

Changes after review for paper:
Explainable Few-Shot Learning for Multiple
Sclerosis Detection in Low-Data Regime

- In the introduction (Section 1):
 - We added references for this sentence “Studies indicate that over half of the patients were misdiagnosed for a period exceeding three years”.
 - We rewrote Wang et al. limitation sentence to be “Explainability remains unexplored”.
 - We added our motivation of using few-shot learning in MS detection “Additionally, the collection of MS and other demyelinating diseases data is challenging due to the variability in disease presentation, limited patient availability, and the high cost of medical imaging. Therefore, the application of few-shot learning is essential to leverage limited data effectively.”
 - We added a sentence about the choice of FLAIR sequence of MRI “Furthermore, a key finding in MS identification is the presence of white matter lesions in the brain, detectable via Fluid Attenuated Inversion Recovery (FLAIR) sequence of MRI.”
 - We wrote the sentence about focusing on binary classification: “This study focuses on distinguishing MS from other demyelinating diseases.”
- In Section 2.1 (Architecture overview)
 - We omitted the mathematical annotation table
- In Section 2.3 (Few-shot learning):
 - We added ProtoNet for abbreviation
 - We articulated ProtoNet implementation details “We employed Euclidian distance for our ProtoNet to calculate the distance between the support samples and query samples.”
 - We removed the extra reference for ADAM optimized
- In Section 3.1:
 - We updated the title to be “Employed datasets” instead of “Evaluation datasets”
 - We emphasized the annotation of the datasets using the word “labeled”
 - We detailed the dataset description for “Fattouma Bourguiba Monastir (FBM), Tunisia” in table 1 (dataset statistics).
- In Section 3.2 (Experimental settings):
 - In parameter settings subsection, we included implementation detail about dropout rate: “We employed dropout with 20% rate”
 - In Hyperparameter settings subsection, we renamed “experiment D and E” with “test 1” and “test 2” for more clarity.
 - In Table 2 (experiments results), we updated the metrics for the experiment asked by the reviewers to avoid data leakage and to show the class specific metrics.
- In Section 3.3 (DemyeliNeXt evaluation):

- We replaced the experiment names to “C and B” instead of “C” in this paragraph: “In contrast, Experiment C and B, which utilized one, and three shots and queries, respectively, demonstrated the lowest performance”.
- We updated the figure and its caption to include the visualization of feature importance for a NON-MS example with an MS example. “Deep SHAP Explanation of MS and NON-MS Examples. A: Explanation of NON-MS example. B: Explanation of MS example. For each of the subfigures (A and B)”
- We emphasized testing the “backbone on unseen MS and NON-MS examples” to show the reviewer that our model is able to generalize well to unseen distribution.
- We added the used features for diagnosis: "We evaluated the explainer results using the key diagnostic features outlined in the McDonald criteria [13], which include lesion size, number of lesions, lesion location, lesion contrast, and lesion shape".
- We updated Deep SHAP discussion to show that it has detected lesions that are responsible for MS and NON-MS diseases: “The Deep SHAP explainer seems to identify some of the key features for classification , specially the lesions in MS example (Fig.2 B).”
- We added a warning about the explanation features “However, one should note that there is a risk that the included features in the explanation could be deemed irrelevant to clinicians.”
- In Limitations and future studies subsection, we add this sentence to address the reviewer request about more quantitative and qualitative benchmarks: “In future studies, we aim to benchmark against state-of-the-art methods.”
- We addressed the reviewer concern of evaluation of the explanation in future works: “... as well as developing more clinically relevant explainability methods with their evaluation.”
- Throughout this paper:
 - We replaced the term “task” with “episode” for harmonization.
 - We disclosed the hospital names “Fattouma Bourguiba Monastir (FBM), Tunisia” and “Sahloul Sousse (SS), Tunisia).
- We disclosed the code source link
- We added the disclosure of interest and acknowledgment