Mathematical and Computational Applications

MDPI

Article

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

Kadhim K. Al-Barazanchi [1,*], Ali H. Al-Timemy [2] and Zahid M. 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; ali.altimemy@kecbu.uobaghdad.edu.iq
[3] College of Medicine, University of Babylon, Babylon 51001, Iraq; med.zahid@uobabylon.edu.iq
* Correspondence: st.kadhim.k.hasan@ced.nahrainuniv.edu.iq

Abstract: Muscle ultrasound quantification is a valuable complementary diagnostic tool for diabetic peripheral neuropathy (DPN), enhancing physicians' diagnostic capabilities. Quantitative assessment is generally regarded as more reliable and sensitive than visual evaluation, which often necessitates specialized expertise. 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 neighbor (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%. CAD systems can effectively diagnose muscle pathology, thereby addressing limitations and identifying issues in individuals with diabetes. This research underscores muscle ultrasound as a promising diagnostic tool to aid physicians in making accurate diagnoses, streamlining workflow, and uncovering muscle-related complications in DPN patients.

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

Citation: Al-Barazanchi, K.K.; Al-Timemy, A.H.; Kadhim, Z.M. Bag of Feature-Based Ensemble Subspace KNN Classifier in Muscle Ultrasound Diagnosis of Diabetic Peripheral Neuropathy. *Math. Comput. Appl.* **2024**, *29*, 95. https://doi.org/10.3390/mca29050095

Academic Editor: Leonardo Trujillo

Received: 7 September 2024
Revised: 14 October 2024
Accepted: 18 October 2024
Published: 20 October 2024

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 computer-aided diagnostic (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 bag of features (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 through a fast-Hessian detector (it is based on the Hessian matrix because of its good performance in computation time and accuracy) 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 as follows: (i) Random subspace, which 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].

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 classification algorithms, support vector machine (SVM), and conventional k-nearest neighbor (KNN) 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.

We have organized the content in this paper into the following sections. Section 2 provides a literature review of related work. Section 3 explains our research methodology, including the dataset used, the experimental setup, and performance evaluation. In Section 4, we thoroughly analyze the results to assess the reliability of our proposed approach. Section 5 discusses the significance of the results achieved and the limitations of this study. Section 6 presents concluding remarks and future work.

2. Related Work

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, SVM, and KNN classifiers for myositis detection and achieved a classification accuracy of 87%. Nodera et al. [19] quantitatively analyzed 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 achieved an accuracy of 98%

for binary and multiclass classification. In a recent study, Noda et al. [21] developed a real-time automatic texture analysis of muscle ultrasound for diagnosing neuromuscular disorders. They extracted texture features (dissimilarity and homogeneity) from the region of interest based on a random selection of size and location for classification using a random forest algorithm. The results achieved an accuracy of 81% in distinguishing patients. 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 DPNs. Table 1 summarizes works operated on different methodologies.

Table 1. Comparison of the latest studies on muscle ultrasound diagnosis.

Study	Algorithm	Model	No. of Subjects/Images	Accuracy
Konig et al. [18]	Machine learning	2D-DWT + PCA + SVM	18/60	87%
Nodera et al. [19]	Machine learning	Texture features + (simple logistic, SVM, and random forest)	51/51	78.4%
Ahmed et al. [20]	Deep learning	YOLO-CSE + SPPF+ ELU	80/3214	98%
Noda et al. [21]	Machine learning	Texture features + random forest	75/17,021	81%

3. 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 the DPN classification system. It starts with collecting muscle ultrasound images, extracting the BOF dataset, classifying through 5-fold cross-validation, and performance evaluation based on metric measurements. 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.

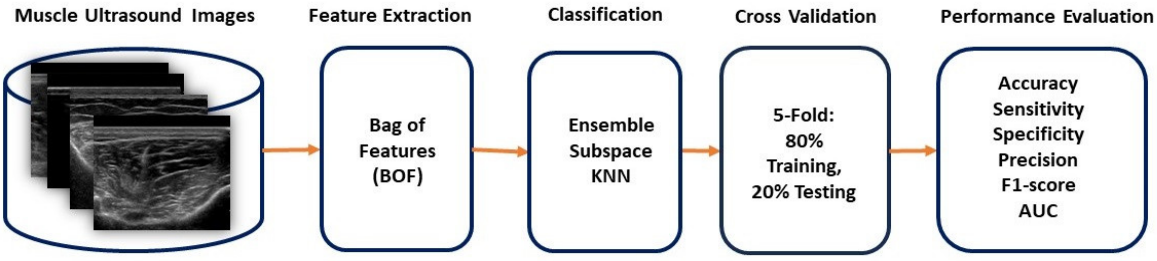

Figure 1. Proposed BOF-based ensemble subspace KNN classification system for DPN diagnosis.

3.1. Dataset

A case–control study comprising 26 DPN patients with type 2 diabetic mellites and 27 healthy controls (CTR) was conducted. Table 2 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. The exclusion criteria of all groups considered all kinds of other neuromuscular disorders and other diseases like

kidney failure and cancer and included only diabetic peripheral neuropathy patients and normal subjects without any disorders and other diseases.

Table 2. 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 was length-dependent, the investigation was required for each leg, from the proximal to distal muscles. Three of the muscles were in the upper limb: biceps brachii (BB), brachioradialis (BR), and abductor digiti minimi (ADM). The other three muscles were in the lower limb: rectus femoris (RF), tibialis anterior (TA), and abductor hallucis brevis (AHB). The total number of images per muscle with the distribution among classes is in Table 3.

Table 3. Ultrasound image distribution among muscles and classes.

Muscle	DPN	Control	Total Number of Images
BB	228	179	407
BR	238	228	466
ADM	308	297	605
RF	281	233	514
TA	258	248	506
AHB	359	364	723

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 [22–27]. Based on the manufacturer's studies, a particular preset (already saved setting in the ultrasound machine) was used for filming muscle ultrasound. The musculoskeletal general preset was used with a setting of 50% gain, compression of 62, medium pressure, and a depth of 3 cm. Ultrasound images of the muscle belly were taken, and the best image displaying muscle fibers in a cross-sectional view with the bone labeled in the background 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. In our study, we concentrated on muscle ultrasound images. We cropped the images to display only the muscle area, removing irrelevant information like annotation and subject data. The image size for all ultrasound data sets was 521×721 pixels for each of the six muscles.

3.2. 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 [28], 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 [29].

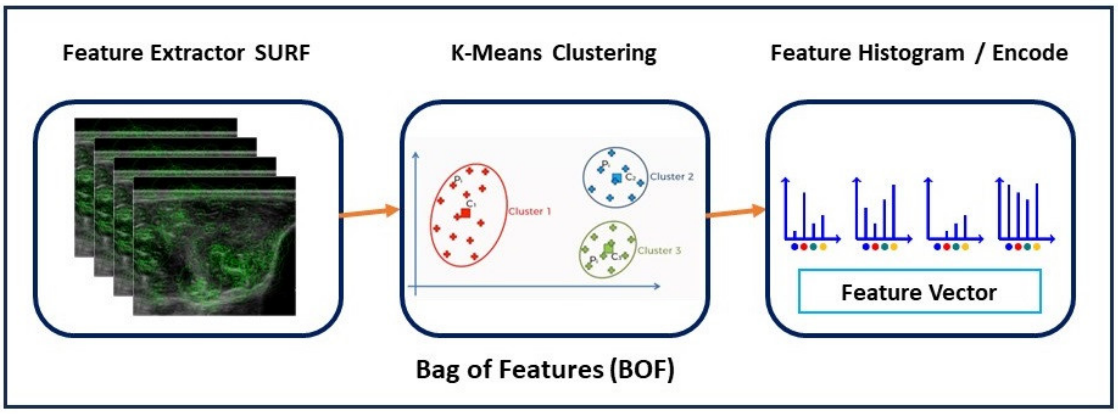

Figure 2. Illustration of BOF steps.

3.2.1. 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 employed the grid method and the SURF descriptor algorithm to meticulously pinpoint key features and extract their unique characteristic vectors. By superimposing a precise, uniform 8×8 grid onto the image, we identified the specific points of interest. The process of extracting features from muscle ultrasound images involves a comprehensive analysis of low-level, middle-level, and high-level attributes. The SURF descriptor operates by summing the Haar wavelet response around the point of interest, leveraging the integral image to minimize the computational load. In detecting interest points, the SURF algorithm utilizes 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, ensuring efficient processing. Additionally, for orientation assignment, the SURF algorithm incorporates wavelet responses in both horizontal and vertical directions, meticulously factoring in 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 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.

3.2.2. 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 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.

3.2.3. 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, …, and d_n. Tre 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, …, and w_k, where k is the number of clusters. The data can be summarized in an N × 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.

3.3. 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 this study was KNN, while the ensemble approach was a random subspace. The ensemble 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 traditional KNN architecture, the selection of neighbors and the choice of distance metric play crucial roles. Ensemble classifiers, depicted in Figure 3, amalgamate the results of several base classifiers to heighten the overall classification accuracy. Although the KNN algorithm is robust, it is susceptible to alterations in the feature set. To address this, ensemble systems fashion individual classifiers from randomly chosen data subspaces and then aggregate their outputs using a majority vote to yield the final result. Subspace KNN encompasses projecting all points onto a specified subspace and calculating distances to ascertain the K nearest neighbors. Utilizing multiple designated subspaces is instrumental in determining the majority vote on the class membership of the test sample by amalgamating the K nearest neighbors [30,31]. In the random subspace ensemble model, multiple KNN classifiers are trained using different subsets of features randomly selected from the complete dataset. This approach enables training with smaller feature subsets and improves the feature-to-instance ratio. For the specific task of classifying muscle ultrasound images, a subset of features with fewer dimensions is randomly selected from the original dataset, and KNN classifiers are trained using these feature subsets. This process is repeated multiple times to train

multiple classifiers, each using different feature subsets. The predictions of these classifiers are then combined using a majority voting approach. This study utilized the subspace KNN method for training, employing 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.

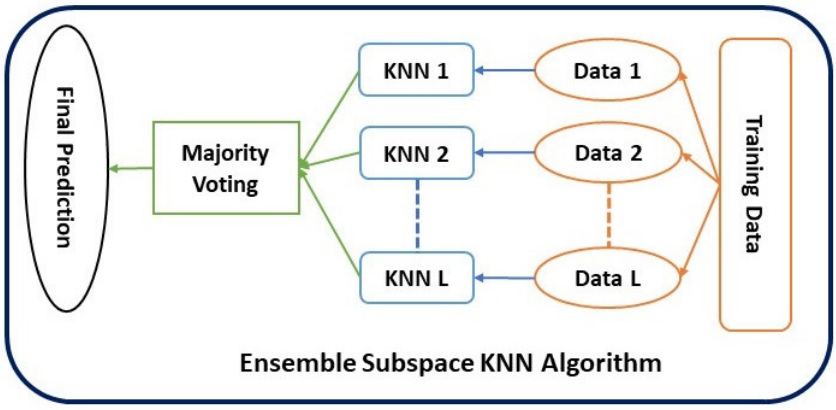

Figure 3. Ensemble subspace KNN algorithm architecture.

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.

3.4. Performance Evaluation

We employed 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 is beneficial for visualizing data with thousands of dimensions, such as feature vectors from a deep transfer learning model [32,33].

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:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{P} + \text{N}} \tag{1}$$

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:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

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:

$$\text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}} \tag{3}$$

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

$$\text{F1-Score} = \frac{2\ \text{TP}}{2\ \text{TP} + \text{FP} + \text{FN}} \tag{4}$$

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

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5}$$

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.

4. 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, which is considered a state-of-the-art model for our research.

4.1. 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 dataset was used to select optimal values using empirical analysis to find the best setting for optimizing the BOF. In this study, 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.

Table 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.

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

K-Value	SVM	KNN	Ensemble Subspace KNN
200	89.09%	93.39%	94.71%
250	88.93%	93.06%	95.04%
300	90.41%	93.88%	95.04%
350	88.76%	94.71%	95.70%
400	90.91%	95.37%	97.23%

Tables 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.

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

K-Value	SVM	KNN	Ensemble Subspace KNN
200	88.14%	92.88%	91.89%
250	89.72%	92.88%	95.25%
300	87.94%	94.66%	94.86%
350	89.13%	92.89%	95.45%
400	90.12%	95.26%	95.85%

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

K-Value	SVM	KNN	Ensemble Subspace KNN
200	88.91%	91.25%	93.77%
250	90.86%	92.41%	95.14%
300	89.88%	91.83%	95.14%
350	91.63%	91.05%	95.11%
400	92.41%	94.16%	95.70%

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

K-Value	SVM	KNN	Ensemble Subspace KNN
200	85.62%	89.06%	88.63%
250	88.63%	89.91%	91.85%
300	88.63%	92.06%	93.99%
350	89.70%	89.06%	91.85%
400	89.27%	93.13%	95.17%

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

K-Value	SVM	KNN	Ensemble Subspace KNN
200	90.46%	89.35%	90.87%
250	88.38%	90.18%	92.39%
300	89.90%	89.21%	91.59%
350	91.01%	90.46%	92.39%
400	92.12%	91.70%	92.67%

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

K-Value	SVM	KNN	Ensemble Subspace KNN
200	81.53%	82.49%	86.81%
250	76.50%	83.69%	86.33%
300	81.77%	83.45%	87.05%
350	83.69%	83.45%	87.05%
400	85.13%	85.61%	89.00%

The proposed ensemble subspace KNN algorithms demonstrated superior classification accuracy compared to KNN and SVM with a K value of 400 visual words BOF. They achieved performance accuracy of 95.85% for TA muscle, 95.70% for RF muscle, 95.17% for BR muscle, 92.67% for AHB, and 89.00% for BB muscle. In comparison, KNN achieved performance accuracy of 95.26%, 94.16%, 93.13%, 91.7%, and 85.61%, while SVM achieved 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.

Increasing the number of visual words from 200 to 400 positively impacted the performance of ensemble, KNN, and SVM algorithms by improving classification accuracy metrics. The analysis revealed that by increasing the number of visual words K, there was an average improvement in performance accuracy of nearly 5% across the six muscle ultrasounds for the classification algorithms.

4.2. 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 4 and 5 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 selecting ensemble subspace KNN classification algorithms. Furthermore, qualitative observations indicated that t-SNE significantly improved the classification accuracy in the training and validation datasets.

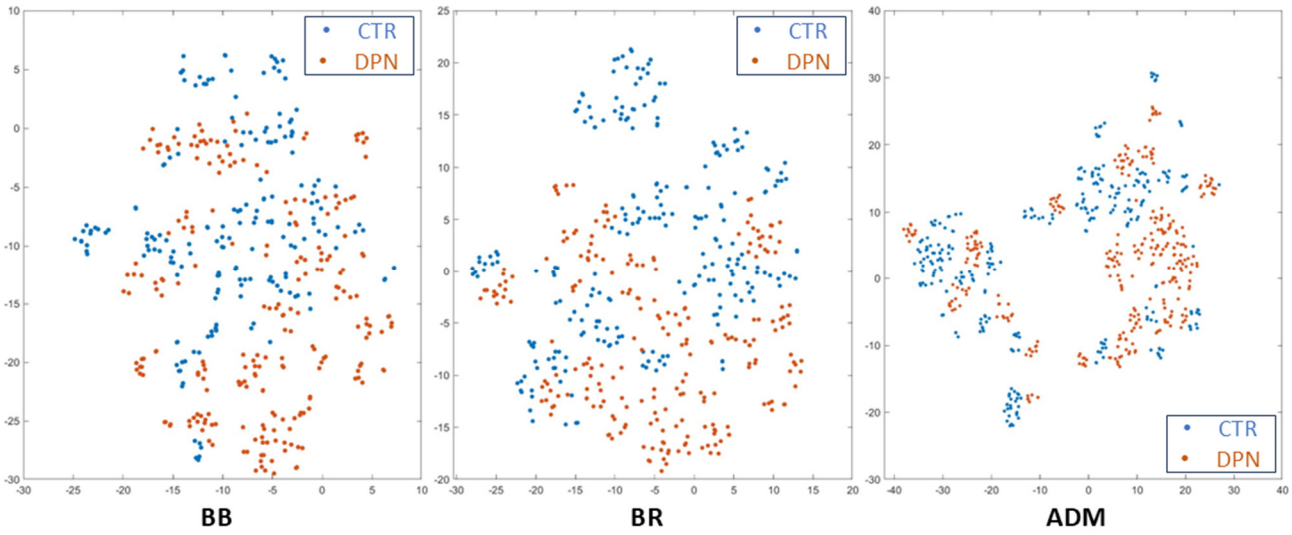

Figure 4. The t-SNE display for the upper limb muscles (BB, BR, and ADM) BOF dataset with 400 visual words.

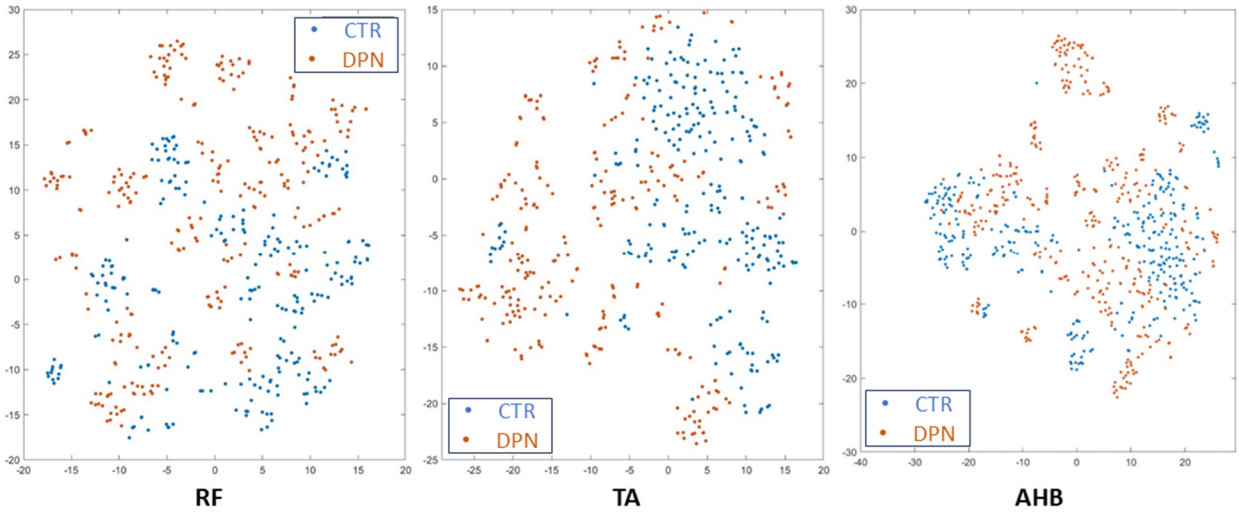

Figure 5. The t-SNE display for the lower limb muscles (RF, TA, and AHB) BOF dataset with 400 visual words.

4.3. Classification Performance Ensemble Subspace KNN

The classification evaluation results are based on the performance metrics listed in Table 10 for the feature datasets. We used five-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 89.00%, sensitivity of 89.70%, 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 92.67%, 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.

Table 10. 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%

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, the 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 6 shows the confusion matrix for the binary classifier for 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 predicted 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 grayscale representation. These changes will be reflected in the detected features and improve the classification accuracy of DPN cases.

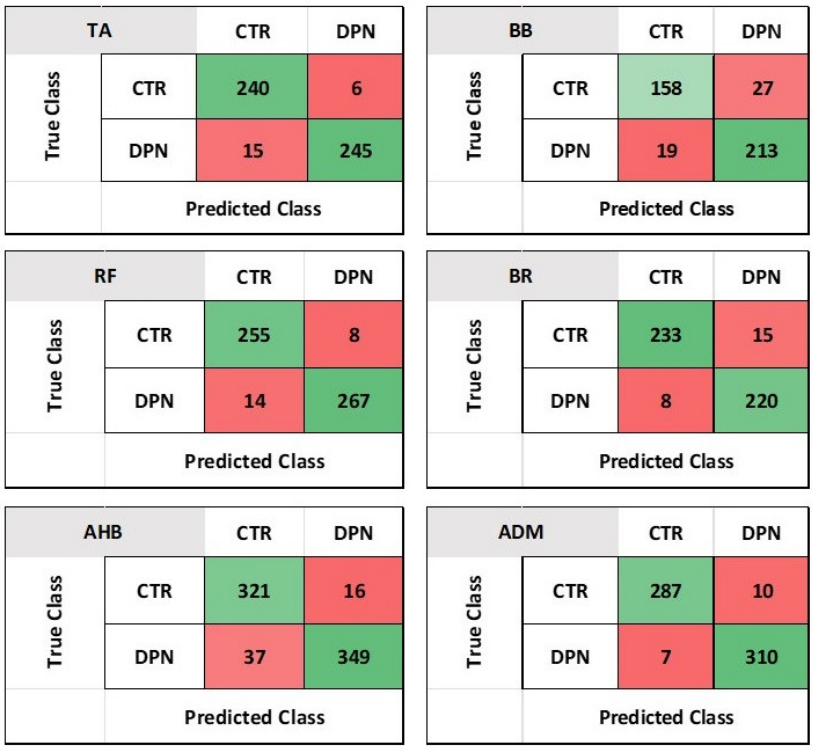

Figure 6. Confusion matrix of the ensemble-BOF-based classifier with 400 visual words.

Additionally, Figure 7 shows the ROC curve for the ensemble subspace KNN classifier based on the muscle ultrasound images coded with 400 visual words for the ADM muscle ultrasound images, where we achieved the highest classification accuracy compared with other muscles.

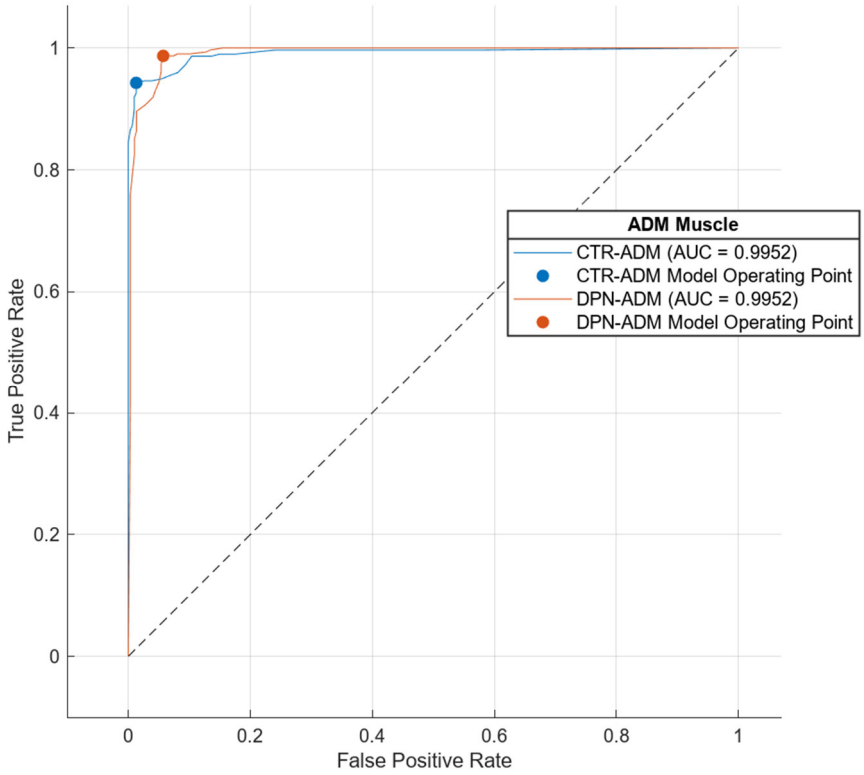

Figure 7. ROC curve of the ensemble-BOF-based classifier with 400 visual words for ADM.

5. 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. The comparison in Table 11 provides insights into various neuromuscular disease detection approaches and their corresponding accuracy rates. Our proposed model for DPN detection achieved the highest accuracy rate of 97.23% in DPN detection among the listed approaches.

Table 11. Comparison of the muscle ultrasound diagnosis approaches.

Author	Approach	No. of Subjects/Images	Accuracy
Konig et al. [18]	2D-DWT + PCA + SVM	18/60	87%
Nodera et al. [19]	Texture features + (simple logistic, SVM, and random forest)	51/51	78.4%
Ahmed et al. [20]	YOLO-CSE + SPPF+ ELU	80/3214	98%
Noda et al. [21]	Texture feature (dissimilarity and homogeneity) + random forest	75/17,021	81%
Our Model	BOF + Ensemble Subspace KNN	53/3231	97.23%

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. This preponderance of the ensemble-BOF-based model against texture-based classification through conventional SVM and random forest models in [18,19,21] and complex deep-learning techniques was examined in [20] based on performance for distinguishing between positive and negative cases and stability across iterations. However, this study has some limitations, including the number of subjects and images used for the analysis, which require further investigation. Also, this proposed method can be applied to other neuromuscular diseases to show its capabilities for detecting pathological changes in muscle ultrasound images. For future work, the authors will investigate adding more muscle ultrasound images and analyze them with texture features to find pathological changes in muscle tissue. Also, it recommended investigating muscle ultrasound with convolutional neural networks (CNNs), YOLO (versions 8–10), and segment anything model (SAM), to improve the accuracy of DPN detection with a larger dataset of MUS.

6. 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 this study, we researched the bag of features classification model. This 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.

Author Contributions: Conceptualization, K.K.A.-B., A.H.A.-T. and Z.M.K.; methodology, K.K.A.-B. and A.H.A.-T.; software, K.K.A.-B. and A.H.A.-T.; validation, K.K.A.-B., A.H.A.-T. and Z.M.K.; formal analysis, K.K.A.-B.; investigation, A.H.A.-T.; resources, K.K.A.-B.; data curation, K.K.A.-B.; writing—original draft preparation, K.K.A.-B.; writing—review and editing, A.H.A.-T.; visualization, A.H.A.-T.; supervision, A.H.A.-T. and Z.M.K.; project administration, K.K.A.-B. and A.H.A.-T.; funding acquisition, K.K.A.-B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data is unavailable due to privacy and ethical restrictions.

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

References

1. Hicks, C.W.; Selvin, E. Epidemiology of Peripheral Neuropathy and Lower Extremity Disease in Diabetes. *Curr. Diab Rep.* **2019**, *19*, 86. [CrossRef] [PubMed]
2. Bansal, V.; Kalita, J.; Misra, U.K. Diabetic Neuropathy. *Postgrad. Med. J.* **2006**, *82*, 95–100. [CrossRef] [PubMed]
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*, 3554. [CrossRef] [PubMed]
4. Yu, Y. Gold Standard for Diagnosis of DPN. *Front. Endocrinol.* **2021**, *12*, 719356. [CrossRef]
5. Dinescu, S.C.; Stoica, D.; Bita, C.E.; Nicoara, A.I.; Cirstei, M.; Staiculesc, M.A.; Vreju, F. Applications of Artificial Intelligence in Musculoskeletal Ultrasound: Narrative Review. *Front. Med.* **2023**, *10*, 1286085. [CrossRef]
6. Katakis, S.; Barotsis, N.; Kastaniotis, D.; Theoharatos, C. 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), Zagori, Greece, 10–12 June 2018.
7. Panda, S.K.; Panda, C.S. A Review on Image Classification Using Bag of Features Approach. *Int. J. Comput. Sci. Eng.* **2019**, *7*, 538–542. [CrossRef]
8. Azhar, R., Tuwohingide; Suciati, N. Batik Image Classification Using SIFT Feature Extraction, Bag of Features and Support Vector Machine. *Proc. Procedia Comput. Sci.* **2015**, *72*, 24–30. [CrossRef]
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. [CrossRef]
10. Bay, H.; Tuytelaars, T.; Gool, L. Van SURF: Speeded Up Robust Features. *Comput. Vis. Image Underst.* **2008**, *110*, 346–359. [CrossRef]
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. [CrossRef]
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*, 21. [CrossRef] [PubMed]
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 J. Eng. Sci.* **2024**, *27*, 84–90. [CrossRef]
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. [CrossRef]
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 Second Conference of Post Graduate Researches (CPGR'2017), Baghdad, Iraq, 4 October 2017; pp. 132–137.
16. Ashour, A.S.; Eissa, M.M.; Wahba, M.A.; Elsawy, 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*, 102656. [CrossRef]
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*, 24. [CrossRef]
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. [CrossRef]
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. *J. Med. Invest.* **2019**, *66*, 237–247. [CrossRef]
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. [CrossRef]
21. Noda, Y.; Sekiguchi, K.; Matoba, S.; Suehiro, H.; Nishida, K.; Matsumoto, R. Real-Time Artificial Intelligence-Based Texture Analysis of Muscle Ultrasound Data for Neuromuscular Disorder Assessment. *Clin. Neurophysiol. Pract.* **2024**, *9*, 242–248. [CrossRef]
22. 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. [CrossRef]
23. 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. [CrossRef] [PubMed]
24. Nosaka, K.; Chan, R.; Newton, M. Measurement of Biceps Brachii Muscle Cross-Sectional Area by Extended-Field-of-View Ultrasound Imaging Technique. *Orig. Artic. Kinesiol. Slov.* **2012**, *18*, 36–44.
25. 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*, 12. [CrossRef] [PubMed]
26. 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*, 304. [CrossRef]

27. Xiao, T.G.; Cartwright, M.S. Ultrasound in the Evaluation of Radial Neuropathies at the Elbow. *Front. Neurol.* **2019**, *10*, 216. [CrossRef]
28. Hiba, C.; Hamid, Z.; Jadida, E.; Alheyane Omar, M. Bag of Features Model Using the New Approaches: A Comprehensive Study. *Int. J. Adv. Comput. Sci. Appl.* **2016**, *7*, 226–234. [CrossRef]
29. Vinoharan, V.; Ramanan, A. An Efficient Bag-of-Feature Representation for Object Classification. *Electron. Lett. Comput. Vis. Image Anal.* **2021**, *20*, 51–68. [CrossRef]
30. Ho, K. The Random Subspace Method for Constructing Decision Forests. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 832–844.
31. 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. *Trait. Signal* **2021**, *38*, 1259–1270. [CrossRef]
32. Hajibabaee, P.; Pourkamali-Anaraki, F.; Hariri-Ardebili, M.A. An Empirical Evaluation of the T-SNE Algorithm for Data Visualization in Structural Engineering. *arXiv* **2021**, arXiv:2109.08795.
33. Arora, S.; Hu, W.; Kothari, P.K. An Analysis of the t-SNE Algorithm for Data Visualization. *Proc. Mach. Learn. Res.* **2018**, *75*, 1455–1462.