Emotion Recognition in Music using Artificial Intelligence

1. **A Regression Approach to Music Emotion Recognition**

Acest articol este relevant, intrucat prezinta o modalitate nu atat de des intalnita pentru a clasifica muzica la o anumita emotie folosind regresia , contrar altor tipuri de algoritmi care recurg la retele neuronale. Totadata, descrie multiple moduri de extragere a caracteristicilor dintr-un sample.

Lucrarea are un format oarecare clasic prezentand urmatoarele capitole:

* Introducere
* Related Works (sunt prezentate alte 3 abordari ale acestei teme cu descrieri bine structurate)
  + Arousal and Valence Modeling
  + Fuzzy Approach
  + System Identification Approach (System ID)
* Regression Approach (sunt prezentate notiunile teoretice, matematice ale acestei abordari)
* System Description (sunt prezentati pasii prin care sample-ul este trecut prin system pentru a ajunge la faza de clasificare)
  + Data Collection and Preprocessing
  + Feature Extraction
  + Subjective Test
  + Regressor Training
  + Emotion Visualization
* Performance Study (sunt prezentate mai multe metrici de performanta, fiind folosite diferite regresoare)
  + Consistency Evaluation of the Ground Truth
  + Data Space
  + Feature Space
  + Performance Evaluation of Regressor
  + Performance Evaluation for MEVD
* Discussion on the subjectivity issue
* Conclusion

Referintele in acest articol sunt in numar de 35. Numele autorilor apare primul, urmat de titlul lucrarii, volumul, numerele paginilor, luna si anul.

1. **SMERS: MUSIC EMOTION RECOGNITION USING SUPPORT VECTOR REGRESSION**

Acest articol este relevant, intrucat prezinta o oarecare abordare similara cu cea anterioara, insa sunt folosite alte caracteristici si metode de a incerca diminuarea datelor clasificate incorrect.

Lucrarea cumprinde urmatoarele capitole:

* Introduction
* Related Work (sunt prezentate 2 abordari, insotite de citatii si descrieri scurte, dar bine structurate)
  + Music and Emotion
  + Music Emotion Recognition
* Implementation (prezinta date legate de descrierea sistemului, setul de date folosit si caracteristicile din setul de date si o descriere referitoare la modul cum se obtin aceste caracteristici)
  + System Description
  + Dataset
  + Musical Features
    - Scale
    - Average Energy
    - Rhythm
    - Harmonics
* Emotion Recognition (prezinta procesul de antrenare, drumul parcurs pana la alegerea unei abordari corecte care poate fi folosita pentru optimizarea modelului, astfel incat sa diminueze numarul de clasificari incoerente)
  + Training Process
  + Classification Methods
    - Support Vector Regression (SVR)-based Training
    - Support Vector Machine (SVM)-based Training
    - Gaussian Mixture Model (GMM)-based Training
* Experiments and results (prezinta experimente si rezultate folosind diferite tipuri de clasificatori si moduri de reprezentare a sistemului de coordinate ales pentru a mapa emotia melodiei)
  + Confusion Matrix
  + Accuracy
* Conclusions and future work
* References

Referintele sunt in numar de 17, numele autorilor apare primul, urmat de titlul lucrarii, volum, numerele paginilor si anul publicarii.

1. **Novel audio features for music emotion recognition**

Acest articol este relevant, intrucat prezinta in detaliu obtinerea caracteristicilor unui set de date cu fisiere audio, in special melodii.

Lucrarea cuprinde urmatoarele capitole:

* Introduction
* Related work (prezinta faptul ca majoritatea caracteristicilor deja existente limiteaza calitatea sistemelor de clasificare)
* Methods (prezinta pasii prin care au trecut pentru procurarea unui nou set de date cu feature-uri multiple cu descrieri foarte detaliate si algoritmi multiplii pentru detectarea acestora, totodata in final este ilustrat si rezultatul cu o acuratete mai mare decat setul de date cu feature-uri baseline)
  + Dataset Acquisition
  + Validation of Emotion Annotations
  + Standard Audio Features
  + Novel Audio Features
    - From the audio signal to MIDI notes
    - Melodic features
    - Dynamic features
    - Rhythmic features
    - Musical texture features
    - Expressivity features
    - Voice Analysis Toolbox (VAT) features
  + Emotion Recognition
* Results and discussion (sunt studiate in detaliu rezultatele si caracteristicile)
  + Classification Results
* Conclusions and future work

Referintele sunt in numar de 54. Numele autorilor apare primul, urmat de titlul lucrarii, volum, numerele paginilor si an.

1. **Musical instrument emotion recognition using deep recurrent neural network**

Acest articol este relevant, intrucat prezinta o abordare diferita fata de celalalte articole, si anume, o metoda de deep learning folosind RNNs.

Lucrarea cuprinde urmatoarele capitole:

* Introduction
* Previous works (sunt prezentate in principiu abordarile cu SVM, algoritmi genetici, dar si CNNs)
* Dataset (prezinta o scurta descriere a setului de date)
* Acoustic features (prezinta clasele caracteristicilor folosite, alaturi de notiunile teoretice descriptive acestora)
  + MFCC
  + Chroma Energy Normalized Statistics (CENS)
  + Chroma STFT
  + Spectral Centroid
  + Spectral Bandwidth
  + Spectral Rolloff
* System structure (prezinta structura si arhitectura sistemului, respective modelului; sunt evidentiate layerele folosite pentru RNN cu 4 layere ascunse, dintre care: 2 sunt bidirectionale LSTM cu 64 noduri si 2 noduri cu 32, respective 16 noduri; se folosesc 2 functii de activare, ReLU si softmax; totodata se prezenta si performanta unui SVM)
  + Instrument emotion recognition with RNN
  + Instrument emotion recognition using SVM
* Results and discussion (prezinta performantele folosirii a RNNs relative la SVMs cu diferite clase de feature-uri precum MFCC)
* Conclusion

Referintele sunt in numar de 24. Numele autorilor apare primul, urmat de titlul lucrarii, anul, volumul, numerele paginilor.

1. **Musical instrument emotion recognition using deep recurrent neural network**

Acest articol este relevant, intrucat prezinta tot o abordare de deep learning si anume, CLDNNs. Acuratetea incredibil de buna de 99.19% este de asemenea un factor ce confera relevanta.

Lucrarea cuprinde urmatoarele capitole:

* Introduction
* New emotional Turkish music database (presents a description regarding the making of a new dataset with 124 30s audio clips representing Turkish traditional songs)
* CLDNN architecture for music emotion recognition (prezinta arhitectura retelei neuronale, vasta, cu 200 de layere ascunse si o combinatie de LSTM+DNN classifiers; totodata descrie cum se extrag feature-urile din fiecare clip audio, feature-uri care sunt pasate CNN-unul)
  + One-dimensional CNN for feature extraction
  + Standard audio features
  + Feature selection
* Experimental setup (prezinta cum au folosit 2 clase de feature-uri: MFCC si log-mel filterbank energies; totodata au adoptat o metoda de 10-fold cross-validation pentru a masura performanta modelului prin acuratete, recall, precizie si F1 score)
* Results and discussions (prezinta diferite rezultate de performanta a modelului)
* Conclusion

Referintele sunt in numar de 52. Numele autorilor apare primul, urmat de titlul lucrarii, anul, numerele paginilor, linkuri catre articole.

Lab3

Cuprins:

1. Introduction
2. Related works
3. Feature extraction
4. Datasets
5. Experiments
6. Proposed method
7. Conclusion

Plan:

Ipoteza de lucru: Fine-tuning-ul unor modele neuronale ar putea imbunatati acuratetea detectiei emotiilor in melodii.

Metodologia:

1. Colectare datelor: Baza de date cu melodii etichetate cu emotii, cu feature-uri similare celor de la Spotify API
2. Preprocesarea datelor: Vom folosi un API de la spotify care extrage caracteristici din melodii, dar si o librarie librosa, in general utilizata pentru a analiza semnale audio muzicale.
3. Selectarea modelelor: Vom alege modele de tip LSTM, RNN, intrucat au prezentat o performanta mult mai buna in detectarea emotiilor. Vom folosi SVMs si abordari de regresie pentru a compara performantele.
4. Fine-tuning
5. Setul de date: Vom imparti datele existente in 80-10-10, antrenament, testare, validare.
6. Metrici de evaluare: Acuratete

Abordare originala: Incercarea fine-tuning-ului in speranta obtinerii unei etichetari cat mai apropiata de cea reala

Descrierea experimentelor:

* Parametrizarea, antrenarea fiecarui model in parte (ANNs vs ANNs vs SVMs vs Regression)
* Evaluarea performantei pe acelasi set de date de test
* Analiza datelor: compararea acuratetii fiecarui model asupra setului de date de test
* Limitari

Posibila contributie originala:

* Poate fine-tuning-ul sa obtina o acuratete mai buna a rezultatelor? In cazul in care raspunsul este da, acest lucru ar scoate in evidenta valori esentiale pentru antrenarea modelelor specializate in lucrul cu spectre audio.

**References**

1. Y. -H. Yang, Y. -C. Lin, Y. -F. Su and H. H. Chen, "A Regression Approach to Music Emotion Recognition," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 16, no. 2, pp. 448-457, Feb. 2008, doi: 10.1109/TASL.2007.911513.
2. B. Han, S. Rho, R. B. Dannenberg, and E. Hwang, “SMERS: Music emotion recognition using support vector regression,” in Proc. of the Intl. Society for Music Information Conf., Kobe, Japan, 2009.
3. R. Panda, R. Malheiro and R. P. Paiva, "Novel Audio Features for Music Emotion Recognition," in IEEE Transactions on Affective Computing, vol. 11, no. 4, pp. 614-626, 1 Oct.-Dec. 2020, doi: 10.1109/TAFFC.2018.2820691.
4. Rajesh, Sangeetha, and N. J. Nalini. "Musical instrument emotion recognition using deep recurrent neural network." *Procedia Computer Science* 167 (2020): 16-25, doi: 10.1016/J.PROCS.2020.03.178
5. Hizlisoy, Serhat, Serdar Yildirim, and Zekeriya Tufekci. "Music emotion recognition using convolutional long short term memory deep neural networks." *Engineering Science and Technology, an International Journal* 24.3 (2021): 760-767, doi: 10.1016/J.JESTCH.2020.10.009
6. Yang, Yi-Hsuan, and Homer H. Chen. Music emotion recognition. CRC Press, 2011.
7. Panda, R., Redinho, H., Gonçalves, C., Malheiro, R. and Paiva, R.P., 2021, July. How does the spotify api compare to the music emotion recognition state-of-the-art?. In 18th Sound and Music Computing Conference (SMC 2021) (pp. 238-245). Axea sas/SMC Network.
8. Amsterdam, Noah. "Analyzing popular music using Spotify’s Machine Learning Audio Features." (2019).
9. Álvarez, P., J. García de Quirós, and S. Baldassarri. "RIADA: A Machine-Learning Based Infrastructure for Recognising the Emotions of Spotify Songs." (2023). <https://doi.org/10.9781/ijimai.2022.04.002>
10. Liu, Tong, Li Han, Liangkai Ma, and Dongwei Guo. "Audio-based deep music emotion recognition." In AIP Conference Proceedings, vol. 1967, no. 1. AIP Publishing, 2018.
11. Laurier, Cyril, Perfecto Herrera, M. Mandel, and D. Ellis. "Audio music mood classification using support vector machine." MIREX task on Audio Mood Classification (2007): 2-4.
12. Malheiro, Ricardo, Renato Panda, Paulo Gomes, and Rui Pedro Paiva. "Emotionally-relevant features for classification and regression of music lyrics." IEEE Transactions on Affective Computing 9, no. 2 (2016): 240-254.
13. Zhang, Fan, Hongying Meng, and Maozhen Li. "Emotion extraction and recognition from music." In 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), pp. 1728-1733. IEEE, 2016.
14. Lartillot, Olivier, and Petri Toiviainen. "A Matlab toolbox for musical feature extraction from audio." In International conference on digital audio effects, vol. 237, p. 244. 2007.
15. Thayer, Robert E. The biopsychology of mood and arousal. Oxford University Press, 1990.
16. McFee, Brian, Colin Raffel, Dawen Liang, Daniel P. Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. "librosa: Audio and music signal analysis in python." In Proceedings of the 14th python in science conference, vol. 8, pp. 18-25. 2015.