

Transfer Learning in Classifying Acoustic Emission Signals from Osteoarthritic Knees

Nazmush Sakib, Md Abdur Rahman, Tawhidul Islam KHAN*

Biological and material engineering Graduate school of science and engineering, Saga University Saga, Japan

Md. Mehedi Hassan, Shuya Ide

Tsuruta orthopedic clinic Ushizu-machi, Ogi Saga Japan

Abstract

Acoustic emission (AE) is a non-destructive evaluation (NDE) method that allows to inspect the internal condition of material by analyzing the signal which are produced due to the internal change in the condition. Compared to the present methods, due to simplicity and immense potential, AE has gained attention in knee health assessment. With the advancement of computational power many researchers have implemented advanced machine learning (ML) algorithms to characterize the AE signals which were generated from human knees. However, most of this research are focused on implementing the unsupervised ML algorithms. The minimal variability between the AE signals from different knee conditions has posed significant challenges in implementing supervised ML algorithms which shows the promise to make the diagnosis significantly simpler than the present approaches. Therefore, this work aims at implementing transfer learning using CNN and wavelet-based images to classify the AE signals which were generated from the knees of the knee osteoarthritis of different Kellgren Lawrence (KL) grades. VGG-16 CNN model has been trained on the images which were generated from AE signals of the participants. The results shows huge promise of transfer learning in classifying the AE signals from different knee health conditions

Contribution of the Paper: This paper shows a novel application of CNN based transfer learning in diagnosis of knee OA from acoustic emission signals

Keywords: Acoustic emission, Knee OA, Transfer learning, CNN, Deep learning

IJCVSP

ISSN: 2186-1390 (Online)
http://cennser.org/IJCVSP

Article History:

Received: 20/2/2025

Revised: 16/7/2025

Accepted: 1/11/2025

Published Online: 23/11/2025

© 2012, IJCVSP, CNSER. All Rights Reserved

1. INTRODUCTION

The knee is one of the most intricate joints in the human body. It comprises three primary bones: the femur, tibia, and patella. Each of these bones is covered with articular cartilage at their ends, facilitating seamless movement and serving as a shock absorber to prevent direct bone contact. Additionally, the synovial fluid and surrounding ligaments contribute to the joint's stability and strength, while the fibrocartilaginous menisci between the tibia and femur further enhance smooth motion and structural support [1][2]. Despite its robust architecture,

the knee remains highly vulnerable to various diseases and age-related degeneration. One of the most prevalent knee disorders is osteoarthritis (OA), characterized by the progressive breakdown of articular cartilage and modifications in the subchondral bone, leading to pain and inflammation that significantly affect daily activities [3]. Diagnostic approaches for knee OA include arthroscopy, radiography, MRI, fluid analysis, and blood testing. However, each method has drawbacks: arthroscopy is an invasive surgical procedure, X-rays fail to capture dynamic knee conditions, and although MRI is highly effective, it is costly, time-intensive, and unsuitable for certain patients [4][5]. This highlights the need for an advanced diagnostic system that overcomes these limitations. The acoustic emission (AE) technique presents a promising alternative and has been

*Corresponding author

Email address: khan@cc.saga-u.ac.jp (Tawhidul Islam KHAN)

Kellgren and Lawrence (KL) Grading System

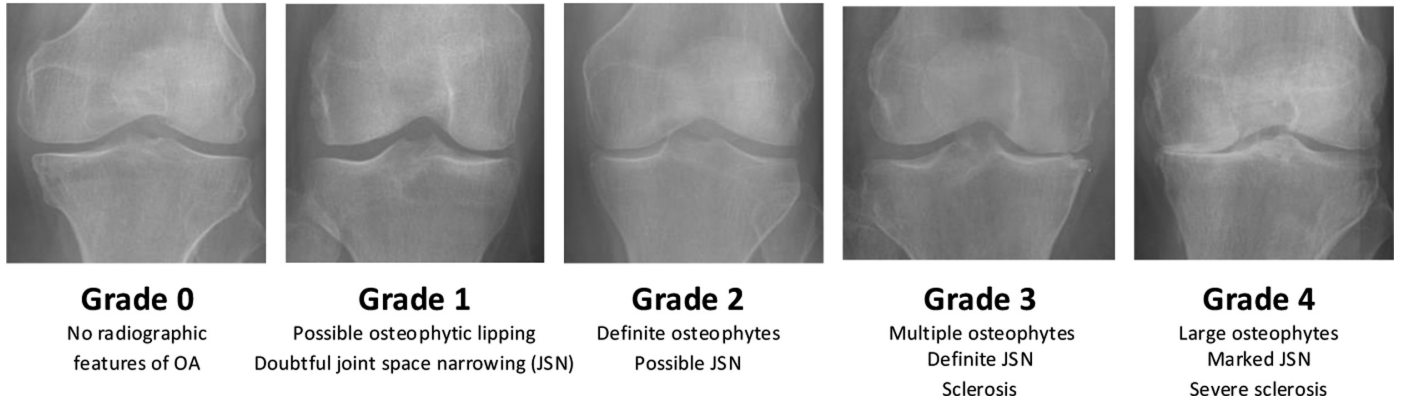


Figure 1: X ray images from differently graded knees using KL method

widely used in various scientific and engineering domains. AE captures acoustic waves generated by mechanical activities within materials, such as crack formation, grain boundary shifts, and friction between solid surfaces [6][7]. These high-frequency signals can be detected using sophisticated AE systems, providing a viable tool for monitoring mechanical behavior and potentially improving knee OA diagnosis [8][9]. The mechanical behavior of the knee produces AE signals, influenced by factors such as bone movement and cartilage integrity. The deterioration of articular cartilage and the formation of osteophytes contribute to an increased number of AE events [10]. Previous research has demonstrated the feasibility of analyzing knee conditions through AE signals. Moreover, studies have confirmed AE's sensitivity to early-stage knee OA [10]. Daniela et al. investigated AE as a biomarker for OA and identified four preliminary AE biomarkers with diagnostic potential. However, they emphasized the necessity for improved AE measurement techniques and further validation of biomarkers [11]. In a pilot study assessing AE's diagnostic accuracy for knee OA, Kiselev et al. reported good to excellent accuracy across multiple knee regions, with specificity values ranging from 0.59 to 0.78 and sensitivity between 0.86 and 1. Nonetheless, the study was constrained by a small sample size [12]. Similarly, Dagyeong et al. demonstrated AE's potential in detecting OA knees by analyzing AE wave amplitude and frequency [13]. Heyon et al. explored the correlation between the prospective b-value of AE hits and knee conditions, obtaining promising results [14]. Additionally, Khan et al. and Hassan et al. examined the source localization of AE signals in OA-affected knees, yielding results that were later corroborated by physicians [15][16]. Despite significant research on AE's role in OA diagnosis, few studies have focused on the uncertainty, noise, or partitioning criteria of AE signals from healthy and OA knees. Previously, large-scale AE data analysis was hindered by limited computational resources and signal processing methods. Even with technological advancements, research on noise characterization in AE signals remains

scarce. Various machine learning techniques have been utilized to differentiate OA knees from healthy ones. For instance, Khan et al. assessed the clustering effectiveness of the Gaussian Mixture Model (GMM) algorithm on AE data from both OA and healthy knees [17]. However, previous studies have explored unsupervised learning and statistical analysis of AE signals from OA knees, none have leveraged convolutional neural networks (CNNs) combined with image classification techniques. With the continuous advancement of computational power and emerging supervised learning methodologies, there is a growing need to investigate their potential in knee OA diagnosis.

Therefore, this paper aims to implement transfer learning in CNN on AE signals obtained from different graded OA knees to classify the signals. AE data were obtained from 5 different graded OA patients. Preprocessing and wavelet transformation were performed in order to generate the images. Later these images were fed as input to the VGG-16 CNN network. The result from the classification has been evaluated on the basis of accuracy, F-score, sensitivity and specificity. The rest of the paper is designed as follows. Section gives introduction to KL-grading in knee OA and the experimental set up. Section 3 discusses the implemented CNN architecture; section 4 discusses the results, and the conclusion is in section 5.

2. KL GRADING BACKGROUND AND EXPERIMENTAL SET UP

2.1. KL Grading

The Kellgren and Lawrence (KL) grading system, accepted by WHO in 1961, is the most commonly used knee OA severity grading system[18]. KL system splits knee OA severity into 5 grades from grade 0 to grade 4. Fig. 4 shows X-ray images of different KL graded knees. Physicians usually inspect a scanned knee X-ray image and then give KL grades to both knee joints in a very short time period. The diagnostic accuracy is highly relied on physicians' experience and carefulness. In addition, the criterion of KL

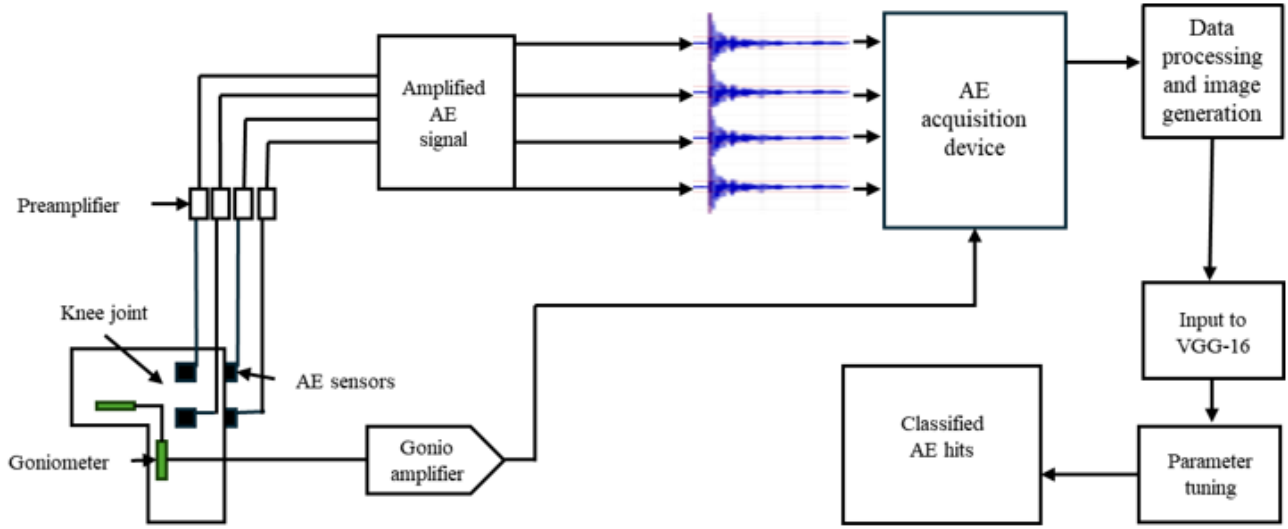


Figure 2: Schematics of the experimental set up. Four AE sensors are attached to the knee which are connected to the AE acquisition system via amplifiers.

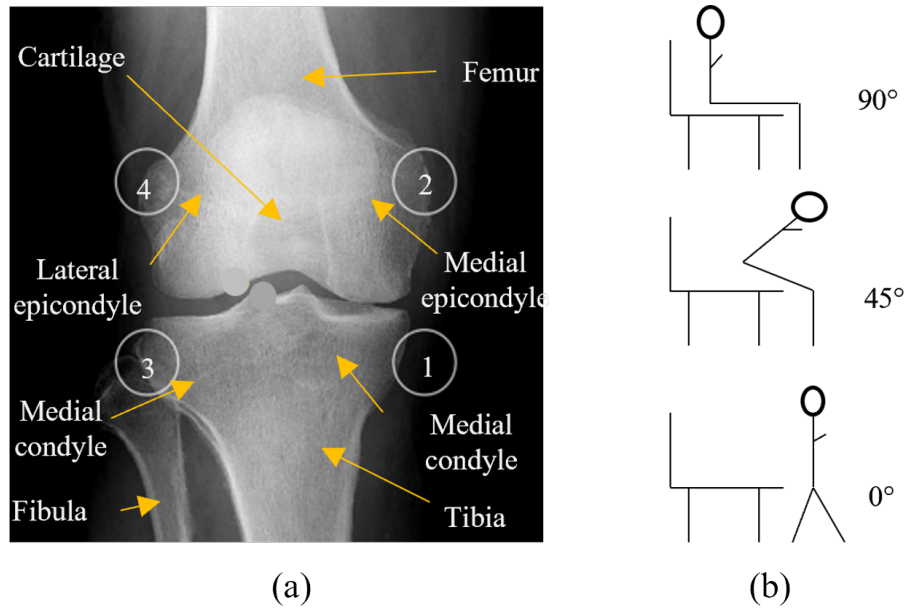


Figure 3: a) Sensor attachment (b) Sit-stand-sit schematics.

grading is very ambiguous. For example, possible osteophytic lipping and doubtful JSN are used as the criterion for KL grade 1. Even the same physician may give different KL grades for the same knee joint when inspecting at different time points. The KL intra-rater reliability ranges from 0.67 to 0.73 in a study conducted by Culvnr et al. We suppose this low reliability of physicians' grading to be

rooted in misclassifying the joint's KL grade to its nearby grades because of the ambiguous criterion. In clinical diagnosis, misclassifying the grade of a knee joint to its nearby grade (e.g., grade 1 to grade 2) is far less serious than misclassifying the grade to be far away (e.g., grade 1 to grade 4).

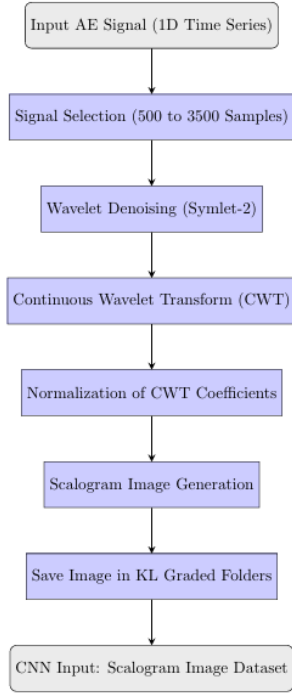


Figure 4: Data processing and image generation steps

2.2. Experimental setup and data preprocessing

AE data were captured from participants' knees in five groups, separated by diagnosed knee condition. The physicians of Tsuruta Orthopedic Hospital diagnosed the knees of the participant and graded according to the knee damage status. The grading system followed the KL grading system. However, for the simplicity of the network KL0, KL1 and KL2 groups were counted as one group. In a result, the classification network is designed to classify AE hits from 3 different groups. Four AE sensors (R6a, Physical Acoustic Corporation) were employed, with an operating frequency range of 35-100 kHz and a resonant frequency of 55 kHz. The acquisition system was managed by AEWin software, with the sensors connected via pre-amplifiers providing a 40 dB gain. Knee movement angles were recorded using two-channel goniometers (SG150, Biometrics Limited). Four AE sensors were placed close to the knee joint; their positioning was deliberated to reduce noise originating from surrounding tissues, muscles, and tendons. Two sensors were mounted to the femur and two to the tibia. The lateral condyle of the tibia, the medial epicondyle of the femur, the medial condyle of the tibia, and the lateral epicondyle of the femur are the corresponding locations of sensors 1 through 4, respectively. A highly elastic medical tape (ELASTPORE-HADA, NICHIBAN CO., LTD.) was used to attach the sensors to the knee. Vacuum-type coupling gel (Shin-Etsu HIVAC-G, Shin-Etsu Chemical Co., Ltd.) was used between the knee and sensor surfaces. The purpose of the elastic tape and gel was to maintain constant contact between the surfaces. At the start of the

experiment, the goniometer was initially set to 90 degrees for the subject's sitting position and 0 degrees for the subject standing position, which was attached to the knee using two-sided tapes. As a result, a sit-stand-sit movement cycle encompassed 180 degrees. That movement's three cycles were regarded as a single set. Five sets of information were obtained from each participant. Furthermore, an amplitude threshold of 45 dB was applied during the data acquisition procedure. The threshold was raised to 50 dB during the preprocessing stage to lessen the skewness of the data. This method's implications allowed for the removal of the majority of the noisy data. A pre-trigger value of 256 was used to set the sampling rate to 5 MSPS throughout the data-capturing process. Peak Definition Time (PDT), Hit Definition Time (HDT), and Hit Lockout Time (HLT) were set to 200, 800, and 1000 microseconds, respectively, to control the obtained envelope of the hit signals, with a maximum data collection time of 1000 milliseconds.

2.3. Data preprocessing and image generation

In this study, continuous wavelet transform (CWT) was employed to generate time-frequency representations of acoustic emission (AE) signals recorded from human knees. These scalograms serve as the input to a convolutional neural network (CNN) designed to classify knee osteoarthritis (OA) severity. The preprocessing steps include signal selection, wavelet denoising, transformation, and image generation. Signal Preprocessing

The AE signals were sampled at a high frequency of $f_s=1$ MHz to capture the intricate mechanical emissions from the knee joints. To ensure consistency across different signal recordings, each signal was truncated to a specific range, denoted as $s(n)$ where discrete time samples represent. The selected segment ranged from 500 to 3500 data points:

$$s(n) = S_{500:3500}(n) \quad (1)$$

where $s(n)$ is the original signal matrix. Wavelet-Based Denoising

To suppress noise while preserving transient features, wavelet denoising was applied using the Symlet-2 (sym2) wavelet basis. The maximum decomposition level was determined as:

$$L_{\max} = \lfloor \log_2(N) \rfloor \quad (2)$$

Where N is the signal length. The denoising process employed a Bayesian thresholding strategy with a hard thresholding rule, ensuring that noise components were minimized while retaining the dominant signal characteristics:

$$s_d(n) = W_{\text{denoise}}(s(n), L_{\max}, \text{Bayes}, \text{Hard}) \quad (3)$$

where W_{denoise} represents the wavelet denoising function.

Continuous Wavelet Transform and Normalization

The preprocessed signal was subjected to CWT using a filter bank with 12 voices per octave. The transformation is defined as:

$$C(a, b) = \int s_d(n) \psi_{a,b}^*(n) dn \quad (4)$$

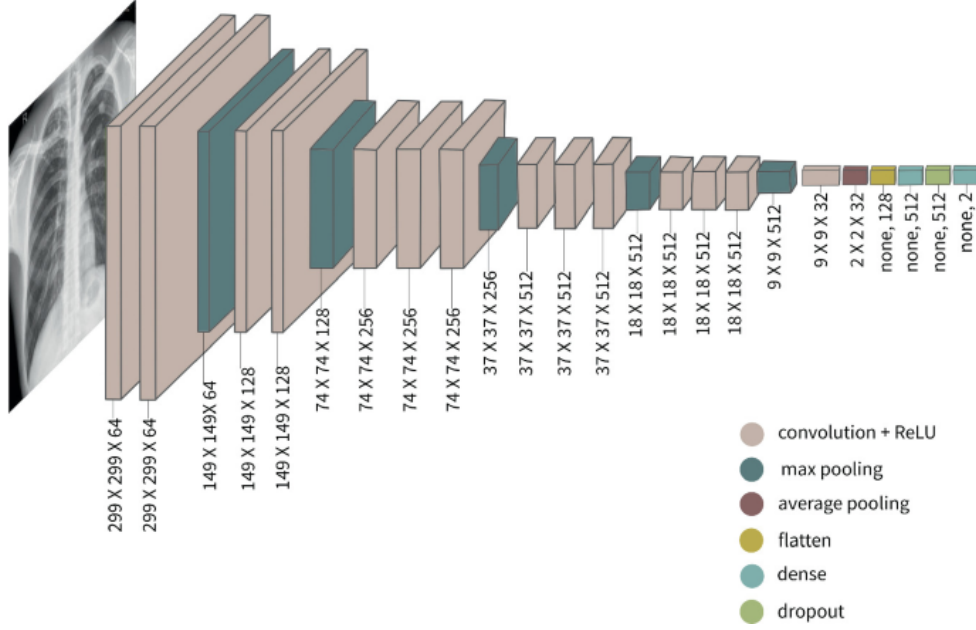


Figure 5: Block diagram of VGG-16 network architecture

Where $C(a,b)$ the wavelet coefficients at scale a and shift b and $\psi_{a,b}^*(n)$ is the complex conjugate of the wavelet function. The wavelet coefficients were subsequently normalized using:

$$C_{\text{norm}}(a, b) = \frac{C(a, b) - \mu}{\sigma} \quad (5)$$

Where μ and σ are the mean and standard deviation of $C(a,b)$, respectively.

Scalogram Image Generation

The normalized wavelet coefficients were visualized using a scalogram, where the x-axis represents time samples, and the y-axis represents the selected scale range (35 to 90). The scalogram was plotted using a colormap, with axes and labels removed to ensure optimal feature extraction by CNN. The images generated were saved as PNG files. These images were stored in predefined directories based on the Kellgren and Lawrence (KL) grading system. The process was repeated for all signals in the dataset, resulting in a robust dataset of scalograms suitable for deep learning-based classification of knee OA severity.

3. VGG-16

VGG-16 is a deep convolutional neural network (CNN) architecture that has significantly influenced image classification and feature extraction in computer vision. Proposed by Simonyan and Zisserman in 2014, VGG-16 gained prominence for its structured and deep architecture, achieving a top 5 accuracy of 92.7% on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC). The model consists of 16 weight layers, including 13 convolutional layers and 3 fully connected layers, followed by a SoftMax classification layer. It employs small 3x3 convolutional filters

True Class	KL012	475	80	88	73.9%	26.1%
	KL3	84	494	22	82.3%	17.7%
	KL4	47	12	541	90.2%	9.8%
		KL012	KL3	KL4	Predicted Class	

Figure 6: Confusion matrix from validation set after the training

with a stride of 1, preserving spatial resolution while capturing fine-grained details. The depth of the network enables hierarchical feature extraction, progressively capturing low-to-high-level representations. Max pooling layers with a 2x2 filter and a stride of 2 are applied after every few convolutional layers, reducing spatial dimensions while maintaining critical information. The fully connected layers, each with 4,096 neurons, contribute to the large number of 138 million parameters, making VGG-16 computationally intensive. Mathematically, the number of parameters in a convolutional layer is given by:

$$P = (K^2 \times C_{\text{in}} \times C_{\text{out}}) + C_{\text{out}} \quad (6)$$

Table 1: Summary of the performance metrics for each class and macro-averaged scores

Class	Precision	Sensitivity	F1 Score	Macro Precision	Macro Sensitivity	Macro F1-score	Overall Accuracy
KL012	0.7838	0.7387	0.7606				
KL3	0.8430	0.8233	0.8331	0.8193	0.8212	0.8195	0.8193
KL4	0.8310	0.9017	0.8649				

where K represents the kernel size (3×3), and N_{in} and N_{out} are the number of input and output channels, respectively. The fully connected layers follow:

$$P = N_{in} \times N_{out} + N_{out} \quad (7)$$

where N_{in} and N_{out} are the number of neurons in adjacent layers. These equations illustrate the computational burden imposed by the fully connected layers. VGG-16's structured design, which maintains uniform kernel sizes throughout, simplifies implementation and interpretation. However, the model's high memory consumption and slow inference speed pose challenges, particularly for real-time applications. Despite these limitations, VGG-16 remains widely utilized in various domains, including medical imaging, autonomous vehicles, biometric recognition, and acoustic emission-based diagnostics, where it serves as a robust feature extractor. Its effectiveness in transfer learning has also contributed to its extensive adoption in deep learning research. While more efficient models such as ResNet and EfficientNet have emerged, VGG-16 continues to be a foundational architecture in CNN-based applications.

4. RESULT AND DISCUSSION

The classification performance of the proposed deep learning model was evaluated using standard metrics, including overall accuracy, precision, sensitivity (recall), and F1-score. The model achieved an overall accuracy of 81.93%, demonstrating its efficacy in distinguishing between different classes of acoustic emission signals associated with knee osteoarthritis (OA).

4.1. Performance metrics analysis

The precision values for the individual classes were 0.7838 for KL012, 0.8430 for KL3, and 0.8310 for KL4, indicating that the model effectively minimizes false positive classifications across all severity levels. The macro-precision of 81.93% further supports the reliability of the classification outcomes. Sensitivity, also known as recall, was observed to be 0.7387 for KL012, 0.8233 for KL3, and 0.9017 for KL4. The relatively high recall values suggest that the model is capable of correctly identifying most instances of each class, with the highest recall achieved in the KL4 class (90.17%), signifying strong recognition ability for that category. The F1-score, which provides a balanced measure of precision and recall, was calculated as 0.7606 for KL012, 0.8331 for KL3, and 0.8649 for KL4. The macro-averaged values for precision, sensitivity, and F1-score were

81.93%, 82.12%, and 81.95%, respectively, indicating consistent performance across the dataset. The slight variation in precision and recall suggests that while the model is effective at correctly identifying instances, there is still room for improvement in reducing misclassifications.

4.2. Confusion matrix interpretation

The confusion matrix (Figure 6) provides further insight into the classification performance by illustrating the distribution of correctly and incorrectly classified instances. The majority of the diagonal elements indicate a high number of correctly classified samples, reinforcing the strong performance of the model. However, some misclassifications were observed, particularly between KL012 and KL3, as well as between KL3 and KL4, highlighting the need for further refinement in feature extraction and augmentation techniques to improve class separability.

5. Conclusion

In this study, a deep learning-based approach utilizing scalogram images was developed to classify acoustic emission signals from knee joints with different severities of osteoarthritis. The model achieved an overall accuracy of 81.93%, demonstrating its effectiveness in distinguishing between KL012, KL3, and KL4 severity levels. The high sensitivity and F1-scores indicate that the approach successfully captures key signal characteristics, making it a promising tool for automated OA assessment. While the model shows strong performance, further improvements can be achieved through enhanced preprocessing techniques, dataset expansion, and optimization of model architecture. Future research could explore transfer learning and hybrid models to further refine classification accuracy. These findings highlight the potential of deep learning methods in biomedical signal processing, paving the way for more advanced, non-invasive diagnostic tools for osteoarthritis detection.

[1] B. K. Madeti, S. R. Chalamalasetti, and S. K. S. siva rao Bolla Pragada, "Biomechanics of knee joint — A review," *Front. Mech. Eng.*, vol. 10, no. 2, pp. 176–186, 2015, doi: 10.1007/s11465-014-0306-x.

[2] R. Nisell, "Mechanics of the knee: A study of joint and muscle load with clinical applications," *Acta Orthop.*, vol. 56, no. S216, pp. 1–42, 1985, doi: 10.3109/17453678509154159.

- [3] B. Chen, W. Huang, and J. Liao, "Osteoarthritis: The Most Common Joint Disease and Outcome of Sports Injury," *J. Clin. Med.*, vol. 12, no. 15, pp. 12–14, 2023, doi: 10.3390/jcm12155103.
- [4] L. S. Lee, P. K. Chan, W. C. Fung, V. W. K. Chan, C. H. Yan, and K. Y. Chiu, "Imaging of knee osteoarthritis: A review of current evidence and clinical guidelines," *Musculoskeletal Care*, vol. 19, no. 3, pp. 363–374, 2021, doi: 10.1002/msc.1536.
- [5] F. W. Roemer, F. Eckstein, D. Hayashi, and A. Guermazi, "The role of imaging in osteoarthritis," *Best Pract. Res. Clin. Rheumatol.*, vol. 28, no. 1, pp. 31–60, 2014, doi: 10.1016/j.berh.2014.02.002.
- [6] S. Gholizadeh, Z. Lemana, and B. T. H. T. Baharudinb, "A review of the application of acoustic emission technique in engineering," *Struct. Eng. Mech.*, vol. 54, no. 6, pp. 1075–1095, 2015, doi: 10.12989/sem.2015.54.6.1075.
- [7] C. B. Scruby, "An introduction to acoustic emission," *J. Phys. E Sci. Instrum.*, vol. 20, no. 8, pp. 3–10, 1987, doi: 10.1088/0022-3735/20/8/001.
- [8] S. Saha, P. Kumar Mallisetty, S. Sen, and S. K. Giri, "Design and development of a high-speed data acquisition system for acquiring acoustic emission signals," *Mater. Today Proc.*, vol. 66, pp. 3830–3837, 2022, doi: 10.1016/j.matpr.2022.06.252.
- [9] L. K. Shark, H. Chen, and J. Goodacre, "Knee acoustic emission: A potential biomarker for quantitative assessment of joint ageing and degeneration," *Med. Eng. Phys.*, vol. 33, no. 5, pp. 534–545, 2011, doi: 10.1016/j.medengphy.2010.12.009.
- [10] L.-K. Shark, "Discovering Differences in Acoustic Emission Between Healthy and Osteoarthritic Knees Using a Four-Phase Model of Sit-Stand-Sit Movements," *Open Med. Inform. J.*, vol. 4, no. 1, pp. 116–125, 2010, doi: 10.2174/1876396001004010116.
- [11] D. K. Schlüter et al., "Use of acoustic emission to identify novel candidate biomarkers for knee osteoarthritis (OA)," *PLoS One*, vol. 14, no. 10, pp. 1–15, 2019, doi: 10.1371/journal.pone.0223711.
- [12] J. Kiselev, B. Ziegler, H. J. Schwalbe, R. P. Franke, and U. Wolf, "Detection of osteoarthritis using acoustic emission analysis," *Med. Eng. Phys.*, vol. 65, pp. 57–60, 2019, doi: 10.1016/j.medengphy.2019.01.002.
- [13] D. Choi, S. Ahn, J. Ryu, M. Nagao, and Y. Kim, "Knee acoustic emission characteristics of the healthy and the patients with osteoarthritis using piezoelectric sensor," *Sensors Mater.*, vol. 30, no. 8, pp. 1629–1641, 2018, doi: 10.18494/SAM.2018.1877.
- [14] H. K. Jeong, D. Whittingslow, and O. T. Inan, "B-Value: A Potential Biomarker for Assessing Knee-Joint Health Using Acoustical Emission Sensing," *IEEE Sensors Lett.*, vol. 2, no. 4, p. 1, 2018, doi: 10.1109/LSSENS.2018.2871981.
- [15] M. Hassan, T. Islam, K. Yuji, and H. Shuya, "Three-Dimensional Source Location of Acoustic Emission Technique for Damage Detection in Osteoarthritic Knee," *J. Nondestruct. Eval.*, vol. 123, 2022, doi: 10.1007/s10921-021-00834-1.
- [16] M. T. I. Khan, M. M. Hassan, Y. Hasemura, N. Matsuo, and S. Ide, "Acoustic emission technique in locating knee osteoarthritis," *Proc. 26th Int. Congr. Sound Vib. ICSV 2019*, pp. 1–4, 2018.
- [17] T. I. Khan, N. Sakib, M. M. Hassan, and S. Ide, "Gaussian mixture model in clustering acoustic emission signals for characterizing osteoarthritic knees," *Biomed. Signal Process. Control*, vol. 87, no. PA, p. 105510, 2024, doi: 10.1016/j.bspc.2023.105510.
- [18] J. H. Kellegren and J. S. Lawrence, "Radiological assessment of rheumatoid arthritis," *Ann Rheum Dis*. 1956 Mar;15(1):55–60. doi: 10.1136/ard.15.1.55.