Responsible NLP Checklist

Paper title: *DIVINE : Coordinating Multimodal Disentangled Representations for Oro-Facial Neurological Disorder Assessment*
Authors: *Mohd Mujtaba Akhtar, Girish, Muskaan Singh*

> How to read the checklist symbols:
>
> ☑ the authors responded 'yes'
>
> ☒ the authors responded 'no'
>
> N/A the authors indicated that the question does not apply to their work
>
> ☐ the authors did not respond to the checkbox question
>
> For background on the checklist and guidance provided to the authors, see the Responsible NLP Checklist page at ACL Rolling Review.

☑ **A. Questions mandatory for all submissions.**

☑ A1. Did you describe the limitations of your work?
This paper has a Limitations section.

☑ A2. Did you discuss any potential risks of your work?
Potential risks are discussed in the Societal Impact section, where we address possible misuse, bias, and risks of deploying the system without clinician oversight.

☑ **B. Did you use or create scientific artifacts? (e.g. code, datasets, models)**

☑ B1. Did you cite the creators of artifacts you used?
Creators of all used artifacts are cited in Section 3 (Representations) and Appendix A.1, with references to the original papers and model repositories.

N/A B2. Did you discuss the license or terms for use and/or distribution of any artifacts?
The paper uses pretrained models and datasets strictly under their original research-use terms. Licensing details are governed by the original providers and are not restated in the paper.

☑ B3. Did you discuss if your use of existing artifact(s) was consistent with their intended use, provided that it was specified? For the artifacts you create, do you specify intended use and whether that is compatible with the original access conditions (in particular, derivatives of data accessed for research purposes should not be used outside of research contexts)?
Consistency with intended use is discussed in Section 3 and the Ethical Statement, where all pretrained models and datasets are used strictly for non-commercial research and evaluation purposes.

☑ B4. Did you discuss the steps taken to check whether the data that was collected/used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect/anonymize it?
Data anonymization, privacy protection, and the absence of personally identifying information are discussed in the Ethical Statement section.

☑ B5. Did you provide documentation of the artifacts, e.g., coverage of domains, languages, and linguistic phenomena, demographic groups represented, etc.?
Dataset characteristics, task design, and demographic information are documented in Section 5 (Experiments) and Appendix A.2.

☑ B6. Did you report relevant statistics like the number of examples, details of train/test/dev splits, etc. for the data that you used/created?
Dataset size, subject counts, task descriptions, and cross-validation protocol are reported in Section 5 and Appendix A.2.

☑ **C. Did you run computational experiments?**

☑ C1. Did you report the number of parameters in the models used, the total computational budget (e.g., GPU hours), and computing infrastructure used?
Model size, parameter counts, training setup, and computing infrastructure are reported in Appendix A.4.

☑ C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?
Experimental setup and fixed hyperparameter values are described in Section 5 (Training Details) and Appendix A.4.

☑ C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?
Results are reported as averages over five-fold cross-validation and multiple random seeds, with statistical significance testing described in Section 5.

☑ C4. If you used existing packages (e.g., for preprocessing, for normalization, or for evaluation, such as NLTK, SpaCy, ROUGE, etc.), did you report the implementation, model, and parameter settings used?
Preprocessing tools and packages (e.g., OpenFace, librosa) and their usage are described in Appendix A.3.

☑ **D. Did you use human annotators (e.g., crowdworkers) or research with human subjects?**

N/A D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?
The study uses an existing clinical dataset; participant instructions were part of the original data collection protocol and are outside the scope of this work.

N/A D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?
Participant recruitment and compensation were handled by the original dataset creators and are documented in the original dataset publication.

☑ D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating (e.g., did your instructions explain how the data would be used)?
Consent procedures are described in the Ethical Statement and referenced via the original dataset documentation.

☑ D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?
Ethics approval and institutional oversight are described in the Ethical Statement section.

☑ D5. Did you report the basic demographic and geographic characteristics of the annotator population that is the source of the data?
Basic participant demographics and clinical group information are reported in Appendix A.2.

☑ **E. Did you use AI assistants (e.g., ChatGPT, Copilot) in your research, coding, or writing?**

☒N/A E1. If you used AI assistants, did you include information about their use?
AI assistants were used only for language polishing.