| PaperID | SessionID         | Title                                                                                                                                      |
|---------|-------------------|--------------------------------------------------------------------------------------------------------------------------------------------|
| 103     | Mon5th10AM(IST)   | Interpreting Song Lyrics with an Audio-Informed Pre-trained Language Model                                                                 |
| 46      | Mon5th10AM(IST)   | Toward postprocessing-free neural networks for joint beat and downbeat estimation                                                          |
| 37      | Mon5th10AM(IST)   | Music Translation: Generating Piano Arrangements in Different Playing Levels                                                               |
| 287     | Mon5th10AM(IST)   | Scaling Polyphonic Transcription with Mixtures of Monophonic Transcriptions                                                                |
| 100     | Mon5th10AM(IST)   | Attention-based audio embeddings for query-by-example                                                                                      |
| 61      | Mon5th10AM(IST)   | SIATEC-C: Computationally efficient repeated pattern discovery in polyphonic music                                                         |
| 109     | Mon5th10AM(IST)   | Tailed U-Net: Multi-Scale Music Representation Learning                                                                                    |
| 85      | Mon5th10AM(IST)   | DDSP-based Singing Vocoders: A New Subtractive-based Synthesizer and A Comprehensive Evaluation                                            |
| 117     | Mon5th10AM(IST)   | Equivariant self-supervision for musical tempo estimation                                                                                  |
| 90      | Mon5th10AM(IST)   | How Music features and Musical Data Representations Affect Objective Evaluation of Music Composition: A Review of CSMT Data Challenge 2020 |
| 153     | Mon5th10AM(IST)   | YM2413-MDB: A Multi-Instrumental FM Video Game Music Dataset with Emotion Annotations                                                      |
| 133     | Mon5th10AM(IST)   | Detecting Symmetries of All Cardinalities With Application to Musical 12-Tone Rows                                                         |
| 234     | Mon5th10AM(IST)   | The power of deep without going deep? A study of HDPGMM music representation learning                                                      |
| 322     | Mon5th10AM(IST)   | Pop Music Generation with Controllable Phrase Lengths                                                                                      |
|         | Mon5th10AM(IST)   | Exploiting Pre-trained Feature Networks for Generative Adversarial Networks in Audio-domain Loop Generation                                |
| 306     | Mon5th10AM(IST)   | Modeling the rhythm from lyrics for melody generation of pop songs                                                                         |
|         | Mon5th1.30PM(IST) | Visualization for Al-Assisted Composing                                                                                                    |
|         | Mon5th1.30PM(IST) | Retrieving musical information from neural data: how cognitive features enrich acoustic ones                                               |
|         | Mon5th1.30PM(IST) | Beat Transformer: Demixed Beat and Downbeat Tracking with Dilated Self-Attention                                                           |
|         | Mon5th1.30PM(IST) | Sketching the Expression: Flexible Rendering of Expressive Piano Performance with Self-Supervised Learning                                 |
|         | Mon5th1.30PM(IST) | Exploiting Device and Audio Data to Tag Music with User-Aware Listening Contexts                                                           |
|         | Mon5th1.30PM(IST) | Jukedrummer: Conditional Beat-aware Audio-domain Drum Accompaniment Generation via Transformer VQ-VAE                                      |
|         | Mon5th1.30PM(IST) | Learning Hierarchical Metrical Structure Beyond Measures                                                                                   |
|         | Mon5th1.30PM(IST) | Mid-level Harmonic Audio Features for Musical Style Classification                                                                         |
|         | Mon5th1.30PM(IST) | Distortion Audio Effects: Learning How to Recover the Clean Signal                                                                         |
|         | Mon5th1.30PM(IST) | End-to-End Full-Page Optical Music Recognition for Mensural Notation                                                                       |
|         | Mon5th1.30PM(IST) | Mel Spectrogram Inversion with Stable Pitch                                                                                                |
|         | Mon5th1.30PM(IST) | Latent feature augmentation for chorus detection                                                                                           |
|         | Mon5th1.30PM(IST) | AccoMontage2: A Complete Harmonization and Accompaniment Arrangement System                                                                |
|         | Mon5th1.30PM(IST) | Supervised and Unsupervised Learning of Audio Representations for Music Understanding                                                      |
|         | Mon5th1.30PM(IST) | Generating Coherent Drum Accompaniment with Fills and Improvisations                                                                       |
|         | Mon5th1.30PM(IST) | Bottlenecks and solutions for audio to score alignment research                                                                            |
|         | Tue6th10AM(IST)   | Raga Classification From Vocal Performances Using Multimodal Analysis                                                                      |
|         | Tue6th10AM(IST)   | Traces of Globalization in Online Music Consumption Patterns and Results of Recommendation Algorithms                                      |
|         | Tue6th10AM(IST)   | Network Analyses for Cross-Cultural Music Popularity                                                                                       |
|         | Tue6th10AM(IST)   | Three related corpora in Middle Byzantine music notation and a preliminary comparative analysis                                            |
|         | Tue6th10AM(IST)   | Playing Technique Detection by Fusing Note Onset Information in Guzheng Performance                                                        |
| 111     | Tue6th10AM(IST)   | KDC: an open corpus for computational research of dastgāhi music                                                                           |

Page 1

| PaperID | SessionID         | Title                                                                                             |
|---------|-------------------|---------------------------------------------------------------------------------------------------|
| 148     | Tue6th10AM(IST)   | Inaccurate Prediction or Genre Evolution? Rethinking Genre Classification                         |
| 334     | Tue6th10AM(IST)   | In Search of Sañcāras: Tradition-informed Repeated Melodic Pattern Recognition in Carnatic Music  |
| 63      | Tue6th10AM(IST)   | Automatic Chinese National Pentatonic Modes Recognition Using Convolutional Neural Network        |
| 81      | Tue6th10AM(IST)   | Teach Yourself Georgian Dataset: A Corpus Of Traditional {\em A Cappella} Vocal Polyphony         |
| 272     | Tue6th10AM(IST)   | Adapting meter tracking models to Latin American music                                            |
| 235     | Tue6th10AM(IST)   | Critiquing Task- versus Goal-oriented Approaches: A Case for Makam Recognition                    |
| 245     | Tue6th10AM(IST)   | A Dataset for Greek Traditional and Folk Music: Lyra                                              |
| 296     | Tue6th10AM(IST)   | Analysis and detection of singing techniques in repertoires of J-POP solo singers                 |
| 335     | Tue6th1.30PM(IST) | Performance MIDI-to-score conversion by neural beat tracking                                      |
| 66      | Tue6th1.30PM(IST) | Symbolic Music Loop Generation with Neural Discrete Representations                               |
| 11      | Tue6th1.30PM(IST) | Automatic music mixing with deep learning and out-of-domain data                                  |
| 139     | Tue6th1.30PM(IST) | Music-STAR: a Style Translation system for Audio-based Re-instrumentation                         |
| 29      | Tue6th1.30PM(IST) | Learning Unsupervised Hierarchies of Audio Concepts                                               |
| 240     | Tue6th1.30PM(IST) | Multi-objective Hyper-parameter Optimization of Behavioral Song Embeddings                        |
| 70      | Tue6th1.30PM(IST) | ATEPP: A Dataset of Automatically Transcribed Expressive Piano Performance                        |
| 180     | Tue6th1.30PM(IST) | PDAugment: Data Augmentation by Pitch and Duration Adjustments for Automatic Lyrics Transcription |
| 107     | Tue6th1.30PM(IST) | Parameter Sensitivity of Deep-Feature based Evaluation Metrics for Audio Textures                 |
|         | Tue6th1.30PM(IST) | Stability of Symbolic Feature Group Importance in the Context of Multi-Modal Music Classification |
|         | Tue6th1.30PM(IST) | Multi-pitch Estimation meets Microphone Mismatch: Applicability of Domain Adaptation              |
|         | Tue6th1.30PM(IST) | Melody transcription via generative pre-training                                                  |
|         | Tue6th1.30PM(IST) | Source Separation of Piano Concertos with Test-Time Adaptation                                    |
|         | Tue6th1.30PM(IST) | Counterpoint Error-Detection Tools for Optical Music Recognition of Renaissance Polyphonic Music  |
|         | Tue6th1.30PM(IST) | A Dataset of Symbolic Texture Annotations in Mozart Piano Sonatas                                 |
|         | Tue6th1.30PM(IST) | Violin Etudes: A Comprehensive Dataset for f0 Estimation and Performance Analysis                 |
|         | Tue6th1.30PM(IST) | Checklist Models for Improved Output Fluency in Piano Fingering Prediction                        |
|         | Wed7th10AM(IST)   | And what if two musical versions don't share melody, harmony, rhythm, or lyrics?                  |
|         | Wed7th10AM(IST)   | Musika! Fast Infinite Waveform Music Generation                                                   |
|         | Wed7th10AM(IST)   | Symphony Generation with Permutation Invariant Language Model                                     |
|         | Wed7th10AM(IST)   | MuLan: A Joint Embedding of Music Audio and Natural Language                                      |
|         | Wed7th10AM(IST)   | MeloForm: Generating Melody with Musical Form based on Expert Systems and Neural Networks         |
|         | Wed7th10AM(IST)   | Towards robust music source separation on loud commercial music                                   |
|         | Wed7th10AM(IST)   | Towards Quantifying the Strength of Music Scenes Using Live Event Data                            |
|         | Wed7th10AM(IST)   | Learning Multi-Level Representations for Hierarchical Music Structure Analysis.                   |
|         | Wed7th10AM(IST)   | Multi-instrument Music Synthesis with Spectrogram Diffusion                                       |
|         | Wed7th10AM(IST)   | DDX7: Differentiable FM Synthesis of Musical Instrument Sounds                                    |
|         | Wed7th10AM(IST)   | Singing beat tracking with Self-supervised front-end and linear transformers                      |
|         | Wed7th10AM(IST)   | EnsembleSet: a new high quality dataset for chamber ensemble separation                           |
| 6       | Wed7th10AM(IST)   | End-to-End Lyrics Transcription Informed by Pitch and Onset Estimation                            |

Page 2

| PaperID | SessionID         | Title                                                                                                      |
|---------|-------------------|------------------------------------------------------------------------------------------------------------|
| 275     | Wed7th10AM(IST)   | Contrastive Audio-Language Learning for Music                                                              |
| 286     | Wed7th10AM(IST)   | MusAV: A dataset of relative arousal-valence annotations for validation of audio models                    |
| 136     | Wed7th10AM(IST)   | What is missing in deep music generation? A study of repetition and structure in popular music             |
| 282     | Wed7th10AM(IST)   | Heterogeneous Graph Neural Network for Music Emotion Recognition                                           |
| 182     | Wed7th1.30PM(IST) | Sonus Texere! Automated Dense Soundtrack Construction for Books using Movie Adaptations                    |
| 262     | Wed7th1.30PM(IST) | A diffusion-inspired training strategy for singing voice extraction in the waveform domain                 |
| 23      | Wed7th1.30PM(IST) | A Model You Can Hear: Audio Identification with Playable Prototypes                                        |
| 155     | Wed7th1.30PM(IST) | An Exploration of Generating Sheet Music Images                                                            |
| 32      | Wed7th1.30PM(IST) | HPPNet: Modeling the Harmonic Structure and Pitch Invariance in Piano Transcription                        |
|         | Wed7th1.30PM(IST) | Generating music with sentiment using Transformer-GANs                                                     |
|         | Wed7th1.30PM(IST) | Improving Choral Music Separation through Expressive Synthesized Data from Sampled Instruments             |
|         | Wed7th1.30PM(IST) | Ethics of Singing Voice Synthesis: Perceptions of Users and Developers                                     |
|         | Wed7th1.30PM(IST) | Emotion-driven Harmonisation And Tempo Arrangement of Melodies Using Transfer Learning                     |
|         | Wed7th1.30PM(IST) | Using Activation Functions for Improving Measure-Level Audio Synchronization                               |
|         | Wed7th1.30PM(IST) | A deep learning method for melody extraction from a polyphonic symbolic music representation               |
|         | Wed7th1.30PM(IST) | A Reproducibility Study on User-centric MIR Research and Why it is Important                               |
|         | Wed7th1.30PM(IST) | Music Separation Enhancement with Generative Modeling                                                      |
|         | Wed7th1.30PM(IST) | SampleMatch: Drum Sample Retrieval by Musical Context                                                      |
|         | Wed7th1.30PM(IST) | A Transformer-Based "Spellchecker" for Detecting Errors in OMR Output                                      |
|         | Wed7th1.30PM(IST) | "More than words": Linking Music Preferences and Moral Values through Lyrics                               |
|         | Thu8th10AM(IST)   | A unified model for zero-shot singing voice conversion and synthesis                                       |
|         | Thu8th10AM(IST)   | Semantic Control of Generative Musical Attributes                                                          |
|         | Thu8th10AM(IST)   | Music Representation Learning Based on Editorial Metadata from Discogs                                     |
|         | Thu8th10AM(IST)   | Melody Infilling with User-Provided Structural Context                                                     |
|         | Thu8th10AM(IST)   | Robust Melody Track Identification in Symbolic Music                                                       |
|         | Thu8th10AM(IST)   | Tracking the Evolution of a Band's Live Performances over Decades                                          |
|         | Thu8th10AM(IST)   | Evaluating Generative Audio Systems and Their Metrics                                                      |
|         | Thu8th10AM(IST)   | Representation Learning for the Automatic Indexing of Sound Effects Libraries                              |
|         | Thu8th10AM(IST)   | Concept-Based Techniques for Musicologist-Friendly Explanations in Deep Music Classifiers                  |
|         | Thu8th10AM(IST)   | Verse versus Chorus: Structure-aware Feature Extraction for Lyrics-based Genre Recognition                 |
|         | Thu8th10AM(IST)   | Transfer Learning of wav2vec 2.0 for Automatic Lyric Transcription                                         |
|         | Thu8th10AM(IST)   | A Novel Dataset and Deep Learning Benchmark for Classical Music Form Recognition and Analysis              |
|         | Thu8th10AM(IST)   | BAF: An audio fingerprinting dataset for broadcast monitoring                                              |
|         | Thu8th10AM(IST)   | Cadence Detection in Symbolic Classical Music using Graph Neural Networks.                                 |
|         | Thu8th10AM(IST)   | Domain Adversarial Training on Conditional Variational Auto-Encoder for Controllable Music Generation      |
|         | Thu8th10AM(IST)   | Modeling perceptual loudness of piano tone: theory and applications                                        |
| 317     | Thu8th10AM(IST)   | On the Impact and Interplay of Input Representations and Network Architectures for Automatic Music Tagging |

Page 3