

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

Hugo S. Oliveira<sup>1</sup>

<sup>1</sup> Universidade do Porto

**Potential competing interests:** No potential competing interests to declare.

Although the work follows a common structure regarding sound classification, there is nothing said about the neural network architecture used to classify the sounds.

Also, nowadays the common standard is to exploit pre-trained networks using the same feature extractor, MFCC, and fine-tune it into the target objective.

For example, AST - Audio Spectrogram Transformer, transformers-based architectures pre-trained on AudioSet.

VIT-based architectures operate on image spectrograms obtained from MFCC, and others such as Inception, ResNet, EfficientNet.

The work lacks the usage of these network architectures, and even the one used is not described.

I suggest making these experiments and updating the document with those more broader experiments.