9-9-2024

Interpretable Classification of Myositis from Muscle Ultrasound Images

Bishwa Karki

Xin Zhong

Yu-Che Chen

Chun-Hua Tsai

Interpretable Classification of Myositis from Muscle Ultrasound Images

Bishwa Karki
University of Nebraska Omaha
Omaha, Nebraska, USA
bishwakarki@unomaha.edu

Xin Zhong
University of Nebraska Omaha
Omaha, Nebraska, USA
xzhong@unomaha.edu

Yu-Ting Chen
Creighton University
Omaha, Nebraska, USA
Yu-TingChen@creighton.edu

Chun-Hua Tsai
University of Nebraska Omaha
Omaha, Nebraska, USA
chunhuatsai@unomaha.edu

ABSTRACT

This study is dedicated to designing and advancing machine learning (ML) algorithms for classifying normal and abnormal muscular tissues, thereby aiding neurologists in diagnosing inclusion body myositis (IBM). Our work mainly aims to leverage machine learning and recent state-of-the-art (SOTA) algorithms to recognize and diagnose myositis from muscle ultrasound images in the preliminary stage and support the traditional diagnostic methodology. Initially, we used an open-source ultrasound image dataset to construct and refine initial models using VGG-16. We employed the Grad-CAM method to annotate muscle ultrasound images and delineate regions of interest (ROI). Subsequent experiments enhanced the VGG16 architecture through extensive layer modifications and parameter adjustments. Our research offers valuable perspectives on utilizing ML to assist neurologists in the early diagnosis of IBM.

CCS CONCEPTS

• **Computing methodologies** → **Neural networks**; **Computer vision**.

KEYWORDS

Deep learning, Interpretable machine learning, Ultrasound imaging

ACM Reference Format:
Bishwa Karki, Xin Zhong, Yu-Ting Chen, and Chun-Hua Tsai. 2024. Interpretable Classification of Myositis from Muscle Ultrasound Images. In *2024 8th International Conference on Medical and Health Informatics (ICMHI 2024), May 17–19, 2024, Yokohama, Japan.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3673971.3673988

1 INTRODUCTION

Inclusion body myositis (IBM) represents a progressive myopathy distinguished by inflammation, weakening, and atrophy of muscles, impacting approximately 200,000 individuals annually within the United States [3]. Recognized as a form of inflammatory myopathies, IBM's hallmark symptoms encompass progressive muscular weakness predominantly affecting the finger and wrist flexors, quadriceps, and ankle dorsiflexion. Patients may also develop facial weakness and difficulty swallowing [14]. The pathogenesis of IBM remains elusive, presumably resulting from a complex interplay of genetic susceptibilities, immunological factors, and environmental influences. While the condition may have a genetic component, it is typically non-hereditary [16]. Muscle weakness is often an overlooked early symptom, commonly attributed to aging [4].

Due to the rarity of myositis, numerous cases of IBM remain undiagnosed or are incorrectly diagnosed for extended periods following the onset of initial symptoms. Diagnosing IBM is intricate, as its clinical presentation often overlaps with other myositis types or neuromuscular conditions. Traditional diagnostic criteria usually include clinical and pathological features [7]. However, muscle biopsy, required to show pathological features, is invasive and expensive, and not 100% sensitive [16]. Recent scholarly pursuits have explored the efficacy of quantitative muscle ultrasound as a more accessible and cost-effective method to distinguish IBM from similar conditions [12]. Preliminary findings suggest that ultrasound is a promising diagnostic tool, capable of differentiating IBM from other inflammatory myopathies like polymyositis (PM)/dermatomyositis (DM) and various neuromuscular disorders, underscoring its potential as a valuable diagnostic resource in identifying early IBM [12, 22].

Ultrasound interpretation in myopathies relies on analyzing muscle texture and grayscale levels to distinguish between healthy and affected muscular structures. However, the accuracy of such visual evaluations is heavily contingent upon the clinician's expertise and subjective interpretation, limiting the sensitivity of visually discerning normal from diseased muscle to approximately 70% [15]. Physicians often encounter interpretative challenges with ultrasound imagery in marginal cases, notably in patients with IBM or elderly individuals with sarcopenia. Research has highlighted the intrinsic difficulties of visual interpretation; while muscles with starkly different conditions (Heckmatt grading scale 1 and 4) can be differentiated relatively easily through increased grayscale levels, distinguishing between more subtly varying conditions (Heckmatt grading scale 2 and 3) often proves challenging, leading to ambiguity in determining muscle normality [12].

ICMHI 2024, May 17–19, 2024, Yokohama, Japan

ACM ISBN 979-8-4007-1687-4/24/05
https://doi.org/10.1145/3673971.3673988

Early and correct diagnosis is critical for potentially benefiting IBM patients, especially those with clinical suymptoms mimicking PM, DM, or amyotrophic lateral sclerosis (ALS) autoimmune conditions. Enhancing the precision of ultrasound image interpretation emerges as a promising avenue. We propose developing and training a machine-learning model to differentiate between normal and pathological muscular tissues and structures through data-driven and computational approaches to address this. Our goal is to provide *visual cues* that assist neurologists in the early detection of myopathies. This advancement could significantly enhance the quality of ultrasound interpretations and, consequently, the overall well-being of individuals with IBM. Our experiment utilized an open-source ultrasound image dataset to develop and fine-tune preliminary models based on the VGG-16 architecture. In subsequent experiments, we further improved the VGG-16 architecture through comprehensive adjustments to its layers and parameters. We employed the Grad-CAM method to annotate muscle ultrasound images and identify *Region of Interest (ROI)*. Based on the experiment results, we discuss the implications of the interpretable model in assisting neurologists' diagnosis.

2 LITERATURE REVIEW

Utilizing machine learning models presents a robust mechanism for elucidating nonlinear features from complex datasets. This methodology is particularly adept at discerning and operationalizing subtle image characteristics, imperceptible to the human eye, for determining muscular normality. The training phase involves adjusting the model's weights across various layers by leveraging training data, enabling the autonomous discovery of features and subsequent model formulation. The validation set further refines network parameters. The process involves extracting features from images to classify them into distinct categories. Such techniques have gained substantial traction in predicting ultrasound characteristics across diverse organ systems, including the breast, liver, kidney, and heart [8–10]. Automated systems have been engineered in breast lesion classification to systematically extract and analyze features from the BI-RADS lexicon, encompassing attributes like shape, margin, orientation, echo pattern, and acoustic shadowing [19].

Machine learning, particularly through Convolutional Neural Networks (CNNs), feature fusion, and bespoke neural architectures, has significantly advanced the analysis and interpretation of medical imagery, including ultrasound. For example, Le et al. [11] demonstrated the utility of deep learning in amalgamating features from various imaging modalities to improve breast cancer detection. Gao et al. [5] investigated the application of CNNs in interpreting ultrasound video content, marking a significant stride in automating dynamic ultrasound image analysis. Liu et al. [13] developed a specialized morphological neural network to enhance pneumonia detection in chest X-ray imaging. In the realm of thyroid ultrasound imaging, Gomes et al. [6] underscored the role of deep learning in mitigating observer variability, particularly in identifying thyroid nodules. Additionally, Nisha et al. [17] introduced an innovative Self-organized Operational Neural Network to detect neurological diseases using transcranial Doppler ultrasound.

There are promising advancements in the field of using machine learning for the classification of myositis from muscle ultrasound images. Burlina et al. [2] investigated the integration of ultrasound imaging with machine learning and deep learning techniques for the automated or semi-automated classification of myositis. Arunkarthick et al. [1] developed a hybrid computer-aided model to distinguish among different myositis types, including PolyMyositis, DermatoMyositis, and IBM, primarily utilizing the You Only Look Once (YOLO) framework. Uccar et al. [21] focused on creating a computer-aided diagnosis system for myositis, employing muscle ultrasound images and deep learning. Their study features an innovative model that merges the architectures of VGG16 and VGG19 to detect inflammatory myopathies.

Current trends particularly focus on enhancing the diagnostic accuracy and efficiency for myositis through muscle ultrasound imaging. Our uniqueness lies in providing explainable results to doctors, emphasizing interpretability with the inclusion of the heatmap of GradCAM. The main contribution is not solely to achieve higher accuracy but also bridging the gap between advanced image analysis techniques and their clinical applicability using interpretability of prediction.

3 METHODOLOGY

3.1 Convolutional Neural Network Modeling

We introduce a novel convolutional model to diagnose Inclusion Body Myositis (IBM) by classifying muscle ultrasound images into normal or abnormal categories. Our model employs transfer learning principles, integrating the well-established VGG-16 architecture, which was previously trained on a large and diverse dataset. This integration, coupled with additional training on a specialized dataset curated for IBM detection, represents an advancement in the automated diagnosis of muscle diseases. We chose the VGG-16 architecture as the foundation of our model due to its exemplary performance in image classification tasks. Introduced by the Visual Geometry Group [20], the VGG-16 achieved a notable 92.7% top-5 accuracy rate on the ImageNet dataset. The architecture consists of 13 convolutional layers and three fully connected layers, totaling approximately 138 million parameters. To tailor VGG-16 for diagnosing IBM, we made strategic modifications to enhance its suitability for muscle ultrasound image analysis, as shown in Fig. 1.

Layer Modification: The original VGG-16 model, designed for classifying images into 1,000 categories, required adjustments for IBM disease diagnosis. We replaced its terminal SoftMax layer with a custom-designed, fully connected layer. This layer is augmented by six additional fully connected layers, each utilizing the Rectified Linear Unit (ReLU) activation function to introduce non-linearity. The final layer of our model is tailored to classify ultrasound images into two or four distinct categories pertinent to IBM disease diagnosis.

Model Optimization: We employed a fine-tuning strategy in adapting VGG-16 to our specific diagnostic task. This involved freezing the pre-trained layers of the VGG-16 model, thereby focusing the training process exclusively on the newly added layers for IBM classification. To optimize the model, we maintained the frozen state of all original layers while applying a minimal learning rate. This approach capitalizes on VGG-16's deep architecture, which is adept at extracting hierarchical features essential for identifying

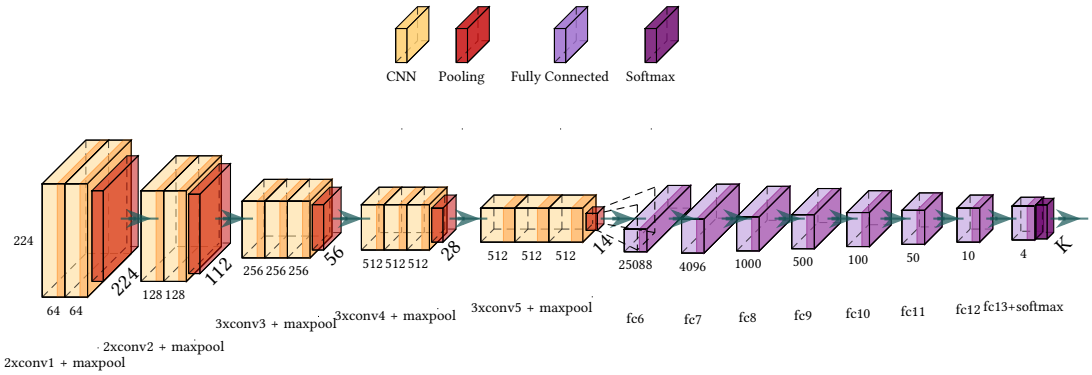

Figure 1: Proposed VGG-based architecture

complex patterns in medical imaging. Through this targeted fine-tuning, our model gains enhanced proficiency in detecting subtle indicators of IBM disease in muscle ultrasound images.

3.2 Grad-CAM: Visual Explanations

Recent progress in the field of medical imaging, particularly in ultrasound imaging for disease prediction [21], has often been criticized for its lack of interpretability. This limitation is a significant barrier to the practical application of such technologies in clinical settings, as it hinders physicians from leveraging these advancements in their practice. To address this challenge, our work introduces an enhancement in interpretability by integrating Grad-CAM [18] to improve the comprehensibility of image classification results.

Specifically, we used Grad-CAM to identify the ROI, crucial for detecting subtle abnormalities in muscles or tissues. The ROI essentially comprises a set of pixels that highlight potential areas of concern in the image. To effectively pinpoint these critical areas, we integrated Grad-CAM to generate a heatmap. This heatmap visually accentuates the core areas of the input image that significantly influence the model's prediction, providing insightful clues about the regions where the model focuses its analysis. Furthermore, the last layer of the CNN, which encapsulates the high-level features learned by the model, plays a pivotal role in this process. By examining the gradients of the predicted class with respect to the feature maps of the last CNN layer, we gain a deeper understanding of the importance of each feature map. This analysis also reveals the correlation between specific feature maps and spatial locations within the image, enhancing our ability to identify the ROI for diagnostic purposes precisely.

Integrating Grad-CAM's interpretative capabilities into the daily operations of medical professionals represents a significant advancement in clinical practice. By highlighting critical regions in ultrasound images provides valuable insights into the decision-making process of deep-learning models. This enhancement bolsters trust in the model's outputs and empowers clinicians with deeper understanding, enabling them to make more informed diagnostic decisions. Furthermore, the clarity offered by Grad-CAM extends beyond mere comprehension of the model's functionality; it plays a crucial role in refining medical diagnostics. Clinicians can leverage the visual cues provided by Grad-CAM to validate and enhance their diagnostic assessments, leading to more accurate and reliable patient outcomes. The adoption of such interpretability techniques is instrumental in addressing the challenges faced in ultrasound-based disease prediction and facilitates the practical application of deep learning models in real-world medical scenarios. Consequently, this fosters a higher degree of trust and augments the overall capabilities of medical diagnostics.

4 EXPERIMENTS AND RESULTS

4.1 Data and Model Preparation

Our study utilized a publicly available dataset, as described in [8], consisting of 3,214 muscle ultrasound images. This dataset is diverse, including images from 33 healthy individuals and patients diagnosed with various myopathies: 19 with Inclusion Body Myositis (IBM), 14 with Polymyositis, and 14 with Dermatomyositis. These images are categorized into four distinct classes: Normal (N), Dermatomyositis (D), Polymyositis (P), and IBM (I). Initially, we divided the images into training and test sets, following a 9:1 ratio. However, this division revealed an imbalanced dataset, as illustrated in Fig. 2. Our first step was to train a binary classification model to address this. In this model, the classes Dermatomyositis, Polymyositis, and IBM were consolidated into a single 'diseased' category. This approach balanced the class distribution and provided a robust framework for binary classification.

Further, we employed data augmentation and oversampling techniques to enhance the training process. These methods included horizontal flipping, random rotation, and random contrast adjustment of images during training. Such augmentation helps build a more robust model by introducing it to a broader spectrum of visual variations, thus mitigating overfitting and enhancing generalization capabilities. Additionally, to ensure the robustness and reliability of our model, especially given the small size of the dataset, we implemented K-fold cross-validation. This process involves dividing the dataset into K subsets and training the model K times. Each iteration uses a different subset as the validation set, offering a more reliable estimate of the model's performance.

4.2 Model building and training

The performance of the VGG-16 models for binary and multi-class classification was evaluated using key metrics, providing a nuanced

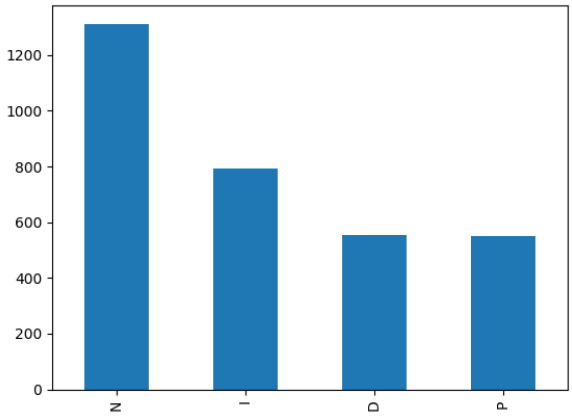

Figure 2: Data distribution

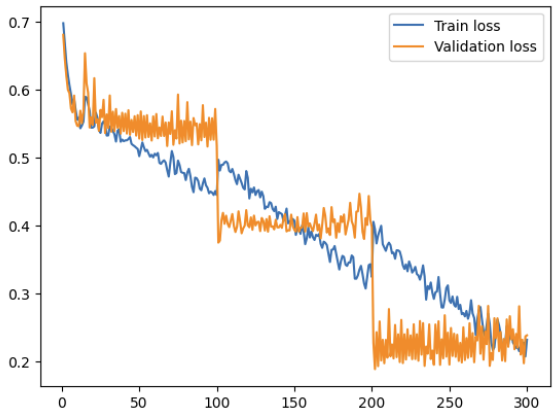

Figure 3: The loss function for K-fold based training

understanding of their effectiveness. We evaluate our model's performance using four key metrics: Accuracy, Precision, Recall, and F1-Score. Accuracy measures the overall correct predictions across all classes and is particularly useful as a stable performance indicator in an unstable loss function. Precision is critical in scenarios like medical diagnoses where minimizing false positives is imperative, providing insight into the model's capability to identify normal cases correctly. Recall becomes crucial in medical contexts to identify all true positive cases, minimizing false negatives, thereby reflecting the model's sensitivity in detecting actual positives amidst fluctuating loss function dynamics. Lastly, the F1-Score, balancing precision and recall, offers a comprehensive performance assessment, especially valuable in class-imbalanced situations, and provides a holistic view of the model's efficacy under varying loss function conditions.

In our study, the binary classification model exhibited exceptional performance, achieving high precision and recall rates of 97%, effectively minimizing false positives and negatives, as shown in Table 1. This was further corroborated by a robust F1-score of 97%, indicating a well-balanced precision and recall, contributing significantly to the model's overall efficacy. In the multi-class scenario, the model demonstrated a remarkable accuracy of 94%,

Model	Classification	Acc	Pre	Re	F1-Score
[21]	Binary	94.09	93.89	93.89	93.39
[21]	Multi	95.12	89.73	88.28	88.91
Ours	Binary	97	97	97	97
Ours	Multi	94	94	94	93

Table 1: Comparaion for binary and multiclass classification

efficiently categorizing ultrasound images into four distinct classes and maintaining balanced precision, recall, and an F1-score of 93% and 94%, respectively. This balance signifies an effective trade-off in minimizing false positives and false negatives. Additionally, employing K-fold cross-validation, particularly with 3-fold validation showing optimal results, underscores the model's robustness and reliability. This validation technique is especially crucial in scenarios with limited data, providing a more accurate estimation of the model's performance.

In comparing our model ([Ours], as shown in Table 1) with existing deep learning-based models and acknowledging the presence of other classical machine learning approaches [1, 2], we particularly emphasize the differences in classification performance in relation to [21]. In binary classification, our model demonstrates superior performance with precision, recall, and F1-Score values of 97%, outperforming [21]. For multi-class classification, our model maintains a high accuracy of 94%. The balanced precision, recall, and F1-Score values of 94% in our model are comparable to, if not better than, those achieved by [21]. This indicates our model's proficiency in categorizing ultrasound images into multiple classes. From a quantitative standpoint, our model's numerical superiority in accuracy, precision, recall, and F1-Score positions it as a more effective solution. Beyond numerical comparisons, our model offers a well-balanced trade-off between minimizing false positives and false negatives, thereby enhancing its overall efficacy. The comprehensive evaluation across multiple metrics underscores the robustness and superiority of our model when compared to existing approaches, including [21] and [2].

4.3 Model Interpretability

Based on the results, we observed that CNN-based models can achieve satisfactory classification outcomes in ultrasound-based disease recognition. These models effectively address the challenge of limited data by leveraging feature transfer from other sources, thereby obviating the need to train the entire network from scratch. Our objective is to provide visual cues that assist neurologists in the early detection of myopathies. We employed the Grad-CAM method to annotate muscle ultrasound images and identify regions of interest (ROI) to achieve this. We present two sets of ultrasound images. In Fig. 4(a), we illustrate ROIs with two labels – heatmap and circle – highlighting the areas our proposed model identifies as significant. In Fig. 4(b), a normal ultrasound image is displayed alongside a potential ROI, allowing the physician to determine if the area warrants further examination. The colors represent the gradients of the predicted class concerning the feature maps of the last CNN layer.

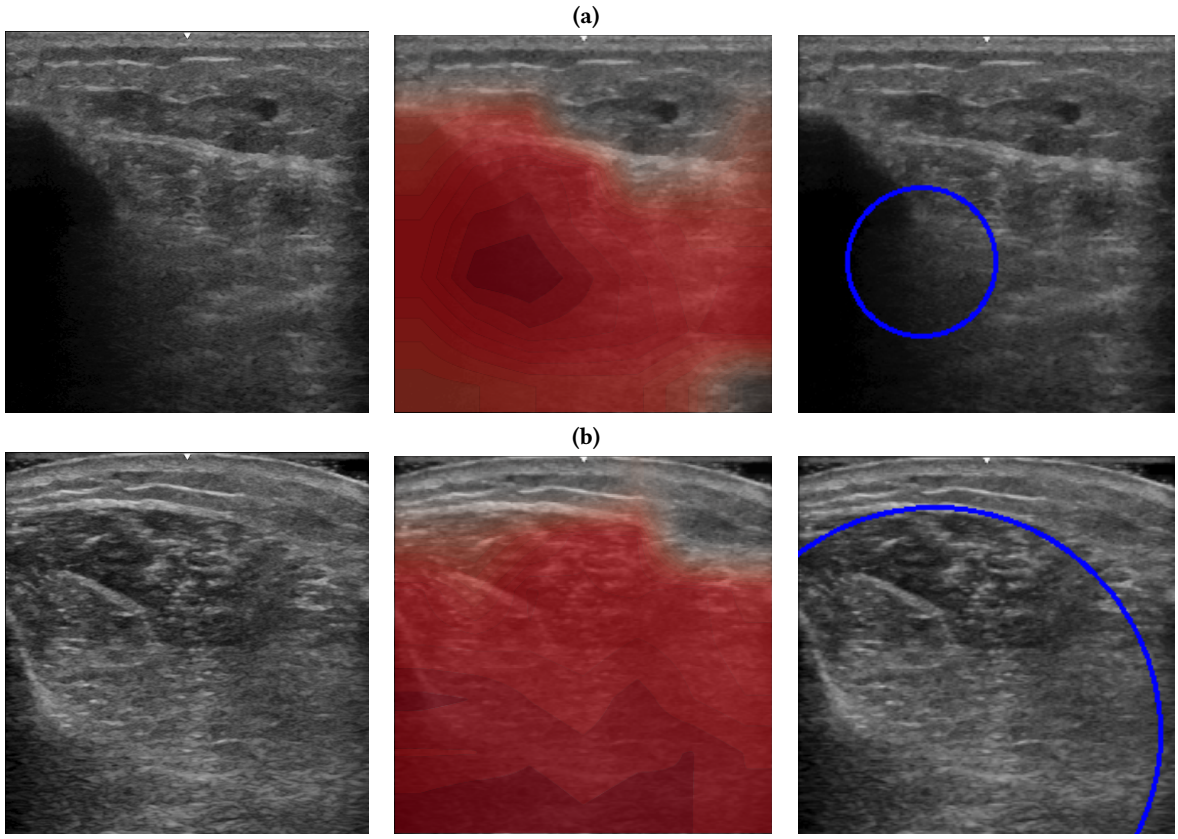

Figure 4: Plots of Grad-CAM-Based Feature Labeling for the Region of Interest (ROI) in Ultrasound Images: Panel (a) displays ultrasound images of a patient diagnosed with Inclusion Body Myositis (IBM), while Panel (b) showcases ultrasound images from a healthy individual for comparison. Each panel presents the images in a sequence from left to right: the original image, the heatmap overlay, and the circular labels.

Our study argued that providing visual cues could achieve two primary objectives. Firstly, it enables physicians to review the labeled ROIs to ascertain their accuracy and determine if further attention or examination is required. This capability is particularly beneficial in interpreting ultrasound images in marginal cases, where traditional interpretation can be challenging. The labeled ROI can be integrated with traditionally manually selected regions (i.e., by plotting the area using a mouse or touchscreen), offering the physician an additional layer of confidence based on the overlap of these layers. Secondly, visual cues provide a low-overhead method for gathering physician feedback on the machine learning models as ground truth. Physicians can more easily assess whether the labeled area aligns with the patient's diagnostic results. This feedback can be systematically collected and incorporated into our proposed model to further refine its accuracy and performance.

5 CONCLUSIONS

Our research demonstrates that leveraging advanced machine learning techniques, particularly the fine-tuned VGG-16 architecture integrated with the Grad-CAM method, can improve the precision of ultrasound image classification and interpretation in diagnosing Inclusion Body Myositis (IBM). This approach presents a promising solution to traditional visual evaluations' inherent limitations, which rely heavily on clinician expertise and often result in subjective interpretations. Our work demonstrates how, by providing clear visual cues and identifying regions of interest in muscle ultrasound images, our model aids neurologists in the early detection of myopathies. This is particularly crucial in challenging cases such as IBM and sarcopenia in elderly patients. The interpretable nature of our model not only enhances the accuracy of IBM diagnosis but also contributes significantly to the broader field of neuromuscular disorders, facilitating early intervention and treatment.

REFERENCES

[1] AK Arunkarthick and Srinivasa Reddy. 2023. Prediction of myositis disease using machine learning algorithm. In *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*. IEEE, 534–539.

[2] Philippe Burlina, Seth Billings, Neil Joshi, and Jemima Albayda. 2017. Automated diagnosis of myositis from muscle ultrasound: Exploring the use of machine learning and deep learning methods. *PloS one* 12, 8 (2017), e0184059.

[3] Enrique J deAndrés Galiana, Francisco Javier Fernández-Ovies, Ana Cernea, Juan Luis Fernández-Martínez, and Andrzej Kloczkowski. 2020. Deep neural networks for phenotype prediction in rare diseases: Inclusion body myositis: A case study. In *Artificial Intelligence in Precision Health*. Elsevier, 189–202.

[4] Mazen M Dimachkie and Richard J Barohn. 2014. Inclusion body myositis. *Neurologic Clinics* 32, 3 (2014), 629–646.
[5] Yuan Gao, Mohammad Ali Maraci, and J Alison Noble. 2016. Describing ultrasound video content using deep convolutional neural networks. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 787–790.
[6] Elmer Jeto Gomes Ataide, Mathews S Jabaraj, Simone Schenke, Manuela Petersen, Sarvar Haghghi, Jan Wuestemann, Alfredo Illanes, Michael Friebe, and Michael C Kreissl. 2023. Thyroid Nodule Detection and Region Estimation in Ultrasound Images: A Comparison between Physicians and an Automated Decision Support System Approach. *Diagnostics* 13, 18 (2023), 2873.
[7] D. Hilton-Jones and S. Brady. 2016. Diagnostic criteria for inclusion body myositis. *Journal of Internal Medicine* 280, 1 (2016), 52–62. https://doi.org/10.1111/joim.12480 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/joim.12480
[8] Yuya Hiramatsu, Chisako Muramatsu, Hironobu Kobayashi, Takeshi Hara, and Hiroshi Fujita. 2017. Automated detection of masses on whole breast volume ultrasound scanner: false positive reduction using deep convolutional neural network. In *Medical imaging 2017: Computer-aided diagnosis*, Vol. 10134. Spie, 717–722.
[9] Yoo Na Hwang, Ju Hwan Lee, Ga Young Kim, Yuan Yuan Jiang, and Sung Min Kim. 2015. Classification of focal liver lesions on ultrasound images by extracting hybrid textural features and using an artificial neural network. *Bio-medical materials and engineering* 26, s1 (2015), S1599–S1611.
[10] Hanan Khamis, Grigoriy Zurakhov, Vered Azar, Adi Raz, Zvi Friedman, and Dan Adam. 2017. Automatic apical view classification of echocardiograms using a discriminative learning dictionary. *Medical image analysis* 36 (2017), 15–21.
[11] Hung Le Minh, Manh Mai Van, and Tran Van Lang. 2019. Deep feature fusion for breast cancer diagnosis on histopathology images. In *2019 11th International Conference on Knowledge and Systems Engineering (KSE)*. IEEE, 1–6.
[12] Kristofoor E Leeuwenberg, Nens van Alfen, Lisa Christopher-Stine, Julie J Paik, Eleni Tiniakou, Christopher Mecoli, Jonne Doorduin, Christiaan GJ Saris, and Jemima Albayda. 2020. Ultrasound can differentiate inclusion body myositis from disease mimics. *Muscle & Nerve* 61, 6 (2020), 783–788.
[13] Shaobo Liu, Frank Y Shih, and Xin Zhong. 2021. Classification of chest X-ray images using novel adaptive morphological neural networks. *International Journal of Pattern Recognition and Artificial Intelligence* 35, 10 (2021), 2157006.
[14] Thomas E Lloyd, Andrew L Mammen, Anthony A Amato, Michael D Weiss, Merrilee Needham, and Steven A Greenberg. 2014. Evaluation and construction of diagnostic criteria for inclusion body myositis. *Neurology* 83, 5 (2014), 426–433.
[15] Jean K Mah and Nens van Alfen. 2018. Neuromuscular ultrasound: clinical applications and diagnostic values. *Canadian Journal of Neurological Sciences* 45, 6 (2018), 605–619.
[16] Elie Naddaf. 2022. Inclusion body myositis: Update on the diagnostic and therapeutic landscape. *Frontiers in Neurology* 13 (2022), 1020113.
[17] Naima Nasrin Nisha, Kanchon Kanti Podder, Muhammad EH Chowdhury, Mamun Rabbani, Md Sharjis Ibne Wadud, Somaya Al-Maadeed, Sakib Mahmud, Amith Khandakar, and Susu M Zughaier. 2023. A Deep Learning Framework for the Detection of Abnormality in Cerebral Blood Flow Velocity Using Transcranial Doppler Ultrasound. *Diagnostics* 13, 12 (2023), 2000.
[18] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. 2019. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *International Journal of Computer Vision* 128, 2 (Oct. 2019), 336–359. https://doi.org/10.1007/s11263-019-01228-7
[19] Juan Shan, S Kaisar Alam, Brian Garra, Yingtao Zhang, and Tahira Ahmed. 2016. Computer-aided diagnosis for breast ultrasound using computerized BI-RADS features and machine learning methods. *Ultrasound in medicine & biology* 42, 4 (2016), 980–988.
[20] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).
[21] Emine Uçar. 2022. Classification of myositis from muscle ultrasound images using deep learning. *Biomedical Signal Processing and Control* 71 (2022), 103277.
[22] Juerd Wijntjes and Nens van Alfen. 2021. Muscle ultrasound: Present state and future opportunities. *Muscle & nerve* 63, 4 (2021), 455–466.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009