22 April 2025 ·

Qeios

Peer Review

Review of: "Synthetic Data Generation of Body Motion Data by Neural Gas Network for Emotion Recognition"

Wellington P. dos Santos[1]

1. Universidade Federal de Pernambuco, Brazil

General Overview

The paper introduces a novel application of the Neural Gas Network (NGN) to synthesize body motion data for emotion recognition, addressing the critical issue of dataset scarcity in this domain. The authors present a comprehensive evaluation framework, comparing NGN with state-of-the-art generative models such as GANs, VAEs, diffusion models, and LSTMs. The results demonstrate that NGN outperforms these methods in terms of accuracy, diversity, fidelity, and computational efficiency. Overall, the paper is well-structured, scientifically sound, and addresses a significant gap in the field of emotion recognition using body motion.

However, while the paper has many strengths, there are areas where improvements could enhance its impact and clarity. Below, I provide a detailed critique organized into several key aspects.

1. Strengths

1.1 Novelty and Contribution

The use of NGN for synthesizing body motion data is innovative and fills a gap in the literature. The paper convincingly argues that NGN offers advantages over existing methods, particularly in terms of computational efficiency and diversity of generated data. The supervised adaptation of NGN to learn skeletal topology and generate realistic synthetic motions is a strong contribution.

1.2 Comprehensive Evaluation

doi.org/10.32388/1OZLEN

The evaluation framework is robust, employing a wide range of metrics (e.g., FID, Diversity, Fidelity, DTW, MPJPE, classification metrics like accuracy, precision, recall, and MCC). This multifaceted approach provides a holistic view of the model's performance. Additionally, the inclusion of kinematic feature analysis and feature importance plots adds depth to the discussion.

1.3 Practical Implications

The paper emphasizes the real-world applicability of NGN-generated data, particularly in enhancing real-time emotion recognition systems. The discussion on scalability and ethical considerations further strengthens the practical relevance of the work.

1.4 Clarity of Methodology

The explanation of the NGN algorithm, including its mathematical formulation and implementation details, is clear and accessible. The step-by-step description of how gas particles are used to represent body joints and generate new motions is commendable.

2. Areas for Improvement

2.1 Dataset Limitations

While the "100Style" dataset is an excellent choice, it has limitations that should be acknowledged more explicitly:

- Single Actor Bias: The dataset contains motion capture data from only one actor (a 182 cm tall individual). This raises concerns about generalizability across different body types, genders, and ages.
- Emotional Categories: The paper focuses on four emotional states (angry, depressed, neutral, proud), which may not fully capture the complexity of human emotions. Including more nuanced or subtle emotions (e.g., frustration, surprise, or mixed emotions) would strengthen the study.

Suggestion: Discuss the potential biases introduced by the single-actor dataset and propose future work to validate the method on more diverse datasets.

2.2 Comparison with Baseline Models

While the comparison with other generative models is thorough, some aspects could be improved:

- Baseline Model Description: The baseline model is briefly mentioned but not described in detail.

Providing more information about its architecture and training process would allow readers to better understand the comparative advantage of NGN.

- Qualitative Analysis: While quantitative metrics are well-presented, qualitative analysis (e.g., visualizations of generated motions for each emotion) is limited. Adding more visual comparisons could help highlight the differences between NGN and other models.

Suggestion: Include additional qualitative evaluations, such as side-by-side videos or animations of generated motions for each emotion, to complement the numerical results.

2.3 Ethical Considerations

The paper mentions the ethical implications of synthetic data generation but does not delve deeply into potential risks:

- Bias Amplification: If the training data contains biases (e.g., cultural or gender-specific movement patterns), the NGN might amplify these biases in the synthetic data.
- Privacy Concerns: While the dataset used is anonymized, the broader implications of generating synthetic human motion data (e.g., misuse in surveillance or deepfake applications) warrant more discussion.

Suggestion: Expand the ethical discussion to address these concerns and propose safeguards (e.g., bias mitigation techniques, usage guidelines).

2.4 Computational Efficiency

One of the claimed advantages of NGN is its computational efficiency. However:

- Hardware Dependency: The runtime results are reported for a specific hardware configuration (Core i7 CPU, 12th generation). Including benchmarks for other configurations (e.g., GPUs) would make the results more generalizable.
- Scalability: The paper briefly mentions scalability but does not test the method on larger datasets or more complex emotional states. Demonstrating scalability would strengthen the argument for NGN's practical utility.

Suggestion: Provide runtime comparisons across different hardware setups and evaluate the method on a larger, more diverse dataset to demonstrate scalability.

2.5 Feature Importance Analysis

The feature importance analysis is insightful but could benefit from additional context:

- Cross-Emotion Comparisons: While velocity is identified as important for anger and acceleration for depression, it would be useful to see how these features interact across different emotions. For example, how do velocity and acceleration jointly contribute to distinguishing pride from anger?
- Statistical Significance: It is unclear whether the differences in feature importance across emotions are statistically significant. Including p-values or confidence intervals would add rigor to the analysis.

Suggestion: Perform statistical tests to confirm the significance of feature importance differences and explore cross-emotion interactions.

3. Writing and Presentation

3.1 Clarity

The writing is generally clear, but some sections could be streamlined:

- Repetitive Phrasing: Certain phrases (e.g., "synthetic and baseline data") are repeated frequently. Using synonyms or rephrasing could improve readability.
- Figures and Tables: While the figures and tables are informative, some (e.g., Figure 9) could benefit from better labeling and annotations to make them more self-explanatory.

Suggestion: Revise repetitive sections for conciseness and improve figure/table annotations for clarity.

3.2 Terminology

Some technical terms (e.g., "Neural Gas Network," "Fréchet Inception Distance") are introduced without sufficient background for readers unfamiliar with these concepts. Providing brief definitions or references would enhance accessibility.

Suggestion: Include a glossary or brief explanations of key terms in the introduction.

4. Future Work

The paper outlines promising directions for future research, such as exploring scalability, optimizing computational efficiency, and addressing ethical concerns. However, additional suggestions could

doi.org/10.32388/1OZLEN

include:

- Multimodal Integration: Investigate how NGN-generated body motion data can be integrated with other modalities (e.g., facial expressions, vocal tone) for more comprehensive emotion recognition.
- Real-Time Applications: Develop a prototype real-time emotion recognition system using NGN-generated data to demonstrate its practical feasibility.

Conclusion

This paper makes a significant contribution to the field of emotion recognition through its innovative application of the Neural Gas Network for synthetic body motion generation. The methodology is rigorous, the results are compelling, and the practical implications are well-articulated. However, addressing the identified limitations—particularly regarding dataset diversity, ethical considerations, and qualitative analysis—would further enhance the paper's impact.

Overall Recommendation: Accept with minor revisions. The paper is well-suited for publication in its current form after addressing the suggested improvements.

Declarations

Potential competing interests: No potential competing interests to declare.