Comments from the editors and reviewers:

-**Reviewer 1**

Recommendation: Accept

The authors have proposed an approach using Capsule Networks for Sign Language Recognition using wearable IMU's. They have done the comparison of the performance of capsule Networks with CNN using game theory. I feel that the following is missing in the paper. Paper, in general, is well written. But I feel that a revised version with the following details will improve the paper.

1. Comparison of results with other existing approaches are missing.

2. The architecture of the CNN used for comparison with the capsule network is not given.

-**Reviewer 2**

Recommendation: Accept

- Authors are requested to address the following comments and suggestions when submitting the camera-ready version:

- Conclusion section is to be expanded little more.

- Use more recent and relevant references such as: Jafar A. Alzubi, 2015. Optimal Classifier Ensemble Design Based on Cooperative Game Theory. Research Journal of Applied Sciences, Engineering and Technology, 11(12): 1336-1343.

- Adding more explanation on ALGORITHM 1

- Figures 3, 4, 5 and 8 are not clear so they need to be re-simulated with higher resolution.

- Adding more explanation on figure 2.

-**Reviewer 3**

Recommendation: Weak Accept

1. How does game theory help in obtaining the target objective in proposed work?

2. In the proposed game theory model, what kind of behavioral relations have been used?

3. How is the computational complexity of the proposed joint model observed to be?

4. The proposed theorem are supportive but could be made more correlated to the proposed work to improve the quality further.

5. How many samples have been taken? Elaborate regarding the heterogenous nature of the data set used? does it incorporate varying style and measurement orientations?

6. If possible some highlighting comments could be provided in figures 6 and 7 to improve readability

7. Elaborate about Nash equilibrium

8. Some recent literature could be used.

RESPONSE FROM AUTHORS

The authors would like to thank the journal for a timely response and the reviewers who took the time to go through our work and present us with valuable feedback. The authors acknowledge the comments provided by the reviewers and have modified the work as mentioned below.

Reviewer-1

The authors have proposed an approach using Capsule Networks for Sign Language Recognition using wearable IMU's. They have done the comparison of the performance of capsule Networks with CNN using game theory. I feel that the following is missing in the paper. Paper, in general, is well written. But I feel that a revised version with the following details will improve the paper.

1. Comparison of results with other existing approaches are missing.
Response- Performance of the proposed Capsule Network (CapsNet) architecture is compared with its immediate predecessor, the Convolutional Neural Network (CNN) having similar trainable hyper-parameters. Since CapsNet was proposed to overcome the problems of CNN, we limited our comparison of the proposed CapsNet architecture with CNN [19]. The architecture of CNN used in this work has been included in more detail.

[19] Sara Sabour, Nicholas Frosst, Geoffrey E. Hinton, "Dynamic Routing Between Capsules", 31st Conference on Advances in Neural Information Processing Systems (NIPS 2017), pp. 1-11.

2. The architecture of the CNN used for comparison with the capsule network is not given.
Response- Architecture of the CNN used for comparison along with its operation has been included in the paper (pg6, para over the heading "NON-COOPERATIVE CAPSNET GAMES" and same is depicted in Fig 2b). Furthermore, dimensional transformations and hyperparameters of the model have also been included here.

Reviewer-2

- Authors are requested to address the following comments and suggestions when submitting the camera-ready version:

- Conclusion section is to be expanded little more.
Response- The conclusion section has been expanded. The proposed approach and its advantages and shortcomings in comparison to the conventional architecture have been

thoroughly discussed. The authors have also elaborated concluding remarks on the applications and implementations of the proposed CapsNet methodology.

- Use more recent and relevant references.

Response- More recent and relevant literature has been added as suggested by the reviewer (references [2], [5], [6], [7], [9], [11], [23], [24], [29], [30]).

- Adding more explanation on ALGORITHM 1

Response- Algorithm 1 describes the non-cooperative pick game constructed in order to make the two models (CNN and CapsNet) compete with each other. An elaborate explanation of the operation and description of the process has been included in the paragraph after Algorithm-1. Here, authors have also extended upon notations, terms and functions used in the algorithm along with their relevance to the main objective.

- Figures 3, 4, 5 and 8 are not clear so they need to be re-simulated with higher resolution.

Response- Figures 3, 4, 5 and 8 (graph plots) have been re-simulated and plotted with higher resolution. Labels in the figures have also been improved for readability.

- Adding more explanation on figure 2.

Response- Fig 2 (Fig 2a of the revised version of the manuscript) represents the architecture of the proposed Caspule Network model. Explanation of the figure has been extended to include more details on its functionality, relevance, parameters used and optimization in the paragraphs after Fig. 2. Theory on the operation of Capsule layers along with the incorporation of Dynamic Routing has also been included in the explanation. A depiction of the CNN model used for comparison has been added in Fig 2b.

Reviewer-3

1. How does game theory help in obtaining the target objective in proposed work?

Response- Game Theory acts as a real-time validation methodology in order to present the performance of both the models used without any external bias. By making use of Game Theory, optimized performance of the proposed CapsNet architecture over the conventional CNN is observed in real-time. Authors have added the usage of Game Theory and its role in achieving the target objective under the headings “Introduction” (paragraph no. 5) and “Non-Cooperative CapsNet Games” (paragraph no.1).

2. In the proposed game theory model, what kind of behavioral relations have been used?

Response- Strategic behavioural relations (behaviour of the model depending upon the strategy adopted, i.e., the trained weights) are used for the Game Theory model. Optimization of the model is carried out during the training phase, which now becomes the strategy of the player. Here, the players are CNN, CapsNet with 3-routings and CapsNet with 5-routings. Games are played on the test data. Finally, Nash equilibrium is used to compare the performance of the three players and declare a winner. The corresponding explanation is expanded under the heading “NON-COOPERATIVE CAPSNET GAMES” in the paragraph before Algorithm-1 as well as under the heading “Definition-1 (Nash Equilibrium)” in the RESULTS section.

3. How is the computational complexity of the proposed joint model observed to be?

Response- Computational complexity of the proposed CapsNet is observed to be higher than that of the conventional CNN approach during the training phase. However, in real-time, CapsNet predicts values faster and more accurately than CNN as a result of increased weight optimization. Remarks on computational complexity for both the models during training and validation phases have been included in the paper as suggested by the reviewer. (in the section on "Results and Discussion", sub-section (b))

4. The proposed theorem are supportive but could be made more correlated to the proposed work to improve the quality further.

Response- The proposed theorems form the basis for the construction of the non-cooperative pick games. In the paper, first, the theorems are mentioned to forma a general framework of the proposed pick game, which is applied subsequently for validation of the considered models. The explanation on the application of these theorems are present in paragraphs after Algorithm 1 and in the Results section, in the two paragraphs after Fig. 8.
The theorems have been kept general so that they may be applicable in other scenario where two or more models are to be validated using pick games.

5. How many samples have been taken? Elaborate regarding the heterogenous nature of the data set used? does it incorporate varying style and measurement orientations?

Response- A total of 20 sentences were recorded with 10 repetitions of each sentence. 10 subjects were used to record the data which provides the total number of samples as 20x10x10= 2000. Explanation of the number of samples and its training and test split has been added in the revised version of the manuscript (under the heading "Signed Sentence Recognition using CapsNet", subsection (a) and in the paragraph above the heading "NON-COOPERATIVE CAPSNET GAMES"). Here, the heterogeneous nature of the dataset constructed has also been included in terms of the subject variability. Variation in gender, age and dominant hand of the subjects have been added. The dataset makes use of multiple sensors as a result of which different signals used possess different styles and orientations. For example, orientation signals for 'Bank' and 'Where' in the sentence 'Where is the Bank?' are in stark contrast with each other (as depicted in the Demo-2 video, uploaded with the manuscript. These justifications have been included in the revised version of the manuscript.

6. If possible some highlighting comments could be provided in figures 6 and 7 to improve readability

Response- Sharp spatial activations with a higher intensity of the colour gradient in Fig.7 indicate better optimized weights of CapsNet in comparison to CNN (Fig. 6). Highlighted remarks for figures 6 and 7 have been incorporated in order to improve readability, as indicated by the reviewer in the explanation included after Fig. 7.

7. Elaborate about Nash equilibrium

Response- The idea of Nash Equilibrium corresponds to the point where both the players in the constructed games have adopted and optimized the single adopted strategy. It denotes that point in the game where both the players are at their best with respect to each other's performance. Nash Equilibrium is an essential indicator in determining the extent of optimization between the strategies. Elaborate explanation on Nash Equilibrium and its role as a significant measure in Game Theory has been added (following Definition-1).

8. Some recent literature could be used.

Response- More recent and relevant literature has been added for reference as suggested by the reviewer. ([2], [5], [6], [7], [9], [11], [23], [24], [29], [30])