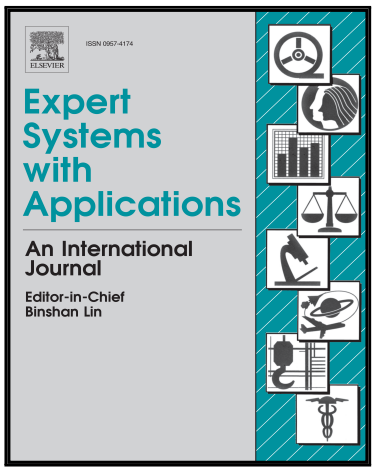

Speech-Music Discrimination Using Deep Visual Feature Extractors

Michalis Papakostas, Theodoros Giannakopoulos

PII: S0957-4174(18)30299-9
DOI: 10.1016/j.eswa.2018.05.016
Reference: ESWA 11969

To appear in: *Expert Systems With Applications*

Received date: 3 January 2018
Revised date: 11 May 2018
Accepted date: 12 May 2018

Please cite this article as: Michalis Papakostas, Theodoros Giannakopoulos, Speech-Music Discrimination Using Deep Visual Feature Extractors, *Expert Systems With Applications* (2018), doi: 10.1016/j.eswa.2018.05.016

Highlights

- CNNs provide state-of-the-art results in Speech-Music Discrimination
- Transfer Learning can be used to transfer knowledge across different modalities
- CNNs & Transfer-Learning can significantly reduce the training data requirements
- Data Augmentation and Representation play important role when training Audio-CNNs