

Article

Not peer-reviewed version

Bag of Feature-Based Ensemble Subspace KNN Classifier in Muscle Ultrasound Diagnosis of Diabetic Peripheral Neuropathy

[Kadhim Kamal Al-Barazanchi](#)*, [Ali Hussein Al-Timemy](#), [Zahid Mohamed Kadhim](#)

Posted Date: 11 September 2024

doi: 10.20944/preprints202409.0925.v1

Keywords: bag of features; ensemble subspace KNN; muscle ultrasound; diabetic peripheral neuropathy; speeded up robust features



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Bag of Feature-Based Ensemble Subspace KNN Classifier in Muscle Ultrasound Diagnosis of Diabetic Peripheral Neuropathy

Kadhim Al-Barazanchi ^{1,*}, Ali Al-Timemy ² and Zahid Kadhim ³

¹ Biomedical Engineering Department, College of Engineering, Al-Nahrain University, Baghdad, 10070, Iraq

² Biomedical Engineering Department, Al-Khwarizmi College of Engineering, University of Baghdad, Baghdad, 10070, Iraq

³ College of Medicine, University of Babylon, Babylon 51001, Iraq

* Correspondence: st.kadhim.k.hasan@ced.nahrainuniv.edu.iq

Abstract: Muscle ultrasound quantification is a valuable complementary diagnostic tool for diabetic peripheral neuropathy (DPN). DPN significantly impacts the lives of individuals with diabetes, leading to pain, lower limb amputation, and disability, affecting patient quality of life. This work develops a computer-aided diagnostic (CAD) system based on muscle ultrasound that integrates the bag of features (BOF) and an ensemble subspace k-nearest neighbour (KNN) algorithm for DPN detection. The BOF creates a histogram of visual word occurrences to represent the muscle ultrasound images and trains an ensemble classifier through cross-validation, determining optimal parameters to improve classification accuracy for the ensemble diagnosis system. The dataset includes ultrasound images of six muscles from 53 subjects, consisting of 27 control and 26 patient cases. An empirical analysis was conducted for each binary classifier based on muscle type to select the best vocabulary tree properties or K values for BOF. The result indicates that ensemble subspace KNN classification, based on the bag of features, achieved an accuracy of 97.23%. This study suggests muscle ultrasound as a promising diagnostic tool with the potential for image recognition and interpretation. CAD systems can accurately diagnose muscle pathology, helping to overcome limitations and identify issues in individuals with diabetes.

Keywords: bag of features; ensemble subspace KNN; muscle ultrasound; diabetic peripheral neuropathy; speeded up robust features

1. Introduction

Diabetic peripheral neuropathy (DPN) is a neuromuscular disorder that impacts individuals with diabetes mellitus, affecting approximately 28% of adults with diabetes in the United States [1]. Major risk factors for DPN include neuropathic pain, reduced sensation, ulcers, amputations, diminished quality of life, limitations in daily activities, and depressive symptoms [2,3]. The diagnosis of DPN is confirmed through quantitative testing and clinical signs, with nerve conduction studies being one of the most dependable methods. These studies assess the ability of peripheral nerves to transmit electrical signals in diabetic neuropathy patients and help identify the occurrence and progression of DPN. The expertise of trained medical professionals is necessary to carry out these studies [4]. Muscle ultrasound, a quantitative tool for diagnosing muscle conditions, has become essential due to its reliability, sensitivity, affordability, and ease of use with high-frequency ultrasound transducers. Computer vision and artificial intelligence (AI) advancements have greatly influenced CAD systems, particularly in medical imaging research. AI methods are used to interpret imaging data for potential diagnosis improvements, especially muscle ultrasonography, for better clinical performance. [5,6].

The BOF methodology, also referred to as Bag of Words (BOW), is a valuable tool in various domains, such as texture recognition and computer vision applications. It revolves around the utilization of an organized set of image features. In image analysis, the BoF model employs a visual

depiction of a word created through vector quantization achieved by clustering low-level visual attributes like color and texture [7,8]. Classifying medical images can be challenging due to the large number of extracted features, making it difficult for machine learning algorithms. Using BOF and clustering can help reduce the number of features for more efficient classification. However, selecting the correct number of clusters is crucial [9]. The BOF work is based on a sped-up robust features (SURF) detector-descriptor approach. It uses integral images to decrease computation time (Fast-Hessian detector) and characterizes Haar-wavelet responses around the point of interest [10]. Ensemble techniques combine multiple classifiers and are popular for enhancing classification performance. They show significant error reductions across real-life applications and handle non-informative features better than individual models [11]. In ensemble learning, several approaches exist: (i) Random subspace: Involves randomizing the learning algorithm by selecting a subset of features before training and combining the models' outputs using a majority vote. (ii) Bagging or bootstrap aggregation creates models trained on random data, and the predictions are aggregated for the final prediction using averaging. (iii) Boosting Involves averaging or voting on multiple models and weights the constructed models based on their performance [12].

Researchers have used muscle ultrasound as a supplementary diagnostic tool for DPN, integrating it with clinical data[13,14]. They have also employed ensemble classifiers for medical data classification [15], and utilized ensemble-based BOF algorithms for medical image classification to enhance the performance and accuracy of the diagnosis system [16,17]. Recently, different feature extraction approaches have been conducted; Konig et al. [18] extracted first-order statistics, wavelet-based, and Haralick's features from ultrasound images and used feature selection and reduction to identify a stable combination. Then, they explored Fisher's, support vector machine (SVM) and KNN classifiers for myositis detection and achieved a classification accuracy of 87%. Nodera et al. [19] quantitatively analyzing lower leg ultrasound images using texture features extracted by calculating the histogram, grey-level co-occurrence matrix, neighborhood grey-level different matrix, grey-level run length matrix, and grey-level zone length matrix. The classification was performed by simple logistic, SVM, and Random Forest, with the Random Forest algorithm achieving a classification accuracy of 78.4%. On the other hand, many researchers utilized deep learning capabilities to classify muscle ultrasound images, such as Ahmed et al. [20] who proposed a modified lightweight YOLOv5 architecture, convolutional block attention module, spatial pyramid pooling-fast plus, and exponential linear unit activation function to detect and classify inflammatory myopathies automatically. This model achieves an accuracy of 98% for binary and multiclass classification.

Recent research has emphasized the importance of employing advanced techniques like feature extraction, machine learning, and deep learning models to analyze muscle ultrasound. Further exploration in this area has the potential to enhance patient care and provide more accurate diagnoses for DPN. As a result, the current study introduces a new integration of a bag of features and ensemble subspace KNN algorithms, which have demonstrated promising outcomes in enhancing the diagnostic accuracy of DPN. This represents a significant advancement in muscle ultrasound research. Moreover, the study includes a comparative analysis of the performances of various classification algorithms and their impact on model accuracy and diagnostic tasks. In addition, it provides insights into utilizing comprehensive performance analysis, reliable evaluation, and innovative approaches for diagnosing DPN using muscle ultrasound.

2. Materials and Methods

The methodology employed in this study is focused on developing and evaluating an integration of BOF with the ensemble subspace KNN classification models for detecting DPN using muscle ultrasound, improving the accuracy and dependability of the DPN diagnostic system, as illustrated in Figure 1. This section details our step-by-step process to create and assess the proposed system for DPN classification system. Our research used the MATLAB platform's computer vision toolbox and machine learning capabilities. The MATLAB framework helped us design and implement machine learning algorithms and supported the validation and testing processes using new datasets. This streamlined environment was essential for conducting efficient and rapid trials,

which, in turn, facilitated the analysis and validation of the proposed method for DPN diagnosis. The muscle ultrasound images were divided into training and testing sets using five-fold cross-validation. This cross-validation splits the dataset into five equal parts to assess the classifier's performance by averaging the results from five separate runs.

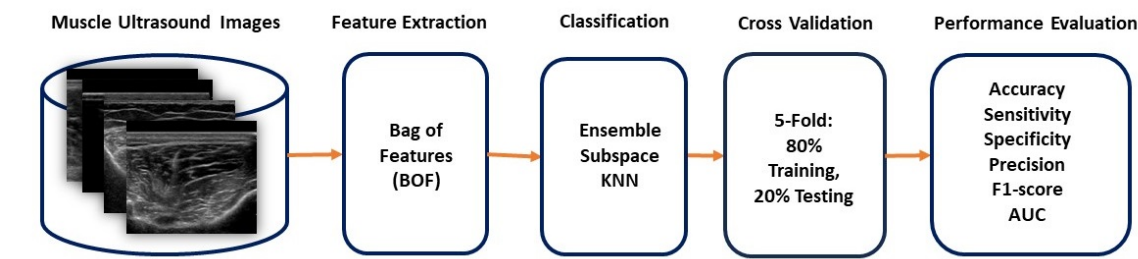


Figure 1. Proposed BOF-Based Ensemble subspace KNN classification System for DPN diagnosis.

2.1. Dataset

A case-control study comprising 26 DPN patients having type 2 diabetic mellites and 27 healthy controls (CTR) was conducted. Table 1 presents the demographics of the data set, including 53 subjects. Our work was carried out in academic collaboration with the Ghazi Al Hariri Surgical Hospital's neurophysiology clinical center at the Medical City Complex in Baghdad, Iraq. The study complied with the Declaration of Helsinki's ethical guidelines for human experimentation and was authorized by the local health ethics council. An electrodiagnostic study confirmed the presence of DPN, and an HbA1c test assessing average blood glucose levels was part of the exclusion criteria.

Table 1. Demographic of the study group Dataset.

Variable	DPN (26)	Control (27)
Age (years)	51.5 ±9.3	42.5 ±9.6
Gender Male/Female Ratio	20/6	20/7

This study involved ultrasound images of six muscles, as the disease under study is length-dependent, the investigation is required for each leg, from the proximal to distal muscles. Three of the muscles were in the upper limb: biceps brachii (BB), a total of 407 images; brachioradialis (BR), a total of 466 images; and abductor digiti minimi (ADM), a total of 605 images. The other three muscles were in the lower limb: rectus femoris (RF), a total of 514 images; tibialis anterior (TA), a total of 506 images; and abductor hallucis brevis (AHB), a total of 723 images. The ultrasonographic examination was performed using a Philips iU22, 2012 ultrasound machine equipped with a linear probe set to a 5-12 MHz frequency. A particular preset for musculoskeletal examination was used to examine all the subjects. The subject was lying prone and relaxed to film the muscles while at rest and in a static position. A skilled physician performed the muscle ultrasound filming, considering the anatomical position of the muscle and referring to a previously published paper that serves as a guide for researchers conducting ultrasound scans of the upper and lower limb muscles [21–26]. The presets installed on the ultrasound machine were used for filming, with the musculoskeletal general preset having a gain of 50%, compression of 62, medium pressure, and a depth of 3 cm. Ultrasound images were acquired from the belly, and the best image showing muscle fibers in a transverse view with the bone marked behind was saved. The images were saved in digital imaging and communications in medicine (DICOM) format on the ultrasound machine, labeled with the patient code number and muscle name for reference and further analysis.

2.2. Bag of Feature

Muscle ultrasound images are represented as compact descriptions through the BOF model, and images are concisely described by creating histograms of local descriptors. This approach involves quantizing the local descriptors of images into visual words based on a visual vocabulary. This

vocabulary is generated by clustering a large set of local descriptors using K-means algorithms. The output is a set of clusters, where each cluster represents a visual word. As a result, an image can be described as a bag of visual words, and a histogram can be constructed with a dimension equal to the size of the visual vocabulary. Each bin in the histogram contains the frequency of occurrence of a visual word within the image [27], as shown in Figure 2. The BOF method is effective for image classification because it resists object orientation changes and eliminates the need for segmentation before classification. It creates a codebook representing all possible visual features in a dataset [16]. A visual codebook maps low-level features into a fixed-length feature vector within a histogram space, allowing for direct application to the classifiers. A discriminative codebook is created by selecting representative key points, while a compact codebook is obtained by eliminating indistinct codewords, reducing computational complexity and improving categorization precision [28].

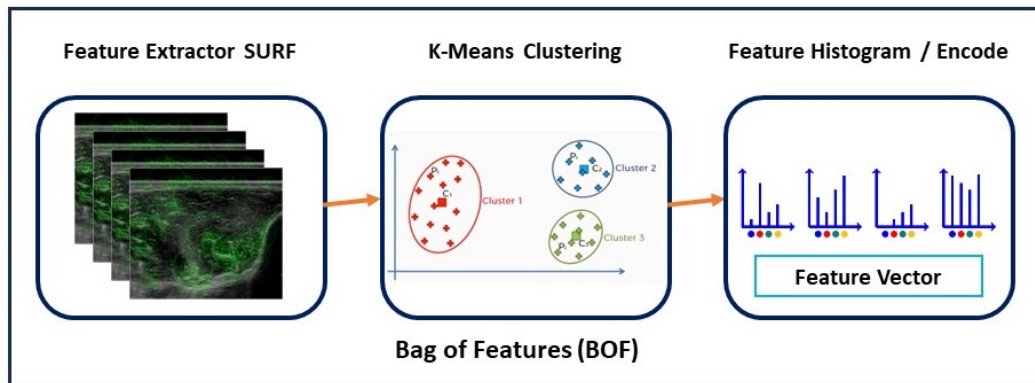


Figure 2. Illustration of BOF steps.

2.2.1. Feature Extraction

In classifying muscle ultrasound images, feature extraction is a crucial initial step, categorized into three types: low-level (color and texture), middle-level (shape), and high-level (semantic interpretation). Color is often used for easy extraction from muscle ultrasound images. Low-level local features are popular in action recognition for handling background noise and being independent in detection and tracking. These local features involve identifying a local area (detector) and describing the identified area (descriptor). We utilized the grid method and the SURF descriptor algorithm to locate key points and extract their feature vectors. A uniform grid with a grid step of 8×8 was overlaid onto the image to find the interest points. Classifying muscle ultrasound images requires feature extraction, which involves low-level, middle-level, and high-level features. We utilized the grid method and the SURF descriptor algorithm to locate key points and extract their feature vectors. The SURF descriptor is based on the sum of the Haar wavelet response around the point of interest by employing the integral image to reduce computational overhead. In interest point detection, SURF employs an integer approximation of the determinant of the Hessian blob detector, which can be computed with just three integer operations using a precomputed integral image. In orientation assignment, the SURF algorithm uses wavelet responses in horizontal and vertical directions, incorporating appropriate Gaussian weights. The dominant orientation is determined by summing all responses within a sliding orientation window of sixty degrees. Additionally, SURF offers a feature known as Upright-SURF or U-SURF, which enhances speed and demonstrates robustness within a range of ± 15 degrees. We used square patches or regions positioned around key points to extract features at multiple scales. These regions of various sizes (32×32 , 64×64 , 96×96 , and 128×128) were divided into sub-regions (8×8 square), and the vertical and horizontal Haar wavelet responses were calculated to create a four-dimensional sub-region feature vector. This vector captured the intensity changes' polarity at each scale.

2.2.2. Quantization and Clustering

Multiple features are generated during the key points feature extraction process, resulting in a 64-dimensional feature vector for each region at every scale. Before creating the codebook, these features are initially reduced to the top 80% of the most robust descriptors based on their scores. After identifying the strongest key points and extracting their features using the SURF descriptor, the next step is to extract the Bag of Features (BOF) from the muscle ultrasound images through vector quantization. The features are then quantized using the K-means clustering algorithm to create a vocabulary of K visual words. This involves grouping the descriptors into K clusters, with the cluster centers representing the K visual words. The initial cluster centers are randomly selected based on the inputted descriptors, and the descriptors are assigned to the nearest cluster center based on the Euclidean distance. The process then iteratively calculates new cluster centers to reduce the sum of the squared Euclidean distances until the cluster centers become stable. The K-means clustering algorithm typically generates K visual words based on the BOF codebook's number of K clusters, also known as codewords.

2.2.3. Encoding

The ultrasound images are represented using visual words obtained from clustering. Histograms are created from the distribution of visual words in each image, forming a feature vector showing how often the K visual words appear. The histograms are generated using a greyscale framework of the muscle ultrasound images. As a result, the input images' histograms represent the visual content distribution using the created codebook. A machine learning algorithm then analyses these histograms and classifies the input muscle ultrasound images as normal or DPN cases. The BOF model is defined by processing a muscle ultrasound D dataset containing N images. Each image is represented as a set of visual features, denoted as $d_1, d_2, \dots, d_n$. K-means unsupervised learning algorithm is then used to group the images based on a fixed number of visual words W denoted as $w_1, w_2, \dots, w_k$, where K is the number of clusters. The data can be summarized in an $N \times K$ co-occurrence table of counts $N_{ij} = n(d_i, w_j)$, where $n(d_i, w_j)$ represents the frequency of word w_j occurring in image d_i .

2.3. Ensemble Subspace KNN

Ensemble learning integrates multiple classification methods to build a powerful composite model. The objective is to achieve higher accuracy compared to using any single model. The base learning method in the study is KNN, while the ensemble approach is a random subspace. KNN algorithm divides the feature space into separate clusters based on the features associated with different classes. When classifying a test feature vector, this classifier considers the k-nearest metric distances between the test sample features and those of the closest classes. In the architecture of KNN, the number of neighbors and the type of distance metric used are significant factors. Ensemble classifiers combine the outputs of multiple base classifiers to improve overall classification accuracy as illustrated in Figure 3. The KNN algorithm is stable but sensitive to feature set changes. Ensemble systems address this by creating individual classifiers from randomly selected data subspaces. These classifiers' outputs are combined using a majority vote to generate the final result. Subspace KNN involves projecting all points onto a specified subspace and calculating distances to determine the k nearest neighbors. Multiple chosen subspaces are used to identify the majority vote on the class membership of the test sample by combining the k nearest neighbors [29,30]. In the random subspace ensemble model, a set of N base KNN classifiers are trained using a randomly selected subset of features (from D_1 to D_m) from the complete dataset (D). This approach allows more accessible training using smaller subspaces and improves the features-to-instance ratio. Specifically, for muscle ultrasound image classification, a p^* -dimensional feature subset is randomly chosen from the given p-dimensional dataset ($p^* < p$), and KNN classifiers are learned using the subspace feature vectors. This process is repeated M times to train M classifiers with new feature vectors each time, and the predictions of N classifiers are evaluated using a majority voting approach. The study used Subspace

KNN for training. It employed the simple majority vote rule and the subspace ensemble method. The learner type was the nearest neighbor, comprising 30 learners and 200 subspace dimensions for classifying the muscle ultrasound images.

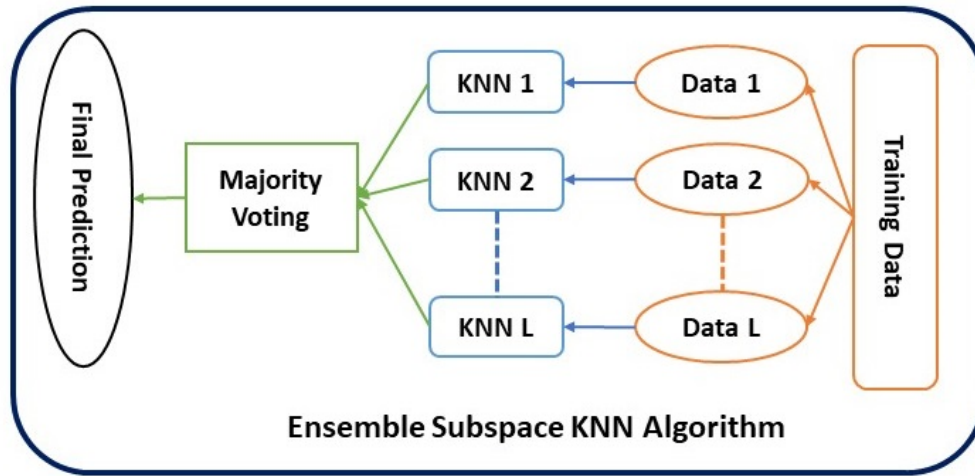


Figure 3. Ensemble Subspace KNN Algorithm Architecture.

2.4. Performance Evaluation

We employ various measures to assess the performance of our proposed method and determine its effectiveness. These metrics are designed to evaluate different aspects of the method's effectiveness and are formulated as follows:

The t-SNE is a technique used to visualize high-dimensional data by projecting it onto a 2D map while preserving its original structure. It's beneficial for visualizing data with thousands of dimensions, such as feature vectors from a deep transfer learning model [31,32].

A confusion matrix is a table arrangement that makes it possible to see how well an algorithm performs. It comprises false positives (FP), false negatives (FN), true positives (TP), and true negatives (TN), where P and N are positive and negative samples in the original dataset.

Accuracy is the ratio of correctly predicted class samples to total samples. It measures the model's classification accuracy using the following formula:

$$Accuracy = \frac{TP + TN}{P + N} \quad (1)$$

Sensitivity measures the accuracy of positive predictions by calculating the ratio of correctly identified positive results to the total number of true positive and false negative results. The formula is as follows:

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

Specificity measures the model's ability to correctly reject healthy patients without a condition by identifying negative instances. It is calculated by dividing the number of true negatives by the sum of true negatives and false positives. The formula is as follows:

$$Specificity = \frac{TN}{FP + TN} \quad (3)$$

The F1-Score measures predictive performance calculated from the test's sensitivity and specificity. The formula is as follows:

$$F1 - Score = \frac{2 TP}{2 TP + FP + FN} \quad (4)$$

Precision is the ratio of correctly predicted observations to the total predicted observations, expressed as follows:

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

A receiver operating characteristic curve (ROC) is a graphical tool used to assess a classifier's performance. The AUC represents the separability of class labels. A higher AUC value indicates better predictive capability, while a lower value suggests a less accurate model.

3. Results

In this section, we will review the performance analysis findings of our proposed integrated model, which utilizes ensemble subspace KNN classification based on a BOF model. The experimental settings involved variations of the BOF with different numbers of levels in the vocabulary tree K values, applied with machine learning algorithms using datasets comprising muscle ultrasound images. We extensively examined the extracted features from BOF, demonstrating the best performance based on classification accuracy using the optimal K value. We compared the performance of our BOF dataset with cross-validation using the ensemble subspace KNN algorithm, considered a state-of-the-art model for our research.

3.1. BOF Empirical Evaluation

The suggested BOF-based ensemble subspace KNN muscle ultrasound classification model is designed to diagnose DPN. The model undergoes training using labeled ultrasound images of six muscles to construct the BOF codebook and set parameters for the classification model. During the testing stage, we utilized ultrasound images as the basis for our key points detection algorithm. This enabled us to derive SURF descriptors from the recognized key points. We then quantized the most dependable features using a created codebook to produce frequency histograms for the testing images. Ultimately, the pre-existing classification model produced the intended result. The data set has been used to select optimal values using empirical analysis to find the best setting for optimizing the BOF. In our research, we examined the impact of varying the K values for quantization of the visual word (K visual word) on the system's performance. We conducted experiments using different K values, from 200 to 400, to explore how increasing the number of extracted features from each muscle ultrasound image affects the classification accuracy.

Additionally, we compared two classification models with the proposed ensemble subspace KNN-based classification in the BOF model. The first model involved fine KNN algorithms with a Euclidean distance and a single neighbor with equal distance weight. The second model used SVM with a Gaussian Kernel function and a Kernel scale parameter set to 32. Figure 4 illustrates the empirical analysis of ADM ultrasound images. The highest accuracy of 97.23% was achieved with 400 visual words among the three classifiers using our proposed ensemble subspace KNN algorithm. The base KNN and SVM algorithms exhibit variation as the K value increases alongside our proposed ensemble model.

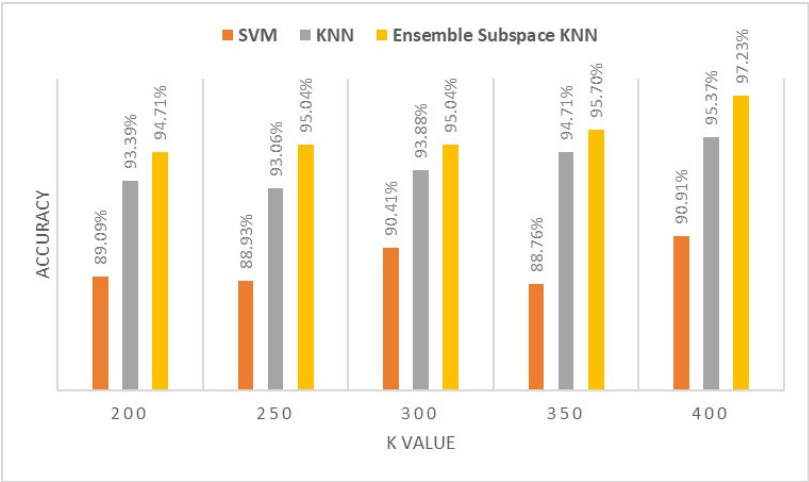


Figure 4. Classification accuracy of different K values for ADM muscle.

Figures 5–9 demonstrate the impact of increasing the number of visual words K from 200 to 400 on the classification models of ultrasound images of TA, RF, BR, AHB, and BB muscle, respectively. This empirical analysis is based on three classification algorithms to indicate the performance of our proposed ensemble model.

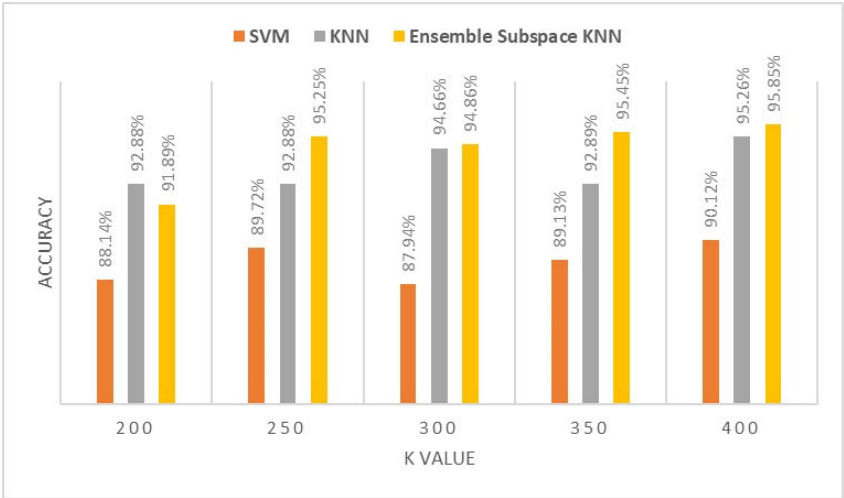


Figure 5. Classification accuracy of different K values for TA muscle.

The proposed ensemble subspace KNN algorithms demonstrate superior classification accuracy compared to KNN and SVM with a K value of 400 visual words BOF. They achieve performance accuracy of 95.85% for TA muscle, 95.7% for RF muscle, 95.17% for BR muscle, 92.67% for AHB, and 89% for BB muscle. In comparison, KNN achieves performance accuracy of 95.26%, 94.16%, 93.13%, 91.7%, and 85.61%, while SVM achieves performance accuracy of 90.12%, 92.41%, 89.27%, 92.12%, and 85.13% for ultrasound images of TA, RF, BR, AHB and BB muscle respectively.

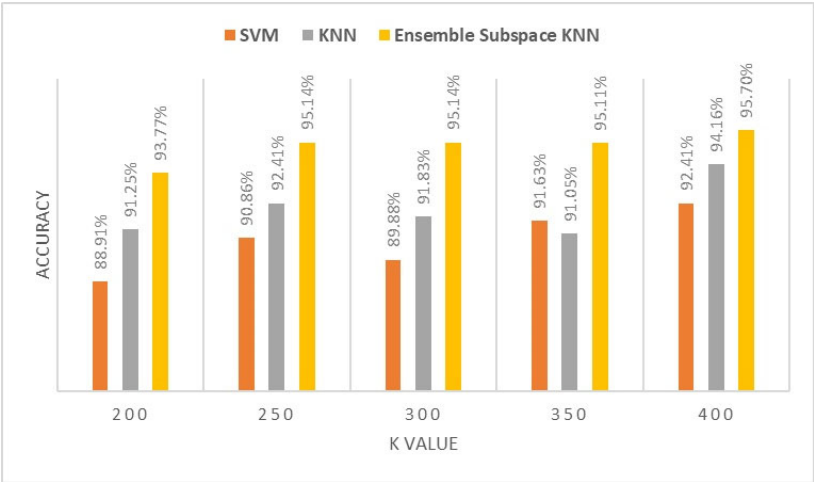


Figure 6. Classification accuracy of different K values for RF muscle.

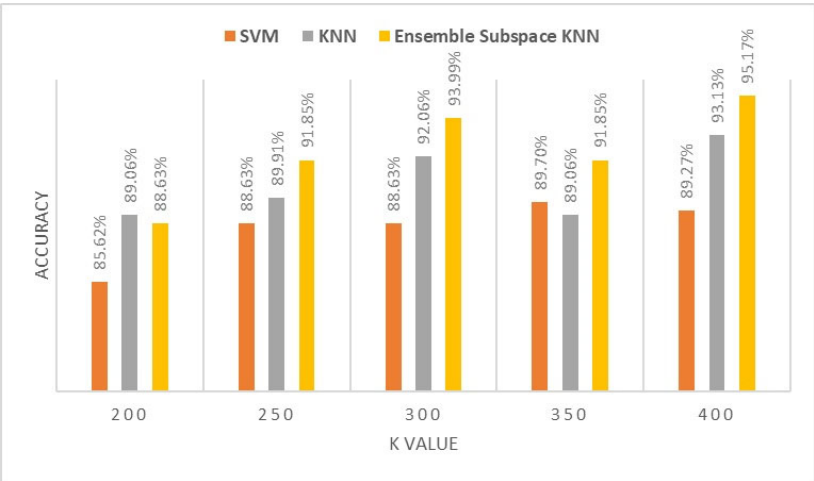


Figure 7. Classification accuracy of different K values for BR muscle.

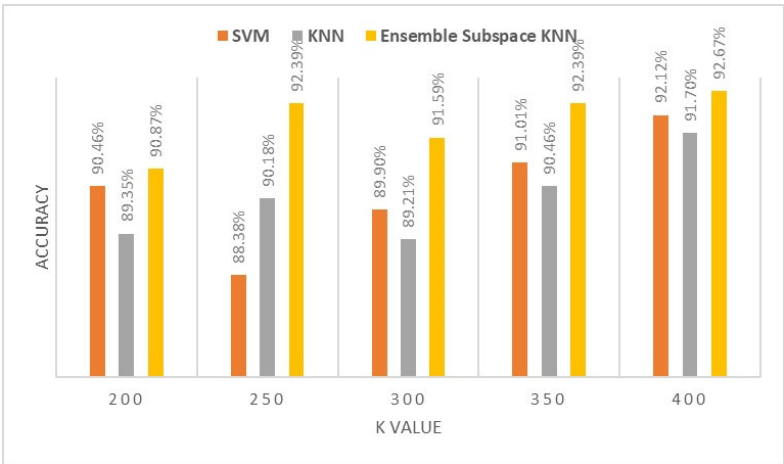


Figure 8. Classification accuracy of different K values for AHB muscle.

Increasing the number of visual words from 200 to 400 positively impacts the performance of ensemble, KNN, and SVM algorithms by improving classification accuracy metrics. The analysis

reveals that by increasing the number of visual words K, there is an average improvement in performance accuracy of nearly 5% across the six muscle ultrasounds for the classification algorithms.

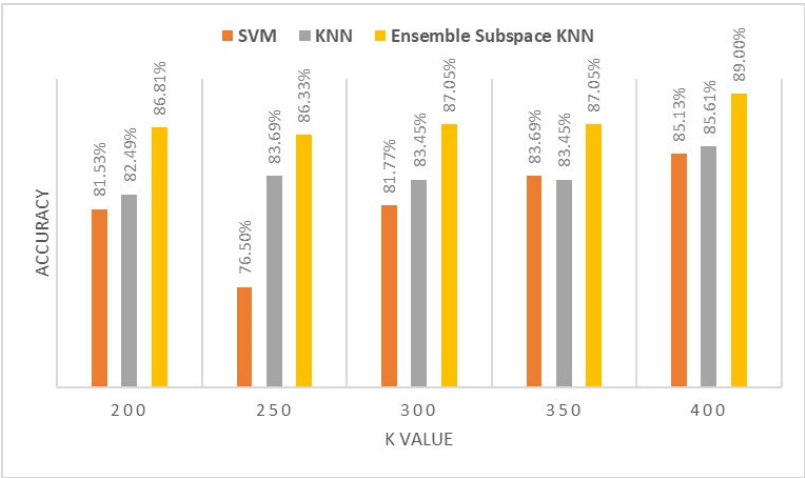


Figure 9. Classification accuracy of different K values for BB muscle.

3.2. BOF Visualization Evaluation

We conducted an analysis of visual words and utilized t-SNE visualization to explore the distribution of high-dimensional feature datasets obtained from ultrasound images using 400 visual words. This method enabled us to map the data onto a 2D representation, offering valuable insights into the data distribution and aiding in the interpretation of predictions. Our analysis revealed that the feature datasets showed non-linear separations, leading to overlapping data clusters in certain areas. Figures 10 and 11 depict the t-SNE distribution of upper limb muscles (BB, BR, and ADM) and lower limb muscles (RF, TA, and AHB) across six different muscles for the BOF dataset with 400 visual words. The data exhibited multiple clustered groups on the 2D map due to BOF features K-means clustering. This approach necessitated a non-linear classifier based on a non-parametric method, which makes predictions by comparing the similarity of data points and identifying the closest groups or nearest points for a query point. This discovery was influential in the selection of ensemble subspace KNN classification algorithms. Furthermore, qualitative observations indicated that t-SNE significantly improved the classification accuracy in the training and validation datasets.

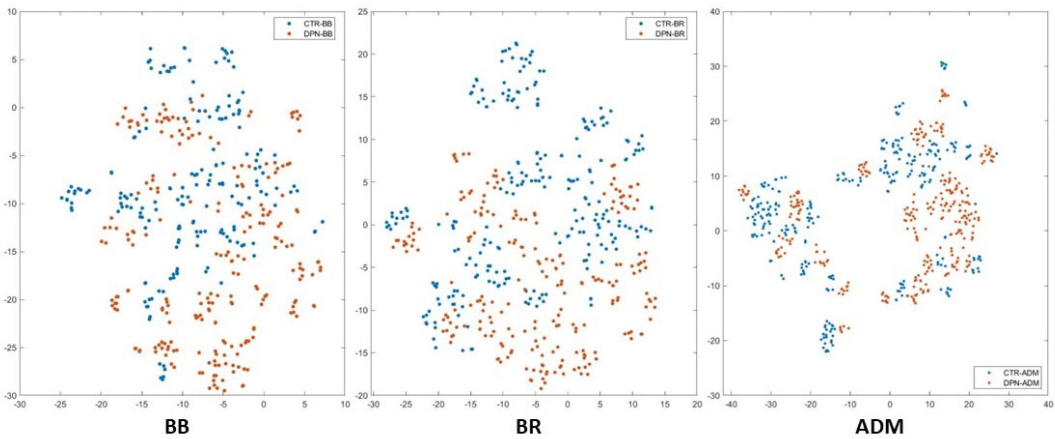


Figure 10. The t-SNE display for the upper limb muscles (BB, BR & ADM) BOF dataset with 400 visual words.

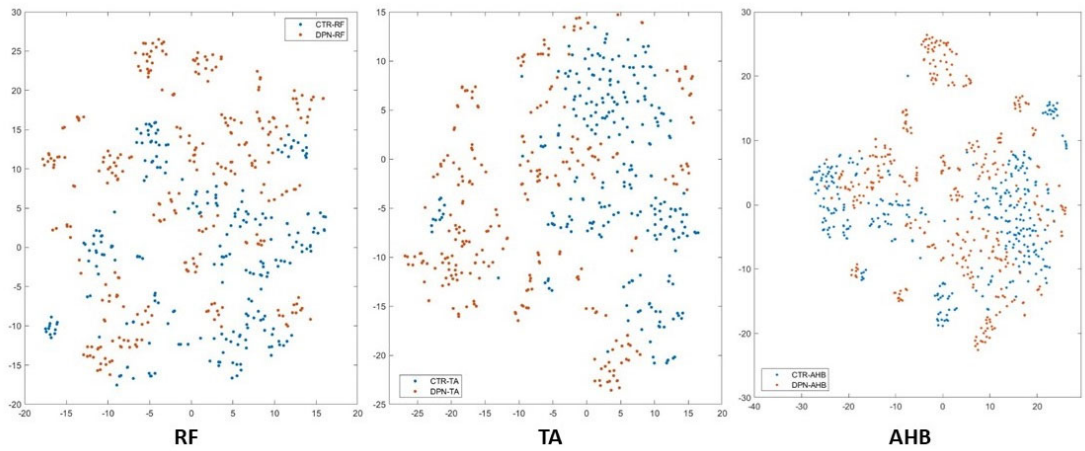


Figure 11. The t-SNE display for the lower limb muscles (RF, TA & AHB) BOF dataset with 400 visual words.

3.3. Classification Performance Ensemble Subspace KNN

The classification evaluation results are based on performance metrics listed in Table 2 for the feature datasets. We used 5-fold cross-validation to ensure the model's performance on an independent dataset and to effectively identify issues such as overfitting or selection bias to gain valuable insights. This evaluation used an ensemble subspace KNN with the extracted features, employing the BOF with 400 visual words. The results show that the ensemble-BOF-based classifier for the ADM muscle achieved the highest classification accuracy of 97.23%, while the sensitivity, specificity, precision, F1-score, and AUC were 97.62%, 96.88%, 96.63%, 97.12% and 99.52% respectively.

On the other hand, our proposed model achieved the lowest result with BB muscle, with an accuracy of 89%, sensitivity of 89.7%, specificity of 88.75%, precision of 85.41%, F1-score of 87.29%, and AUC of 99.62%. Additionally, the evaluation metrics of RF, TA, BR, and AHB achieved an accuracy of 95.7%, 95.85%, 95.17%, and 92.76%, while the AUC was 99.06%, 99.21%, 99.04%, and 92.83% respectively. The sensitivity, specificity, precision and F1-score finding for the TA muscle ultrasound images were 94.12%, 97.61%, 97.56% and 95.81% and for RF muscle ultrasound images were 94.80%, 97.09%, 96.96% and 95.86% respectively based on our proposed ensemble model. Moreover, the sensitivity, specificity, precision and F1-score finding for the BR muscle ultrasound images were 96.68%, 93.62%, 93.95% and 95.3% and for AHB muscle ultrasound images values were 89.66%, 95.62%, 95.25% and 92.37% respectively. The sensitivity, specificity, precision, F1-score, AUC, and accuracy can vary within a reasonable range across different muscle ultrasound images due to variations in anatomy, physiology, and the pathological effects of DPN on each muscle. Our proposed ensemble-BOF-based classifier can accurately classify cases as either DPN or healthy. Figure 12 confusion matrix for the binary classifier per each muscle, based on the proposed ensemble-BOF-based model. The numbers in the matrix indicate the count of images per class. Upon five-fold cross-validation, our model consistently predicts DPN images more accurately than healthy/control images. This is due to anatomical changes in the muscle fibers of DPN patients, resulting in increased muscle echogenicity or gray scale representation. These changes will be reflected in the detected features and improve the classification accuracy of DPN cases.

Table 2. Metric Analysis of Ensemble Subspace KNN with 400 visual word BOF Dataset for 5-fold Cross-validation.

Muscle	Sensitivity	Specificity	Precision	F1-Score	AUC	Accuracy
TA	94.12%	97.61%	97.56%	95.81%	99.21%	95.85%
RF	94.80%	97.09%	96.96%	95.86%	99.06%	95.70%

BB	89.27%	88.75%	85.41%	87.29%	99.62%	89.00%
BR	96.68%	93.62%	93.95%	95.30%	99.04%	95.17%
ADM	97.62%	96.88%	96.63%	97.12%	99.52%	97.23%
AHB	89.66%	95.62%	95.25%	92.37%	92.83%	92.67%



Figure 12. Confusion Matrix of Ensemble-BOF-Based classifier with 400 visual words.

4. Discussion

Our research addresses limitations in previous studies by significantly advancing the DPN diagnosis method. By integrating ensemble machine learning and using a bag of features or visual word approaches, our model offers a robust and precise technique for detecting DPN. The bag of features method analyzes ultrasound images and generates a representation using a group of visual words to illustrate muscle characteristics and extract relevant features from complex image data. Additionally, the ensemble random subspace KNN algorithm combines multiple base learners to enhance the classification accuracy of the base classifier, leading to improved predictions and overcoming the bias-variance tradeoff problem. A comparative analysis highlighting our proposed system's performance compared to other state-of-the-art studies demonstrates the superiority of the bag of features ensemble subspace KNN-based DPN detection model as in table 2 and figures 4 to 9. This preponderance of ensemble-BOF-based model against texture-based classification through conventional KNN and SVM models in [18,19]and complex deep-learning techniques in [20] based on performance for distinguishing between positive and negative cases and stability across iterations.

5. Conclusions

Muscle ultrasound is an important imaging technique for diagnosing various neuromuscular disorders, including diabetic peripheral neuropathy. Ultrasound imaging of muscles can identify changes in tissue structure caused by the degeneration of healthy muscle fibers and their replacement with fatty deposits and fibrous tissue. Muscles affected by DPN appear differently on ultrasound compared to healthy muscles. Patients with DPN typically have higher muscle echogenicity, meaning

their muscles appear whiter on ultrasound scans. However, its full potential is hindered by limitations in visual interpretation and the necessity for device-specific reference values. Developing specialized ultrasound systems and using artificial intelligence in computer-aided diagnosis improve usability and enhance this imaging technique's effectiveness. In our study, we researched the bag of features classification model. The model used labelled ultrasound images of six muscles to construct a codebook and set parameters for the classification model. The training process involved detecting key points, extracting descriptors, and constructing a codebook using K-means clustering. Then, the obtained histograms were used to train the classification model. We examined how the number of visual words impacts classification performance for six muscles in each subject's upper and lower limbs. Increasing the number of visual words enhances classification accuracy by incorporating more distinctive features. In the proposed BOF model, ensembles were used in the classification process to efficiently detect muscle ultrasound images in normal or DPN cases. Ensembles of random subspace KNN were investigated and compared with KNN and SVM base classifiers. The results have shown that the ensemble subspace KNN achieved superior performance with a maximum accuracy of 97.23% for ADM muscle ultrasound images when the number of visual words K was set to 400.

Conflicts of Interest: The authors declare no conflicts of interests.

References

1. Hicks, C.W.; Selvin, E. Epidemiology of Peripheral Neuropathy and Lower Extremity Disease in Diabetes. *Curr Diab Rep* 2019, 19.
2. Bansal, V.; Kalita, J.; Misra, U.K. Diabetic Neuropathy. *Postgrad Med J* 2006, 82, 95–100.
3. Galiero, R.; Caturano, A.; Vetrano, E.; Beccia, D.; Brin, C.; Alfano, M.; Di Salvo, J.; Epifani, R.; Piacevole, A.; Tagliaferri, G.; et al. Peripheral Neuropathy in Diabetes Mellitus: Pathogenetic Mechanisms and Diagnostic Options. *Int J Mol Sci* 2023, 24.
4. Yu, Y. Gold Standard for Diagnosis of DPN. *Front Endocrinol (Lausanne)* 2021, 12, doi:10.3389/fendo.2021.719356.
5. Dinescu, S.C.; Stoica, D.; Bitu, C.E.; Nicoara, A.I.; Cirstei, M.; Staiculescu, M.A.; Vreju, F. Applications of Artificial Intelligence in Musculoskeletal Ultrasound: Narrative Review. *Front Med (Lausanne)* 2023, 10.
6. Sofoklis Katakis; Nikolaos Barotsis; Dimitrios Kastaniotis; Christos Theoharatos Muscle Type Classification on Ultrasound Imaging Deep Convolutional Neural Networks. In *Proceedings of the 2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*; 2018.
7. Panda, S.K.; Panda, C.S. A Review on Image Classification Using Bag of Features Approach. *International Journal of Computer Sciences and Engineering* 2019, 7, 538–542, doi:10.26438/ijcse/v7i6.538542.
8. Azhar, R.; Tuwohingide, D.; Kamudi, D.; Sarimuddin; Suciati, N. Batik Image Classification Using SIFT Feature Extraction, Bag of Features and Support Vector Machine. In *Proceedings of the Procedia Computer Science*; Elsevier, 2015; Vol. 72, pp. 24–30.
9. Chatterjee, S.; Dey, N.; Shi, F.; Ashour, A.S.; Fong, S.J.; Sen, S. Clinical Application of Modified Bag-of-Features Coupled with Hybrid Neural-Based Classifier in Dengue Fever Classification Using Gene Expression Data. *Med Biol Eng Comput* 2018, 56, 709–720, doi:10.1007/s11517-017-1722-y.
10. Bay, H.; Tuytelaars, T.; Gool, L. Van SURF: Speeded Up Robust Features;
11. Gul, A.; Perperoglou, A.; Khan, Z.; Mahmoud, O.; Miftahuddin, M.; Adler, W.; Lausen, B. Ensemble of a Subset of KNN Classifiers. *Adv Data Anal Classif* 2018, 12, 827–840, doi:10.1007/s11634-015-0227-5.
12. Ashour, A.S.; Guo, Y.; Hawas, A.R.; Xu, G. Ensemble of Subspace Discriminant Classifiers for Schistosomal Liver Fibrosis Staging in Mice Microscopic Images. *Health Inf Sci Syst* 2018, 6, doi:10.1007/s13755-018-0059-8.
13. Kamal, K.; Al-Timemy, A.H.; Kadhim, Z.M.; Raoof, K. A Complementary Diagnostic Tool for Diabetic Peripheral Neuropathy Through Muscle Ultrasound and Machine Learning Algorithms. *Al-Nahrain Journal for Engineering Sciences* 2024, 27, 84–90, doi:10.29194/NJES.27010084.
14. Kadhim, Z.M.; Alkhafaji, M.M. The Role of Muscle Thickness and Echogenicity in the Diagnosis of Diabetic Peripheral Neuropathy. *NeuroQuantology* 2021, 19, 113–118, doi:10.14704/nq.2021.19.8.NQ21121.
15. Al-barazanchi, K.K.; Al-Neami, A.Q.; Al-timemy, A.H. Diagnosing of Neuromuscular Disorders for Iraqi Patients Based on Single Channel Electromyography Signal Classification. In *Proceedings of the The Second Conference of Post Graduate Researches (CPGR'2017)*; 2017; pp. 132–137.
16. Ashour, A.S.; Eissa, M.M.; Wahba, M.A.; Elsayy, R.A.; Elgnainy, H.F.; Tolba, M.S.; Mohamed, W.S. Ensemble-Based Bag of Features for Automated Classification of Normal and COVID-19 CXR Images. *Biomed Signal Process Control* 2021, 68, doi:10.1016/j.bspc.2021.102656.

17. Sitaula, C.; Aryal, S. New Bag of Deep Visual Words Based Features to Classify Chest X-Ray Images for COVID-19 Diagnosis. *Health Inf Sci Syst* 2021, 9, doi:10.1007/s13755-021-00152-w.
18. König, T.; Steffen, J.; Rak, M.; Neumann, G.; von Rohden, L.; Tönnies, K.D. Ultrasound Texture-Based CAD System for Detecting Neuromuscular Diseases. *Int J Comput Assist Radiol Surg* 2015, 10, 1493–1503, doi:10.1007/s11548-014-1133-6.
19. Nodera, H.; Sogawa, K.; Takamatsu, N.; Hashiguchi, S.; Saito, M.; Mori, A.; Osaki, Y.; Izumi, Y.; Kaji, R. Texture Analysis of Sonographic Muscle Images Can Distinguish Myopathic Conditions; 2019; Vol. 66;.
20. Ahmed, A.H.; Youssef, S.M.; Ghatwary, N.; Ahmed, M.A. Myositis Detection From Muscle Ultrasound Images Using a Proposed YOLO-CSE Model. *IEEE Access* 2023, 11, 107533–107547, doi:10.1109/ACCESS.2023.3320798.
21. Varghese, A.; Bianchi, S. Ultrasound of Tibialis Anterior Muscle and Tendon: Anatomy, Technique of Examination, Normal and Pathologic Appearance. *J Ultrasound* 2014, 17, 113–123, doi:10.1007/s40477-013-0060-7.
22. Bianchi, S.; Beaulieu, J.Y.; Poletti, P.A. Ultrasound of the Ulnar–Palmar Region of the Wrist: Normal Anatomy and Anatomic Variations. *J Ultrasound* 2020, 23, 365–378, doi:10.1007/s40477-020-00468-5.
23. Nosaka, K.; Chan, R.; Newton, M. Measurement of Biceps Brachii Muscle Cross-Sectional Area by Extended-Field-of-View Ultrasound Imaging Technique. *Original article Kinesiology Slovenica* 2012, 18, 36–44.
24. Mickle, K.J.; Nester, C.J.; Crofts, G.; Steele, J.R. Reliability of Ultrasound to Measure Morphology of the Toe Flexor Muscles. *J Foot Ankle Res* 2013, 6, doi:10.1186/1757-1146-6-12.
25. Deng, M.; Liang, C.; Yin, Y.; Shu, J.; Zhou, X.; Wang, Q.; Hou, G.; Wang, C. Ultrasound Assessment of the Rectus Femoris in Patients with Chronic Obstructive Pulmonary Disease Predicts Poor Exercise Tolerance: An Exploratory Study. *BMC Pulm Med* 2021, 21, doi:10.1186/s12890-021-01663-8.
26. Xiao, T.G.; Cartwright, M.S. Ultrasound in the Evaluation of Radial Neuropathies at the Elbow. *Front Neurol* 2019, 10.
27. Hiba, C.; Hamid, Z.; Jadida, E.; ALHEYANE Omar, M. Bag of Features Model Using the New Approaches: A Comprehensive Study; 2016; Vol. 7;.
28. Vinoharan, V.; Ramanan, A. An Efficient Bag-of-Feature Representation for Object Classification. *Electronic Letters on Computer Vision and Image Analysis* 2021, 20, 51–68, doi:10.5565/rev/elcvia.1403.
29. Ho, K. The Random Subspace Method for Constructing Decision Forests; 1998; Vol. 20;.
30. Rashid, M.; Mustafa, M.; Sulaiman, N.; Abdullah, N.R.H.; Samad, R. Random Subspace K-NN Based Ensemble Classifier for Driver Fatigue Detection Utilizing Selected EEG Channels. *Traitement du Signal* 2021, 38, 1259–1270, doi:10.18280/ts.380501.
31. Hajibabae, P.; Pourkamali-Anaraki, F.; Hariri-Ardebili, M.A. An Empirical Evaluation of the T-SNE Algorithm for Data Visualization in Structural Engineering. 2021.
32. Arora, S.; Hu, W.; Kothari, P.K. An Analysis of the T-SNE Algorithm for Data Visualization. In *Proceedings of the Proceedings of Machine Learning Research*; ML Research Press, 2018; Vol. 75, pp. 1455–1462.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.