Review of: "Songs Classification Problem Research by Genre Based on Neural Network"

Javier Martinez-Moguerza[1]

1 Universidad Rey Juan Carlos

Potential competing interests: No potential competing interests to declare.

In this paper, the use of neural networks for classifying musical compositions by genre based on their audio characteristics is explored.

I think there is a lot of room for improvement in this work. My suggestions to improve the quality of this paper are reflected in the following comments:

- Regarding the title, I suggest "NEURAL NETWORK-BASED RESEARCH ON THE PROBLEM OF SONG CLASSIFICATION BY MUSICAL GENRE. At least, I think it is important to include the term "musical genre" in the title rather than just the term "genre."

- Page 1: An interesting reference for the introduction, treating this and other problems, is:

Meinard Müller (2007). Information Retrieval for Music and Motion. Springer.

- Page 1: The sentence "Traditional classification methods may be ineffective for such diverse content" should be supported with some references.

- Page 2: The whole section "Formulation of the problem" may fit better after the section "Formulation of the article goal."

Page 2: If the previous suggestion is considered by the authors, the first time Mel-Frequency Cepstral Coefficients (MFCC) appear on page 2 in the sentence "Most research focuses on using audio features such as MFCC," a reference on MFCC must be included. A classical one may be:

C. G. M. Fant (1959), "Acoustic description and classification of phonetic units," Ericsson Technics, vol. 1, 1959.

- Page 3: The procedure to initialize the neural network's weights and biases should be described.

- Page 4: I suggest including a numerical results section with figures 1, 2, and 3 in this section. This section should also include a comparison of the current proposal with other approaches in the literature.

- Page 4: The sentence "For testing and training the network, there was prepared data representing songs with genres of rock, pop, and jazz that were used." should be better written as "To test and train the network, data representing songs with rock, pop, and jazz genres were prepared and used."

- Page 4: The database used should be described in detail: total number of songs, number of songs by genre, etc., and made available somewhere to guarantee reproducibility of the results by other researchers.

- Page 4: Figure 1 is not mentioned in the text of the article. It should be mentioned in the sentence "A diagram based on tempo and loudness metrics was constructed to visualize

the results."

- Page 5: In my opinion, figures 2 and 3 are not enough to show the numerical results. I miss some tables showing explicit numerical results, including measures such as accuracy, precision, and recall for the whole dataset and for each musical genre. In addition, as mentioned above, it is advisable to compare with other approaches.

- Page 5: In the conclusions, I suggest including threats to the validity of the research and some recommendations for practitioners.