

We propose a CNN-based approach to classify ten genres of ballroom dances given audio recordings, five latin and five standard, namely Cha Cha Cha, Jive, Paso Doble, Rumba, Samba, Quickstep, Slow Foxtrot, Slow Waltz, Tango and Viennese Waltz. We utilize a spectrogram of an audio signal and we treat it as an image that is an input of the CNN. The classification is performed independently by 5-seconds spectrogram segments in sliding window fashion and the results are then aggregated. The method was tested on following datasets: Publicly available Extended Ballroom dataset collected by Marchand and Peeters, 2016 and two YouTube datasets collected by us, one in studio quality and the other, more challenging, recorded on mobile phones. The method achieved accuracy 93.9%, 96.7% and 89.8% respectively. The method runs in real-time. We implemented a web application to demonstrate the proposed method.