CHAPTER X

RESEARCH FINDINGS

The primary goal of this study is to improve the accuracy of HAR from video sequences by creating a strong deep-learning model. The primary goal of this work is to generate video descriptors for human activity recognition by learning features such as body joints, trajectory points, geometric elements, and spatiotemporal data. In order to improve the accuracy of human activity recognition, five models are proposed here to learn various aspects associated with human activities. Experiments are conducted with the Penn Action dataset. The following conclusions can be drawn from the experimental data:

- By learning both body joints and trajectories in video frames from the Penn Action dataset, the JTDD model achieved 98.7% recognition accuracy, compared to the JDD model.
- Based on the learning of relevant spatial dissimilarities among various video frames belonging to different classes from the Penn Action dataset, the JTPADBRD model attained 99.4% recognition accuracy, compared to the JTDD model.
- By learning long-range body joint correlations across various classes of video sequences in the Penn Action dataset, the JTDPAHBRD model achieves 99.6% recognition accuracy, which is an improvement over the JTPADBRD.
- The JTDGPAHBRD model, by recording different geometries together with the body joints and trajectories in several video frames, outperformed the JTDPAHBRD model on the Penn Action dataset, with a 99.7% recognition accuracy.
- Finally, the JTDGPAHBRD-GCN reaches the maximum recognition accuracy of 99.82% compared to the other models on the Penn Action dataset, because of learning body joints, trajectory points, geometrics, as well as spatiotemporal dependencies from long-range video sequences for video descriptor generation.
- When comparing models, the JTDGPAHBRD-GCN outperforms the JTDD, JTDPABRD, JTDPAHBRD, and JTDGPAHBRD models on the Penn Action

dataset by a margin of 1.13%, 0.42%, 0.22%, and 0.12, respectively, when concatenating features from the conv5b and conv4b layers.

- Thus, it is realized that the JTDGPAHBRD-GCN model obtains the maximum recognition accuracy, whereas the JTDD produces the minimum results.

10.1 CONCLUSION

This research work address issues and problems related to inconsistencies in the literature survey. In several contexts, Human Activity Recognition (HAR) has risen to prominence. The model developed must satisfy the objectives specified in the research work proposed and function well in sports applications.

- The existing deep learning models are analyzed, and the JTDD model is proposed for learning the trajectory points of human activities along with the body joints efficiently.
- In order to learn the important spatiotemporal connections between body joints and trajectory points from the various classes of human activity video sequences, the JTDPABRD model was developed.
- The JTDPAHBRD model is developed for enhancing the feature pooling strategy in the C3D network.
- Different geometries, body joints, and trajectory points can be extracted from various video sequence types using the JTDGPAHBRD model.
- The JTDGPAHBRD-GCN model is developed for learning the spatiotemporal dependencies among various geometrics from the human activities and improving the feature learning ability for video descriptor generation in the HAR.
- Using the Penn Action dataset, the accuracy, precision, recall, and f-measure of five different recognition models were evaluate.
- The JTDGPAHBRD-GCN model showed an accuracy of 99.82%, precision of 0.995, recall of 0.998 and f-measure of 0.997, compared with the existing models.

10.2 FUTURE WORK

This proposed research can involve the following future enhancements:

1. The work be extended to integrate metaheuristic optimization algorithms to optimize the proposed models according to various datasets.
2. Data fusion algorithms can be explored to effectively integrated information from diverse modalities like face, palmprint, eye movements, etc.
3. Third, the scope can be broadened to incorporate the creation of transfer learning mechanisms that can transport expertise from a domain with an abundance of labeled data to one with fewer such labels. This can make models more robust to different environments and user variations.
4. HAR scenarios often involve continuous data streams. In the future, online learning can be introduced that can adapt the model to changing activity patterns over time without the need to retrain from scratch.