classified using wavelet analysis and deep transfer learning. In their study on classifying AE signals from a broken wire based on damage modes – fracture and friction, Ren and Chen obtained a classification accuracy of 92.35% using a VGG-Net deep transfer learning of the wavelet images (Ren and Chen 2021). Xin et al. used the same process to identify fracture within stay-cables using GoogLeNet based transfer learning and obtained an accuracy of 99.53% (Xin et al. 2020). It is expected that in stay cables for bridge applications, the AE signals due to noise would be very distinct from those due to fracture, whereas as stated previously, there are feature characteristics that are present in both burnishing and scratching damage modes. This could explain why they obtained a higher classification accuracy than the one obtained in this study. König et al also used CNN for multi-class classification of AE events from sliding bearing systems and obtained an accuracy of 82.5% (König et al. 2021).

With the CNN-SVM model where features are obtained from the trained GoogLeNet-based CNN, a classification accuracy of 99.3% is obtained. This is also comparable with the study by Niu and Suen, where they developed a novel CNN-SVM classifier to classify handwritten digits (Niu and Suen 2012). Using CNN-SVM the way it has been utilised here, they obtained a classification accuracy of 99.81%. Although the application is different from the one used here, the similar accuracy obtained shows the potential for using deep transfer learning of CWT images to classify AE signals from biotribological contacts. In his study on combining CNN and SVM for image classification, Agarap proposed a new CNN-SVM architecture where instead of using them as two separate models, the softmax classification function which is usually the last layer of a CNN is replaced with a L2-SVM for an integrated CNN-SVM architecture (Agarap 2017). Although when compared with normal CNN, CNN-SVM had a slightly lower classification accuracy (99.23% for CNN-Softmax and 99.04% for CNN-SVM), the difference is not much and the base CNN model was a basic one. He

concluded that better pre-processing and a more sophisticated base CNN can improve the results. By denoising the acquired raw AE signals and using a highly sophisticated base CNN model like GoogLeNet, there is potential for achieving even higher classification accuracy through the use of an integrated CNN-SVM architecture.

Despite the novelty of using deep transfer learning to classify AE signals from orthopaedic applications, deep learning has been succesfully used in medical diagnosis (Wong, Fortino, and Abbott 2020; Lundervold and Lundervold 2019; Aggarwal et al. 2021) showing there is potential for using it to classify AE signals for the purpose of diagnosing natural and artificial joint pathologies.

6.4. Conclusion

This chapter has shown how AE testing can be used to classify AE signals from simulated artificial knee joint articulations based on two damage modes – burnishing and scratching. The following conclusions were reached.

- One way analysis of variance of discrete AE features show that amplitude and frequency (peak frequency) data are significant for distinguishing between damage modes.
- Raw AE signals can be transformed into wavelet images and then classified using deep transfer learning with GoogLeNet as the base CNN model.
- A higher classification accuracy is obtained with a CNN-SVM model when compared to a standard CNN (99.3% and 96.5% respectively).

Coming up in chapter 7 will be the overall discussion, potential further studies and conclusions.

Chapter 7

Overall Conclusions

7.1 Introduction

Having discussed the results of the experimental work carried out for this project in chapters 4 to 6, this chapter presents the overall conclusions reached and potential further work.

7.2 Overall Conclusions

This thesis aimed to investigate the potential of acoustic emission testing as a tool for diagnosing tribological phenomena in artificial joints. This was achieved by undertaking four main objectives through a combination of tribo-acoustic testing and machine learning and deep learning techniques for classifying AE signals based on simulated damage modes.

The proof-of-concept test showed that time-dependent AE features could be used with NARX neural networks (Training R2 values of 99.95% for PEEK and 98.58% for UHMWPE) to predict the shape of the coefficient of friction profile of a simulated joint articulation surface. This achievement is particularly significant for the clinical usability of the technique. It shows the potential of using AE to evaluate the frictional behaviour of an artificial joint *in vivo,* which is not currently achievable. The proof-of-concept test also satisfied this thesis's second objective, which was to use AE to observe the progression of damage in a simulated artificial joint articulating surface. It was found that this can be achieved by clustering AE signals based on the emission type, with the continuous emissions emerging as the best emission type for identifying test stages and observing the evolution of surface damage. This shows that AE testing could be used for the condition monitoring of an artificial joint *in vivo* from implantation through the life of the device.

Having shown the capability of AE testing as a tool for understanding the tribological behaviour of an artificial joint *in vivo,* the second experimental work presented in this thesis showed how machine learning techniques can be used to classify AE signals based on wear

mechanisms simulated under controlled joint conditions. This is required to achieve AE testing's full potential as a tool for diagnosing tribological phenomena in artificial joints. K-means clustering was successfully used to group AE hits based on the different tribological processes present. This shows that different stages of a tribo-acoustic test can be diagnosed using AE testing. Before the classification process, PCA was used to extract significant AE features first. Of the three classification models tested, the BP neural network emerged as the one with the highest classification accuracy (96%). In addition to having a high classification accuracy, the BP neural network is also the simplest and most dynamic of the classification models making it the optimum choice of pattern recognition technique for classifying AE signals. The high classification accuracy obtained also shows that AE hits from tribo-acoustic tests carried out under controlled joint conditions have hidden relationships that can be identified using pattern recognition techniques. This is clinically significant because being able to differentiate between AE signals based on wear mechanisms takes us one step closer to achieving the potential of AE as a tool for diagnosing natural and artificial joint pathologies.

Having shown that machine learning techniques can be used to classify AE signals based on wear mechanisms, the final experimental work explored the use of deep learning (a more advanced version of machine learning) for classifying AE signals from simulated artificial knee articulating joint surface based on two damage modes – scratching and burnishing. The advantage of the deep learning technique over traditional machine learning techniques is that complex feature engineering is not required in deep learning, making it a faster and more intuitive classification process. On a statistical basis, a one-way ANOVA analysis showed that amplitude and peak frequency are the two most significant AE features for differentiating between damage modes. Continuous wavelet transform images of denoised AE signals were then trained using deep transfer learning models with GoogLeNet as the base CNN model in two forms – (i) A simple GoogLeNet transfer learning model and (ii) a

combination of GoogLeNet and support vector machine model. The GoogLeNet-SVM (with trained features) emerged as the best model with a 99.3% classification accuracy. The classification accuracy achieved here is comparable with that obtained in other applications where transfer learning was used to classify AE signals proving the technique's effectiveness.

Overall, the research presented in this thesis has shown that AE testing can be used to interpret the biotribological phenomena in artificial joint materials. In addition, machine learning and deep learning techniques were used to build models with high sensitivity, specificity, and accuracy to classify AE signals according to simulated damage modes. This has proven the initial hypothesis that with AE testing, a more dynamic, highly specific, and highly sensitive process of identifying and diagnosing artificial joint pathologies can be developed thereby reducing patient discomfort and NHS expenditure.

7.3 Potential Further Work

The research presented in this thesis suggests an enormous potential for using AE as a tool for monitoring artificial joint tribological behaviour with the possibility to be extended to the monitoring of natural joint pathologies. Still, it is not without its limitations.

The tribo-acoustic test parameters were simplified to align them with the operating envelope of the TE77 test rig and for ease of analysis. Linear reciprocating sliding motion and a flat-on-flat contact surface was chosen for all tests for ease of interpretability of the AE signals acquired. Further developmental work will need to be carried out to better replicate the complex kinematics of actual joint prostheses using bench test geometries and motions that are closer to those of the actual devices.

Traditionally, *in vitro* biotribological tests are carried out using bovine serum to represent the synovial fluid present in the joints. Bovine serum contains some chemical reagents, and there is a chance these could also be a source of AE signals which is why tests presented in this thesis were carried out using Ringer's solution instead to make sure that the

AE signals acquired are solely due to the tribological interactions in the contacting surface alone. Further work using bovine serum would have to be carried out to investigate whether the bovine serum affects the AE signals acquired. This would also indicate how the presence of synovial fluid around the joints might affect the quality of the AE signals acquired during *in vivo* application of the technology.

The machine learning techniques used for the classification could be further developed to build more intuitive and robust classification models. Discrete features used were chosen based on what has been reported in literature and it is worth noting that there are limited biotribological applications of AE signal analysis using machine learning. A more developed feature engineering process could be used to derive more significant and higher-order features, which could further improve the classification models' accuracy.

For the AE testing method to be transferable to clinical settings, there is considerable work still to be done.

Since tests presented in this thesis are bench tests, further developmental work using the actual devices is required to build a profile of the AE signature of different artificial joints as a reference for the interpretation of signals acquired *in vivo.* Exploration of the effects of wear throughout the device's lifetime would also be a further step toward a condition monitoring system.

Tests presented in this thesis were carried out under simplified conditions, but that is not the case for an actual artificial joint where there will be more than one source of AE signals. To this end, building a machine learning model that will be robust to differing conditions is important. This could be done by simulating different test conditions reminiscent of what is expected in an actual artificial joint to acquire AE signals that can be used to train more machine learning models.

The deep transfer learning models presented in this thesis were built using wavelet transforms of the AE signals. More training data can be acquired by also creating spectrograms of the signals and stitching them together with the wavelet transforms thereby creating a deep learning model that is more robust to differing conditions. This is particularly useful for the CNN-SVM model where more significant features can be extracted from the CNN model for training and testing using SVM.

To summarise, further development of machine learning techniques and higher-order analysis to observe the relationship between AE features and simulated damages will reveal additional insights. Ultimately, the transition of the proposed method from *in vitro* to *in vivo* will require refinement of signal acquisition techniques, data processing and analysis.

APPENDICES

Appendix A1. Test Geometry Determination

At a maximum possible load of 2000 N (BS ISO 18192-1, 2011), the maximum contact pressure of a ball and socket Charite Lumbar Spinal Implant (with ball radius of 14 mm and clearance of 0.35 mm (P. M. Moghadas *et al.*, 2012)) was calculated to be 9.7 MPa and a contact area radius of 9.91 mm using Hertzian contact mechanics (Johnson, 1985). The load and corresponding contact area radius required to achieve an equivalent contact pressure on the TE77 for a sphere-on-plane configuration can be found in Table A1.1. A 20 mm diameter (and above) sphere is required to simulate a close enough maximum contact pressure on the TE77 and this is more than the 6 mm & 10 mm diameter ball the TE77 is configured for. Also, the resulting contact area radius simulated is much lower than that of the Charite Lumbar Spinal Implant bearing surface. Moreover, such a low load could make the friction calculation unstable. For these reasons, a disc-on-plate configuration was decided upon.

Table A1.1: Load and Contact Mechanics for a TE77 sphere-on-plane configuration

Sphere Diameter, mm	Force, N	Maximum Contact Pressure, MPa	Contact Area Radius, mm
6	1	25.6	0.137
10	1	18.2	0.162
20	1	11.5	0.204

Appendix A2. MATLAB Codes

A2.1 PCA Code

```
load test1;
W=A(:,:); % Each row is a sample
CRate=0.95; %Contribution rate

X = normalize(W);
C=cov(X); %Calculate covariance matrix

[V, D] =eig(C); %Calculate eigenvector V, eigenvalue D
[dummy, order] =sort(diag(D),'descend'); %Sort feature vectors in descending order
V=V(:,order); %Arrange feature vectors in descending order according to the size of feature
values
d=diag(D); %Take out the eigenvalues to form a column vector
newd=d(order); %Arrange the column vectors of eigenvalues in descending order

%Take the first n eigenvectors to form the transformation matrix
sumd=sum(newd); %Sum of eigenvalues
for j=1: length(newd)
   i=sum(newd(1:j,1))/sumd;%Calculate contribution rate, contribution rate = sum of the first n
eigenvalues / sum of the total eigenvalues
   if i>CRate; %When the contribution rate is greater than 95%, the cycle ends and how many
eigenvalues are taken
       cols=j;
       break;
  end
end
T=V(:,1:cols); %Take the first 'cols' eigenvectors to form the transformation matrix T
newX=X*T;  %Dimension reduction of X by transformation matrix T

figure;
percent_explained = 100*newd/sum(newd); %cumsum(latent)./sum(latent)
bar(percent_explained);
xlabel('Principal Component');
ylabel('Variance Explained (%)');
save newX1   %Save the dimension reduced data of PCA to 'newX.mat'.
```

A2.2 BP Neural Network Code

```
tic %start of timing
load newX1.mat; % Import feature data after dimension reduction by PCA
load label.mat; % Import label
% Generating test data
X=newX(:,:);
label=labels(:,:);

C_O=randperm(24075); % Randomly scramble data
% 85% is the training set
train_wine = X(C_O(1:20460),:);
% The label of the corresponding training set is also separated
train_wine_labels = label(C_O(1:20460),:);
% 15% is the test set
test_wine = X(C_O(20461:24075),:);
 % The label of the corresponding test set is also separated
test_wine_labels = label(C_O(20461:24075),:);
X=train_wine';  % Give the train set to X
Y=train_wine_labels'; % Give the labels of train set to Y
Xt=test_wine'; % Give the test set to Xt
Yt=test_wine_labels'; % Give the label of test set to Yt
%Create neural network
net = newff( minmax(X) , [10 1] , { 'logsig' 'purelin' } , 'trainrp' )  ;  %The number of nodes in
the hidden layer is 10, and the number of nodes in the output layer is 1.
```

BP network training

```
% Setting up training parameters
net.trainparam.show = 50 ;%Display frequency, set here to display once every 50 trainings
net.trainparam.epochs = 1000 ; % 1000 training times
net.trainparam.goal = 0.01 ; % Minimum error of training target is set to 0.01
net.trainParam.lr = 0.01 ;% The learning rate is set to 0.01
net.divideFcn = 'dividerand'


% starts training
net = train( net,X,Y ) ; %Train BP Neural Network with the Features and Labels of Training
Sets
%simulation
Yd2 = sim( net , X );
Yd1=round(Yd2);
% Print training set classification accuracy
 total = length(Y);% The total number of train set which need to classify
 right = sum( Yd1== Y);% Calculate the sum of the correct number of train set classifications
 shang=right/total*100; % Calculate the accuracy of classification for train set
 disp('Print training set classification accuracy');
 str = sprintf( 'Accuracy =%g%%(%d/%d)', shang,right,total);
 disp(str);
```

```
%simulation
Yd0 = sim(net, Xt);
Yd=round(Yd0);
% Print test set classification accuracy
total = length(Yt);% The total number of test set which need to classify
right = sum(Yd== Yt); % Calculate the sum of the correct number of test set classifications
shang=right/total*100; % Calculate the accuracy of classification for test set
disp('Print test set classification accuracy');
str = sprintf('Accuracy =%g%%(%d/%d)', shang, right, total);
disp(str);
```

```
[x1,y1,t2,AUC] = perfcurve(Yt,Yd,1);
AUC
plot(x1,y1)
figure(1)
plotconfusion(Yt,Yd)
figure(2)
plotroc(Yt,Yd)
figure(3)
plotconfusion(Y,Yd1)
figure;
hold on;
plot(Yt,'+');%The actual label of test set
plot(Yd,'square'); %The predicted label of test set
xlabel('Testing set sample','FontSize',12);
ylabel('Category label','FontSize',12);
legend('The actual classification','The prediction classification');
title('The actual and prediction classification of the testing set','FontSize',12);
grid on;
toc
t=toc
```

A2.3 CNN Deep Transfer Learning Code

create data files and view signals

```
parentFolder = '../DeepLearningClassification'
dataFolder = 'data'
helperCreateAEDirectories(AEData,parentFolder,dataFolder)
```

create rgb scalogram

```
helperCreateRGBfromTF(AEData,parentFolder,dataFolder) % function for creating rgb forms of wavelet transforms
```

divide into training and testing data set

```
allImages = imageDatastore(fullfile(parentFolder,dataFolder),...
    'IncludeSubfolders',true,...
    'LabelSource','foldernames');
rng default
[imgsTrain,imgsValidation] = splitEachLabel(allImages,0.8,'randomized');
disp(['Number of training images: ',num2str(numel(imgsTrain.Files))]);
disp(['Number of validation images: ',num2str(numel(imgsValidation.Files))]);
```

GoogLeNet

```
net = googlenet;
lgraph = layerGraph(net);
numberOfLayers = numel(lgraph.Layers);
figure('Units','normalized','Position',[0.1 0.1 0.8 0.8]);
plot(lgraph)
title(['GoogLeNet Layer Graph: ',num2str(numberOfLayers),' Layers']);
net.Layers(1)
```

Edit Network layers

```
lgraph = removeLayers(lgraph,{'pool5-drop_7x7_s1','loss3-classifier','prob','output'});
numClasses = numel(categories(imgsTrain.Labels));
newLayers = [
    dropoutLayer(0.6,'Name','newDropout')
        fullyConnectedLayer(numClasses,'Name','fc','WeightLearnRateFactor',10,'BiasLearnRateFactor',
10)
    softmaxLayer('Name','softmax')
    classificationLayer('Name','classoutput')];
lgraph = addLayers(lgraph,newLayers);
lgraph = connectLayers(lgraph,'pool5-7x7_s1','newDropout');
inputSize = net.Layers(1).InputSize;
```

Training Options

```
options = trainingOptions('sgdm',...
    'MiniBatchSize',28,...
    'MaxEpochs',30,...
    'InitialLearnRate',1e-4,...
    'Shuffle','every-epoch',...
    'ValidationData', imgsValidation,...
    'ValidationFrequency',20,...
    'Verbose',1,...
    'ExecutionEnvironment','cpu',...
    'Plots','training-progress');
```

```
% train network
rng default
trainedGN = trainNetwork(imgsTrain,lgraph,options);
```

inspect network

```
trainedGN.Layers(end-2:end)
cNames = trainedGN.Layers(end).ClassNames
```

```
% accuracy
[YPred,probs] = classify(trainedGN,imgsValidation);
accuracy = mean(YPred==imgsValidation.Labels);
display(['GoogLeNet Accuracy: ',num2str(accuracy)])
```

A2.4 Code for extracting features from GoogLeNet for training on SVM

```
layer = 'pool5-7x7_s1';

featuresTrain2 = activations(trainedGN,allImages,layer,'OutputAs','rows'); % extract feature representation of training images
featuresTest2 = activations(trainedGN,testImages,layer,'OutputAs','rows'); % extract feature representation of test images


YTrain = allImages.Labels;
YTest = testImages.Labels;
```

train SVM with generated function

```
[trainedClassifier, validationAccuracy] =
trainGaussianSVMClassifier(featuresTrain,YTrain)
```

predict test images using trained SVM

```
yfit = trainedGaussSVMModel.predictFcn(featuresTest);
accuracy = mean(yfit==YTest);
display(['Test Accuracy: ',num2str(accuracy)])
```

confusion matrix

```
plotconfusion(YTest,yfit)
```

Bibliography

Abu-Amer, Yousef, Isra Darwech, and John C. Clohisy. 2007. “Aseptic Loosening of Total Joint Replacements: Mechanisms Underlying Osteolysis and Potential Therapies.” *Arthritis Research and Therapy* 9 (SUPPL.1): 1–7. https://doi.org/10.1186/ar2170.

Agarap, Abien Fred. 2017. “An Architecture Combining Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for Image Classification.” http://arxiv.org/abs/1712.03541.

Agcaoglu, Serife, and Ozan Akkus. 2013. “Acoustic Emission Based Monitoring of the Microdamage Evolution During Fatigue of Human Cortical Bone.” *Journal of Biomechanical Engineering* 135 (8): 081005. https://doi.org/10.1115/1.4024134.

Aggarwal, Ravi, Viknesh Sounderajah, Guy Martin, Daniel S.W. Ting, Alan Karthikesalingam, Dominic King, Hutan Ashrafian, and Ara Darzi. 2021. “Diagnostic Accuracy of Deep Learning in Medical Imaging: A Systematic Review and Meta-Analysis.” *Npj Digital Medicine* 4 (1). https://doi.org/10.1038/s41746-021-00438-z.

Aggelis, Dimitrios G., Maria Strantza, Olivia Louis, Frans Boulpaep, Demosthenes Polyzos, and Danny van Hemelrijck. 2015. “Fracture of Human Femur Tissue Monitored by Acoustic Emission Sensors.” *Sensors (Switzerland)* 15 (3): 5803–19. https://doi.org/10.3390/s150305803.

Ai, Qingsong, Quan Liu, Wei Meng, and Sheng Quan Xie. 2018. *Neuromuscular Signal Acquisition and Processing*. *Advanced Rehabilitative Technology*. https://doi.org/10.1016/b978-0-12-814597-5.00003-5.

Alnaimat, F. A., D. E.T. Shepherd, and K. D. Dearn. 2016. “The Effect of Synthetic Polymer Lubricants on the Friction between Common Arthroplasty Bearing Biomaterials for

Encapsulated Spinal Implants." *Tribology International* 98: 20–25. https://doi.org/10.1016/j.triboint.2016.02.014.

Arun, Mike W J, Narayan Yoganandan, Brian D. Stemper, and Frank A. Pintar. 2014. "A Methodology to Condition Distorted Acoustic Emission Signals to Identify Fracture Timing from Human Cadaver Spine Impact Tests." *Journal of the Mechanical Behavior of Biomedical Materials* 40: 156–60. https://doi.org/10.1016/j.jmbbm.2014.08.023.

Asamene, Kassahun, and Mannur Sundaresan. 2012. "Analysis of Experimentally Generated Friction Related Acoustic Emission Signals." *Wear* 296: 607–18. https://doi.org/10.1016/j.wear.2012.07.019.

ASTM International. 2017. "F732-17(2017) Standard Test Method for Wear Testing of Polymeric Materials Used in Total Joint Prostheses." *ASTM International*. United States. https://doi.org/10.1520/F0732-17.

Baccar, D., and D. Söffker. 2015. "Wear Detection by Means of Wavelet-Based Acoustic Emission Analysis." *Mechanical Systems and Signal Processing* 60: 198–207. https://doi.org/10.1016/j.ymssp.2015.02.012.

Baccar, D, and D Soffker. 2013. "Application of Acoustic Emission Technique for Online Evaluation and Classification of Wear State." In *Structural Health Monitoring 2013, Volume 1 and 2- A Roadmap to Intelligent*, 1218–25. DEStech Publications.

Barão, Valentim A.R., Remya Ampadi Ramachandran, Adaías Oliveira Matos, Ravindra Vijaykumar Badhe, Carlos R. Grandini, Cortino Sukotjo, Didem Ozevin, and Mathew Mathew. 2021. "Prediction of Tribocorrosion Processes in Titanium-Based Dental Implants Using Acoustic Emission Technique: Initial Outcome." *Materials Science and Engineering C* 123 (February). https://doi.org/10.1016/j.msec.2021.112000.

Baykal, D., R. S. Siskey, H. Haider, V. Saikko, T. Ahlroos, and S. M. Kurtz. 2014. “Advances in Tribological Testing of Artificial Joint Biomaterials Using Multidirectional Pin-on-Disk Testers.” *Journal of the Mechanical Behavior of Biomedical Materials* 31: 117–34. https://doi.org/10.1016/j.jmbbm.2013.05.020.

Belyi, V A, O V Kholodilov, and A I Sviridyonok. 1981. “ACOUSTIC SPECTROMETRY AS USED FOR THE EVALUATION OF TRIBOLOGICAL SYSTEMS.” *Wear* 69: 309–19.

Boness, R J, and S L McBride. 1991. “Adhesive and Abrasive Wear Studies Using Acoustic Emission Techniques.” *Wear* 149: 41–53.

Boness, RJ, S L McBride, and M Sobczyk. 1990. “Wear Studies Using Acoustic Emission Techniques.” *Tribology International* 23 (5): 291–95. https://doi.org/10.1016/0301-679X(90)90001-6.

Boon, Maurits J.G.N., Dimitrios Zarouchas, Marcias Martinez, Daniel Gagar, Rinze Benedictus, and Peter Foote. 2014. “Temperature and Load Effects on Acoustic Emission Signals for Structural Health Monitoring Applications.” *7th European Workshop on Structural Health Monitoring, EWSHM 2014 - 2nd European Conference of the Prognostics and Health Management (PHM) Society*, 1997–2003.

Briscoe, B.J., A. Chateauminois, J. Chiu, and S. Vickery. 2001. “Acoustic Noise Emission in a Model PMMA/Steel Fretting Contact.” *Tribology Series* 39: 673–81. https://doi.org/10.1016/s0167-8922(01)80149-8.

Brown, Tim, Qi Bin Bao, C. Mauli Agrawal, and Nadim James Hallab. 2011. “An in Vitro Assessment of Wear Particulate Generated from Nubac: A Peek-on-Peek Articulating Nucleus Replacement Device: Methodology and Results from a Series of Wear Tests Using Different Motion Profiles, Test Frequencies, and Environmental Conditions.” *Spine*

36 (26): 1675–85. https://doi.org/10.1097/BRS.0b013e31821ac8a0.

Bryant, M, R Farrar, R Freeman, K Brummitt, J Nolan, and A Neville. 2014. "Galvanically Enhanced Fretting-Crevice Corrosion of Cemented Femoral Stems." *Journal of the Mechanical Behavior of Biomedical Materials* 40: 275–86. https://doi.org/10.1016/j.jmbbm.2014.08.021.

BS ISO 14242-1. 2014. "BSI Standards Publication Implants for Surgery — Wear of Total Hip-Joint Disc Prostheses Part 1 : Loading and Displacement Parameters for Wear-Testing Machines and Corresponding Environmental Conditions for Test."

BS ISO 18192-1. 2011. "BSI Standards Publication Implants for Surgery — Wear of Total Intervertebral Spinal Disc Prostheses Part 1 : Loading and Displacement Parameters for Wear Testing and Corresponding Environmental Conditions for Test."

Cadot, Christophe, Jean-François Saillant, and Bernard Dulmet. 2016. "Method for Acoustic Characterization of Materials in Temperature." *Wcndt 2016*, 1–9.

Chen, S., S.A. Billings, and P.M. Grant. 1989. "Non-Linear Systems Identification Using Neural Networks." *Research Report. Acse Report 370*. https://eprints.whiterose.ac.uk/78225/.

Choudhury, Dipankar, Robert Walker, Taposh Roy, Sweety Paul, and Rajshree Mootanah. 2013. "Performance of Honed Surface Profiles to Artificial Hip Joints: An Experimental Investigation." *International Journal of Precision Engineering and Manufacturing* 14 (10): 1847–53. https://doi.org/10.1007/s12541-013-0247-z.

Crivelli, Davide, Mario Guagliano, and Alberto Monici. 2014. "Development of an Artificial Neural Network Processing Technique for the Analysis of Damage Evolution in Pultruded Composites with Acoustic Emission." *Composites Part B* 56: 948–59.

https://doi.org/10.1016/j.compositesb.2013.09.005.

Curry, B., and D. Rumelhart. 1990. “MSnet: A Neural Network Which Classifies Mass Spectra.” *Tetrahedron Computer Methodology* 3 (3–4): 213–37.

Curry, Bo, and David E. Rumelhart. 1990. “MSnet: A Neural Network Which Classifies Mass Spectra.” *Tetrahedron Computer Methodology* 3 (3–4): 213–37. https://doi.org/10.1016/0898-5529(90)90053-B.

David, Thierry. 2007. “Long-Term Results of One-Level Lumbar Arthroplasty: Minimum 10-Year Follow-up of the CHARITÉ Artificial Disc in 106 Patients.” *Spine* 32 (6): 661–66. https://doi.org/10.1097/01.brs.0000257554.67505.45.

Devin, Clinton J., Thomas G. Myers, and James D. Kang. 2008. “Chronic Failure of a Lumbar Total Disc Replacement with Osteolysis: Report of a Case with Nineteen-Year Follow-Up.” *Journal of Bone and Joint Surgery - Series A* 90 (10): 2230–34. https://doi.org/10.2106/JBJS.G.01712.

Dowson, Duncan. 2012. “Bio-Tribology,” 9–30. https://doi.org/10.1039/c2fd20103h.

Ech-Choudany, Y., M. Assarar, D. Scida, F. Morain-Nicolier, and B. Bellach. 2017. “Unsupervised Clustering for Building a Learning Database of Acoustic Emission Signals to Identify Damage Mechanisms in Unidirectional Laminates.” *Applied Acoustics* 123: 123–32. https://doi.org/10.1016/j.apacoust.2017.03.008.

Eckold, D. G., K. D. Dearn, and D. E T Shepherd. 2015. “The Evolution of Polymer Wear Debris from Total Disc Arthroplasty.” *Biotribology* 1–2: 42–50. https://doi.org/10.1016/j.biotri.2015.04.002.

Engh, Gerard A., Rebecca L. Zimmerman, Nancy L. Parks, and C. Anderson Engh. 2009. “Analysis of Wear in Retrieved Mobile and Fixed Bearing Knee Inserts.” *Journal of*

Arthroplasty 24 (6 SUPPL.): 28–32. https://doi.org/10.1016/j.arth.2009.03.010.

Ferrer, C., F. Salas, M. Pascual, and J. Orozco. 2010. "Discrete Acoustic Emission Waves during Stick-Slip Friction between Steel Samples." *Tribology International* 43 (1–2): 1–6. https://doi.org/10.1016/j.triboint.2009.02.009.

Fisher, J., D. Dowson, H. Hamdzah, and H. L. Lee. 1994. "The Effect of Sliding Velocity on the Friction and Wear of UHMWPE for Use in Total Artificial Joints." *Wear* 175 (1–2): 219–25. https://doi.org/10.1016/0043-1648(94)90185-6.

FitzPatrick, A.J., G.W. Rodgers, G.J. Hooper, and T.B.F. Woodfield. 2017. "Development and Validation of an Acoustic Emission Device to Measure Wear in Total Hip Replacements In-Vitro and in-Vivo." *Biomedical Signal Processing and Control* 33: 281–88. https://doi.org/10.1016/j.bspc.2016.12.011.

Foresee, F Dan, and Martin T Hagan. 1997. "GAUSS-NEWTON APPROXIMATION TO BAYESIAN LEARNING." In *International Conference on Neural Networks (ICNN'97)*, 1930–35. https://doi.org/10.1109/ICNN.1997.614194.

Franke, R P, P Dorner, H. J. Schwalbe, and B Ziegler. 2004. "Acoustic Emission Measurement System for the Orthopedical Diagnostics of the Human Femur and Knee Joint." *Journal of Acoustic Emission* 22: 236–42.

Geetha, M., A. K. Singh, R. Asokamani, and A. K. Gogia. 2009. "Ti Based Biomaterials, the Ultimate Choice for Orthopaedic Implants - A Review." *Progress in Materials Science* 54 (3): 397–425. https://doi.org/10.1016/j.pmatsci.2008.06.004.

Grecu, Dan, Iulian Antoniac, Octavian Trante, Marius Niculescu, and Olivera Lupescu. 2016. "Failure Analysis of Retrieved Polyethylene Insert in Total Knee Replacement." *Materiale Plastice* 53 (4): 776–80.

Gutkin, R., C. J. Green, S. Vangrattanachai, S. T. Pinho, P. Robinson, and P. T. Curtis. 2011. "On Acoustic Emission for Failure Investigation in CFRP: Pattern Recognition and Peak Frequency Analyses." *Mechanical Systems and Signal Processing* 25 (4): 1393–1407. https://doi.org/10.1016/j.ymssp.2010.11.014.

Hase, Alan, Hiroshi Mishina, and Masaki Wada. 2012. "Correlation between Features of Acoustic Emission Signals and Mechanical Wear Mechanisms." *Wear* 292–293: 144–50. https://doi.org/10.1016/j.wear.2012.05.019.

———. 2013. "Microscopic Study on the Relationship between AE Signal and Wear Amount." *Wear* 308: 142–47. https://doi.org/10.1016/j.wear.2013.08.005.

Hase, Alan, Masaki Wada, and Hiroshi Mishina. 2008. "Acoustic Emission Signals and Wear Phenomena on Severe-Mild Wear Transition." *Tribology Online* 3 (5): 298–303. https://doi.org/10.2474/trol.3.298.

Hellier, C J. 2003. "Chapter 10: Acoustic Emission Testing." In *Handbook of Nondestructive Evaluation,* 10.1-10.39. Boston: The McGraw-Hill Companies, Inc.

Hoskins, T J, K D Dearn, S N Kukureka, and D Walton. 2011. "Acoustic Noise from Polymer Gears – A Tribological Investigation." *Materials and Design* 32 (6): 3509–15. https://doi.org/10.1016/j.matdes.2011.02.041.

Hossain, Fahad, Shelain Patel, and Fares Sami Haddad. 2010. "Midterm Assessment of Causes and Results of Revision Total Knee Arthroplasty." *Clinical Orthopaedics and Related Research* 468 (5): 1221–28. https://doi.org/10.1007/s11999-009-1204-0.

Howell, J. R., L. A. Blunt, C. Doyle, R. M. Hooper, A. J.C. Lee, and R. S.M. Ling. 2004. "In Vivo Surface Wear Mechanisms of Femoral Components of Cemented Total Hip Arthroplasties: The Influence of Wear Mechanism on Clinical Outcome." *Journal of*

Arthroplasty 19 (1): 88–101. https://doi.org/10.1016/S0883-5403(03)00278-X.

Hua, Zikai, Yongwei Fan, and Zhongmin Jin. 2014. "A Biotribo-Acoustic Testing Method for Ceramic Orthopaedic Biomaterials." *Tribology International* 71: 1–6. https://doi.org/10.1016/j.triboint.2013.10.015.

Huang, Shao Fu, Wan Rung Chen, and Chun Li Lin. 2016. "Biomechanical Interactions of Endodontically Treated Tooth Implant-Supported Prosthesis under Fatigue Test with Acoustic Emission Monitoring." *BioMedical Engineering Online* 15 (1): 1–10. https://doi.org/10.1186/s12938-016-0140-y.

Jiaa, C L, and D A Dornfeld. 1990. "Experimental Studies of Sliding Friction and Wear via Acoustic Emission Signal Analysis." *Wear* 39.

Jin, Z. M., M. Stone, E. Ingham, and J. Fisher. 2006. "(V) Biotribology." *Current Orthopaedics* 20 (1): 32–40. https://doi.org/10.1016/j.cuor.2005.09.005.

Jin, Z.M., J. Zheng, W. Li, and Z.R. Zhou. 2016. "Tribology of Medical Devices." *Biosurface and Biotribology* 2 (4): 173–92. https://doi.org/10.1016/j.bsbt.2016.12.001.

Johnson, K.L. 1985. *Contact Mechanics*. Cambridge: Cambridge University Press.

Kanaga Karuppiah, K. S., Angela L. Bruck, Sriram Sundararajan, Jun Wang, Zhiqun Lin, Zhi Hui Xu, and Xiaodong Li. 2008. "Friction and Wear Behavior of Ultra-High Molecular Weight Polyethylene as a Function of Polymer Crystallinity." *Acta Biomaterialia* 4 (5): 1401–10. https://doi.org/10.1016/j.actbio.2008.02.022.

Kapur, Richard A. 2016. "Acoustic Emission in Orthopaedics: A State of the Art Review." *Journal of Biomechanics* 49 (16): 4065–72. https://doi.org/10.1016/j.jbiomech.2016.10.038.

Karl, Knahr. 2013. "Tribology and Total Hip Arthroplasty Implants." *Orthopedics* 36 (11):

854–55. https://doi.org/10.3928/01477447-20131021-05.

Khan-Edmundson, Ataif, Geoffrey W. Rodgers, Tim B F Woodfield, Gary J. Hooper, and J. Geoffrey Chase. 2012. “Tissue Attenuation Characteristics of Acoustic Emission Signals for Wear and Degradation of Total Hip Arthroplasty Implants.” In *IFAC Proceedings Volumes (IFAC-PapersOnline)*, 45:355–60. IFAC. https://doi.org/10.3182/20120829-3-HU-2029.00046.

Khan, Tawhidul Islam, and Harino Yoho. 2016. “Integrity Analysis of Knee Joint by Acoustic Emission Technique.” *Journal on Multimodal User Interfaces* 10 (4): 319–24. https://doi.org/10.1007/s12193-015-0206-3.

Kleinbaum, David G., and Mitchel Klein. 2002. *Logistic Regression: A Self-Learning Text. Survival*. Second Edi. Springer.

König, F., C. Sous, A. Ouald Chaib, and G. Jacobs. 2021. “Machine Learning Based Anomaly Detection and Classification of Acoustic Emission Events for Wear Monitoring in Sliding Bearing Systems.” *Tribology International* 155 (November 2020). https://doi.org/10.1016/j.triboint.2020.106811.

Lee, Christine, Lu Zhang, Dalton Morris, Kai Yuan Cheng, Remya Ampadi Ramachandran, Mark Barba, Divya Bijukumar, Didem Ozevin, and Mathew T. Mathew. 2021. “Non-Invasive Early Detection of Failure Modes in Total Hip Replacements (THR) via Acoustic Emission (AE).” *Journal of the Mechanical Behavior of Biomedical Materials* 118 (February 2020): 104484. https://doi.org/10.1016/j.jmbbm.2021.104484.

Leuridan, Steven, Quentin Goossens, Tom Vander Sloten, Koen De Landsheer, Hendrik Delport, Leonard Pastrav, Kathleen Denis, Wim Desmet, and Jos Vander Sloten. 2017. “Vibration-Based Fixation Assessment of Tibial Knee Implants: A Combined in Vitro and in Silico Feasibility Study.” *Medical Engineering and Physics* 49: 109–20.

https://doi.org/10.1016/j.medengphy.2017.08.007.

Li, Dongsheng, Jinping Ou, Chengming Lan, and Hui Li. 2012. “Monitoring and Failure Analysis of Corroded Bridge Cables under Fatigue Loading Using Acoustic Emission Sensors.” *Sensors* 12 (4): 3901–15. https://doi.org/10.3390/s120403901.

Lingard, S, C W Yu, and C F Yau. 1993. “Short Communication Sliding Wear Studies Using Acoustic Emission.” *Wear* 164: 597–604.

Liu, Gongde, Ming Xiang, and Huilin Li. 2004. “A Study on Sliding Wear of Ultrahigh Molecular Weight Polyethylene / Polypropylene Blends” 44 (1): 197–208.

Lorena, Ana C., Luis F.O. Jacintho, Marinez F. Siqueira, Renato De Giovanni, Lúcia G. Lohmann, André C.P.L.F. De Carvalho, and Missae Yamamoto. 2011. “Comparing Machine Learning Classifiers in Potential Distribution Modelling.” *Expert Systems with Applications* 38 (5): 5268–75. https://doi.org/10.1016/j.eswa.2010.10.031.

Lu, Shi bao, Yong Hai, Chao Kong, Qing yi Wang, Qingjun Su, Lei Zang, Nan Kang, Xiang long Meng, and Yu Wang. 2015. “An 11-Year Minimum Follow-up of the Charite III Lumbar Disc Replacement for the Treatment of Symptomatic Degenerative Disc Disease.” *European Spine Journal* 24 (9): 2056–64. https://doi.org/10.1007/s00586-015-3939-5.

Lundervold, Alexander Selvikvåg, and Arvid Lundervold. 2019. “An Overview of Deep Learning in Medical Imaging Focusing on MRI.” *Zeitschrift Fur Medizinische Physik* 29 (2): 102–27. https://doi.org/10.1016/j.zemedi.2018.11.002.

Malham, Gregory M., and Rhiannon M. Parker. 2017. “Early Experience with Lateral Lumbar Total Disc Replacement: Utility, Complications and Revision Strategies.” *Journal of Clinical Neuroscience* 39 (September 2011): 176–83.

https://doi.org/10.1016/j.jocn.2017.01.033.

Masmoudi, Sahir, Abderrahim El Mahi, and Sa??d Turki. 2015. “Use of Piezoelectric as Acoustic Emission Sensor for in Situ Monitoring of Composite Structures.” *Composites Part B: Engineering* 80: 307–20. https://doi.org/10.1016/j.compositesb.2015.06.003.

Mavrogordato, Mark, Mark Taylor, Andrew Taylor, and Martin Browne. 2011. “Real Time Monitoring of Progressive Damage during Loading of a Simplified Total Hip Stem Construct Using Embedded Acoustic Emission Sensors.” *Medical Engineering and Physics* 33 (4): 395–406. https://doi.org/10.1016/j.medengphy.2010.10.025.

McCrory, John P., Safaa K. Al-Jumaili, Davide Crivelli, Matthew R. Pearson, Mark J. Eaton, Carol A. Featherston, Mario Guagliano, Karen M. Holford, and Rhys Pullin. 2015. “Damage Classification in Carbon Fibre Composites Using Acoustic Emission: A Comparison of Three Techniques.” *Composites Part B: Engineering* 68: 424–30. https://doi.org/10.1016/j.compositesb.2014.08.046.

Mishina, H., and Alan. Hase. 2019. “Effect of the Adhesion Force on the Equation of Adhesive Wear and the Generation Process of Wear Elements in Adhesive Wear of Metals.” *Wear*, no. 432–433.

Mishina, H., and Alan Hase. 2013. “Wear Equation for Adhesive Wear Established through Elementary Process of Wear.” *Wear* 308 (1–2): 186–92.

Moghadas, Parshia M., Duncan E T Shepherd, David W L Hukins, and Aziza Mahomed. 2012. “Polymer-on-Metal or Metal-on-Polymer Total Disc Arthroplasty: Does It Make a Difference?” *Spine* 37 (21): 1834–38. https://doi.org/10.1097/BRS.0b013e318257fdd3.

Moghadas, Parshia, Aziza Mahomed, David W.L. Hukins, and Duncan E.T. Shepherd. 2012. “Friction in Metal-on-Metal Total Disc Arthroplasty: Effect of Ball Radius.” *Journal of*

Biomechanics 45 (3): 504–9. https://doi.org/10.1016/j.jbiomech.2011.11.045.

Moghadas, Parshia, Aziza Mahomed, Duncan E T Shepherd, and David W L Hukins. 2015. "Wear of the Charité® Lumbar Intervertebral Disc Replacement Investigated Using an Electro-Mechanical Spine Simulator." *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of Engineering in Medicine* 229 (3): 264–68. https://doi.org/10.1177/0954411915576537.

Momon, S., N. Godin, P. Reynaud, M. R'Mili, and G. Fantozzi. 2012. "Unsupervised and Supervised Classification of AE Data Collected during Fatigue Test on CMC at High Temperature." *Composites Part A: Applied Science and Manufacturing* 43 (2): 254–60. https://doi.org/10.1016/j.compositesa.2011.10.016.

Mukhopadhyay, C. K., T. Jayakumar, Baldev Raj, and S. Venugopal. 2012. "Statistical Analysis of Acoustic Emission Signals Generated During Turning of a Metal Matrix Composite Acoustic." *Journal of the Brazilian Society of Mechanical Sciences and Engineering* XXXIV (2): 145–54. https://doi.org/10.1590/S1678-58782012000200006.

Nakahara, Ichiro, Masaki Takao, Shunichi Bandoh, Nicky Bertollo, William R. Walsh, and Nobuhiko Sugano. 2013. "In Vivo Implant Fixation of Carbon Fiber-Reinforced PEEK Hip Prostheses in an Ovine Model." *Journal of Orthopaedic Research* 31 (3): 485–92. https://doi.org/10.1002/jor.22251.

Narendra, Kumpati S., and Kannan Parthasarathy. 1991. "Learning Automata Approach to Hierarchical Multiobjective Analysis." *IEEE Transactions on Systems, Man and Cybernetics*. https://doi.org/10.1109/21.101158.

National Joint Registry. 2021. "NJR 18th Annual Report 2021." reports.njrcentre.org.uk/Downloads.

Niu, Xiao Xiao, and Ching Y. Suen. 2012. “A Novel Hybrid CNN-SVM Classifier for Recognizing Handwritten Digits.” *Pattern Recognition* 45 (4): 1318–25. https://doi.org/10.1016/j.patcog.2011.09.021.

Olorunlambe, K A, D E T Shepherd, and K D Dearn. 2019. “A Review of Acoustic Emission as a Biotribological Diagnostic Tool.” *Tribology - Materials, Surfaces & Interfaces* 13 (3): 161–71. https://doi.org/10.1080/17515831.2019.1622914.

Olorunlambe, K A, D G Eckold, D E T Shepherd, and K D Dearn. 2022. "Bio-Tribo-Acoustic Emissions: Condition Monitoring of a Simulated Joint Articulation." *Biotribology*. https://doi.org/10.1016/j.biotri.2022.100217.

Olorunlambe, Khadijat A., Zhe Hua, Duncan E. T. Shepherd, and Karl D. Dearn. 2021. “Towards a Diagnostic Tool for Diagnosing Joint Pathologies: Supervised Learning of Acoustic Emission Signals.” *Sensors* 21 (23): 8091. https://doi.org/10.3390/s21238091.

Ossi, Zannar, Wael Abdou, Robert L. Reuben, and Richard J. Ibbetson. 2012. “In Vitro Assessment of Bone-Implant Interface Using an Acoustic Emission Transmission Test.” *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* 226 (1): 63–69. https://doi.org/10.1177/0954411911428696.

———. 2013. “Transmission of Acoustic Emission in Bones, Implants and Dental Materials.” *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* 227 (11): 1237–45. https://doi.org/10.1177/0954411913500204.

Ouyang, Huei Tau. 2017. “Nonlinear Autoregressive Neural Networks with External Inputs for Forecasting of Typhoon Inundation Level.” *Environmental Monitoring and Assessment* 189 (8). https://doi.org/10.1007/s10661-017-6100-6.

Patzer, Gregor, and Mathias Woydt. 2021. “New Methodologies Indicating Adhesive Wear in

Load Step Tests on the Translatory Oscillation Tribometer." *Lubricants* 9 (10). https://doi.org/10.3390/lubricants9100101.

Petrica, Marcela, Bernadette Duscher, Thomas Koch, and Vasiliki Maria Archodoulaki. 2016. "Studies on Tribological Behavior of PEEK and PE-UHMW." *AIP Conference Proceedings* 1779 (October 2016). https://doi.org/10.1063/1.4965533.

Polat, Çiğdem, and Mehmet Siraç. 2018. "Introduction to Wavelets and Their Applications in Signal Denoising." *Journal of Science and Technology* 8 (1): 1–10.

Price, E. D., A. W. Lees, and M. I. Friswell. 2005. "Detection of Severe Sliding and Pitting Fatigue Wear Regimes through the Use of Broadband Acoustic Emission." *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology* 219 (2): 85–98. https://doi.org/10.1243/135065005X9817.

Puloski, S. K.T., R. W. McCalden, S. J. MacDonald, C. H. Rorabeck, and R. B. Bourne. 2001. "Tibial Post Wear in Posterior Stabilized Total Knee Arthroplasty." *Journal of Bone and Joint Surgery - Series A* 83 (3): 390–97. https://doi.org/10.2106/00004623-200103000-00011.

Punt, Ilona M., Violette M. Visser, Lodewijk W. Van Rhijn, Steven M. Kurtz, Jop Antonis, Geert Willem H. Schurink, and André Van Ooij. 2008. "Complications and Reoperations of the SB Charité Lumbar Disc Prosthesis: Experience in 75 Patients." *European Spine Journal* 17 (1): 36–43. https://doi.org/10.1007/s00586-007-0506-8.

Qiao, X., W. Weng, and Q. Li. 2019. "Acoustic Emission Monitoring and Failure Behavior Discrimination of 8YSZ Thermal Barrier Coatings under Vickers Indentation Testing." *Surface and Coatings Technology*, 913–22.

Rao, A.K. 1990. "Acoustic Emission and Signal Analysis." *Defence Science Journal* 40 (1):

55–70. http://www.publications.drdo.gov.in/ojs/index.php/dsj/article/view/4450.

Ray, Sunil. 2015. “8 Proven Ways for Boosting the ‘Accuracy’ of a Machine Learning Model.” Analytics Vidhya. 2015.

Reeks, John, and Hong Liang. 2015. “Materials and Their Failure Mechanisms in Total Disc Replacement.” *Lubricants* 3 (2): 346–64. https://doi.org/10.3390/lubricants3020346.

Ren, Beining, and Jinbin Chen. 2021. “Fracture Acoustic Emission Signals Identification of Broken Wire Using Deep Transfer Learning and Wavelet Analysis.” *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3469213.3470230.

Rodgers, Geoffrey W., Roger Welsh, Laura J. King, Anthony J. FitzPatrick, Tim B F Woodfield, and Gary J. Hooper. 2017. “Signal Processing and Event Detection of Hip Implant Acoustic Emissions.” *Control Engineering Practice* 58 (May 2015): 287–97. https://doi.org/10.1016/j.conengprac.2016.09.013.

Rodgers, Geoffrey W., Jade L. Young, Anna V. Fields, Riki Z. Shearer, Tim. B.F. Woodfield, Gary J. Hooper, and J. Geoffrey Chase. 2014. “Acoustic Emission Monitoring of Total Hip Arthroplasty Implants.” In *IFAC Proceedings Volumes*, 47:4796–4800. IFAC. https://doi.org/10.3182/20140824-6-ZA-1003.00928.

Rousseeuw, P. 1987. “Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis.” *Journal of Computational and Applied Mathematics* 20: 53–65.

Rowland, Chris, Martin Browne, Andy Taylor, and U.K. Leatherhead. 2004. “Dynamic Health Monitoring of Metal on Metal Hip Prostheses Using Acoustic Emission.” In *26th European Conference on Acoustic Emission Testing*, 465–67.

Rubtsov, V E, E A Kolubaev, A V Kolubaev, and V L Popov. 2013. “Using Acoustic Emission for the Analysis of Wear Processes during Sliding Friction.” *Technical Physics Letters* 39

(2): 223–25. https://doi.org/10.1134/S1063785013020235.

Saikko, Vesa. 1998. “A Multidirectional Motion Pin-on-Disk Wear Test Method for Prosthetic Joint Materials.” *Journal of Biomedical Materials Research* 41 (1): 58–64. https://doi.org/10.1002/(SICI)1097-4636(199807)41:1<58::AID-JBM7>3.0.CO;2-P.

———. 2017. “Effect of Contact Area on the Wear and Friction of UHMWPE in Circular Translation Pin-on-Disk Tests.” *Journal of Tribology* 139 (6): 061606. https://doi.org/10.1115/1.4036448.

Sause, Markus G R. 2011. “Investigation of Pencil-Lead Breaks as Acoustic Emission Sources.” *Journal of Acoustic Emission* 29: 184–96.

Schwalbe, H. J., G. Bamfaste, and R. P. Franke. 1999. “Non-Destructive and Non-Invasive Observation of Friction and Wear of Human Joints and of Fracture Initiation by Acoustic Emission.” *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* 213 (1): 41–48. https://doi.org/10.1243/0954411991534799.

Shark, L. K., H. Chen, and J. Goodacre. 2011. “Knee Acoustic Emission: A Potential Biomarker for Quantitative Assessment of Joint Ageing and Degeneration.” *Medical Engineering and Physics* 33 (5): 534–45. https://doi.org/10.1016/j.medengphy.2010.12.009.

Shark, Lik-Kwan, Hongzhi Chen, and John Goodacre. 2010. “Discovering Differences in Acoustic Emission between Healthy and Osteoarthritic Knees Using a Four-Phase Model of Sit-Stand-Sit Movements.” *The Open Medical Informatics Journal* 4: 116–25. https://doi.org/10.2174/1874431101004010116.

Simon, Sheldon R, Igor L Paul, Robert M Rose, and Eric L Radin. 1975. “‘Stiction-Friction’ of Total Hip Prostheses and Its Relationship to Loosening.” *The Journal of Bone and Joint*

Surgery 57-A (2).

Siskey, Ryan, Lauren Ciccarelli, Melissa K.C. Lui, and Steven M. Kurtz. 2016. "Are PEEK-on-Ceramic Bearings an Option for Total Disc Arthroplasty? An In Vitro Tribology Study." *Clinical Orthopaedics and Related Research* 474 (11): 2428–40. https://doi.org/10.1007/s11999-016-5041-7.

Siskey, Ryan, Jonathan Peck, Hitesh Mehta, Allison Kosydar, Steven Kurtz, and Genevieve Hill. 2016. "Development of a Clinically Relevant Impingement Test Method for a Mobile Bearing Lumbar Total Disc Replacement." *Spine Journal* 16 (9): 1133–42. https://doi.org/10.1016/j.spinee.2016.05.004.

Strantza, Maria, Olivia Louis, Demosthenes Polyzos, Frans Boulpaep, Danny van Hemelrijck, and Dimitrios G. Aggelis. 2014. "Wave Dispersion and Attenuation on Human Femur Tissue." *Sensors (Switzerland)* 14 (8): 15067–83. https://doi.org/10.3390/s140815067.

Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2015. "Going Deeper with Convolutions." In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1–9. https://doi.org/10.1109/CVPR.2015.7298594.

Tang, Yichuan. 2013. "Deep Learning Using Linear Support Vector Machines." http://arxiv.org/abs/1306.0239.

Tensi, Hans Maria. 2004. "The KAISER - EFFECT and Its Scientific Background Outline of This Paper Work of Joseph Kaiser." *EWGAE 2004, 26th European Conference on Acoustic Emission Testing*, 31–42. http://www.ndt.net/article/ewgae2004/html/htmltxt/l00tensi.htm.

Toen, C. Van, J. Street, T. R. Oxland, and P. A. Cripton. 2012. "Acoustic Emission Signals

Can Discriminate between Compressive Bone Fractures and Tensile Ligament Injuries in the Spine during Dynamic Loading." *Journal of Biomechanics* 45 (9): 1643–49. https://doi.org/10.1016/j.jbiomech.2012.03.025.

Unnporsson, Runar. 2013. "Hit Detection and Determination in AE Bursts." In *Acoustic Emission,* edited by Wojciech Sikorski. Rjeka: IntechOpen. https://doi.org/10.5772/54754.

Wang, Ling, Robert Wood, and Jun Sun. 2008. "Acoustic Emissions from Oil Lubricated Metal on Metal Sliding Contacts." *Insight: Non-Destructive Testing and Condition Monitoring* 50 (9): 506–11. https://doi.org/10.1784/insi.2008.50.9.506.

Wong, Kelvin K.L., Giancarlo Fortino, and Derek Abbott. 2020. "Deep Learning-Based Cardiovascular Image Diagnosis: A Promising Challenge." *Future Generation Computer Systems* 110: 802–11. https://doi.org/10.1016/j.future.2019.09.047.

Xin, H, D E T Shepherd, and K D Dearn. 2012. "PEEK (Polyether-Ether-Ketone) Based Cervical Total Disc Arthroplasty: Contact Stress and Lubrication Analysis." *The Open Biomedical Engineering Journal* 6: 73–79. https://doi.org/10.2174/1874230001206010073.

Xin, Haohui, Lu Cheng, Robin Diender, and Milan Veljkovic. 2020. "Fracture Acoustic Emission Signals Identification of Stay Cables in Bridge Engineering Application Using Deep Transfer Learning and Wavelet Analysis." *Advances in Bridge Engineering* 1 (1): 1–16. https://doi.org/10.1186/s43251-020-00006-7.

Yang, G., and W. Garrison. 1989. "A Comparison of Microstructural Effects on Two-Body and Three-Body Abrasive Wear." *Wear* 129 (1): 93–103.

Yao, Y., X. Li, and Z. Yuan. 1999. "Tool Wear Detection with Fuzzy Classification and

Wavelet Fuzzy Neural Network." *International Journal of Machine Tools and Manufacture* 39 (10): 1525–38.

Yu, Hui, David C. Samuels, Ying yong Zhao, and Yan Guo. 2019. "Architectures and Accuracy of Artificial Neural Network for Disease Classification from Omics Data." *BMC Genomics* 20 (1): 1–12. https://doi.org/10.1186/s12864-019-5546-z.

Zhang, H., L. T. Brown, L. A. Blunt, X. Jiang, and S. M. Barrans. 2009. "Understanding Initiation and Propagation of Fretting Wear on the Femoral Stem in Total Hip Replacement." *Wear* 266 (5–6): 566–69. https://doi.org/10.1016/j.wear.2008.04.076.

Zhao, Longcai, Qiangzi Li, Yuan Zhang, Hongyan Wang, and Xin Du. 2019. "Integrating the Continuous Wavelet Transform and a Convolutional Neural Network to Identify Vineyard Using Time Series Satellite Images." *Remote Sensing* 11 (22). https://doi.org/10.3390/rs11222641.

Zhou, Z.R., and Z.M. Jin. 2015. "Biotribology: Recent Progresses and Future Perspectives." *Biosurface and Biotribology* 1 (1): 3–24. https://doi.org/10.1016/j.bsbt.2015.03.001.

Zykova, Lucie, Pavel Mazal, and Lubos Pazdera. 2006. "Identification of Contact Fatigue Stages with Acoustic Emission Method." In *European Conference on Non-Destructive Testing*, 1–8.